



U.S. Food and Drug Administration
Protecting and Promoting Public Health

www.fda.gov

Substance Identifiers in Regulatory Practice

Frank Switzer, PhD
US Food & Drug Administration
frank,switzer@fda.hhs.gov



The views expressed in this presentation are those of the author and do not reflect the official policy or position of the Department of Health & Human Services, US Food & Drug Administration or the US Government



Why ISO 11238 and DTS 19844?

- Substances form the essence or basis of every product
- Interactions between substances are responsible for nearly all pharmacological activity.
- Substances are a lynchpin on which to organize regulatory information but many regulatory agencies only identify/define them by names or codes
- Names and codes are not sufficient to manage substance information.
- Regulatory agencies need to be able to define substances scientifically in an unambiguous manner
- ISO 11238 provides a framework for defining all substances and related regulatory information
- DTS 19844 provides many detailed examples of which data is collected according to the standard.



Names

- Often Ambiguous
 - Different meanings in different domains
 - Lime
 - Different meanings in different jurisdictions
 - Amoxicillin



PDF's and Package Inserts

- Paper or Electronic Paper
- Information not accessible
 - Difficult to read
 - More difficult for computers to read
- Need for Structured Information



Approach to Substances

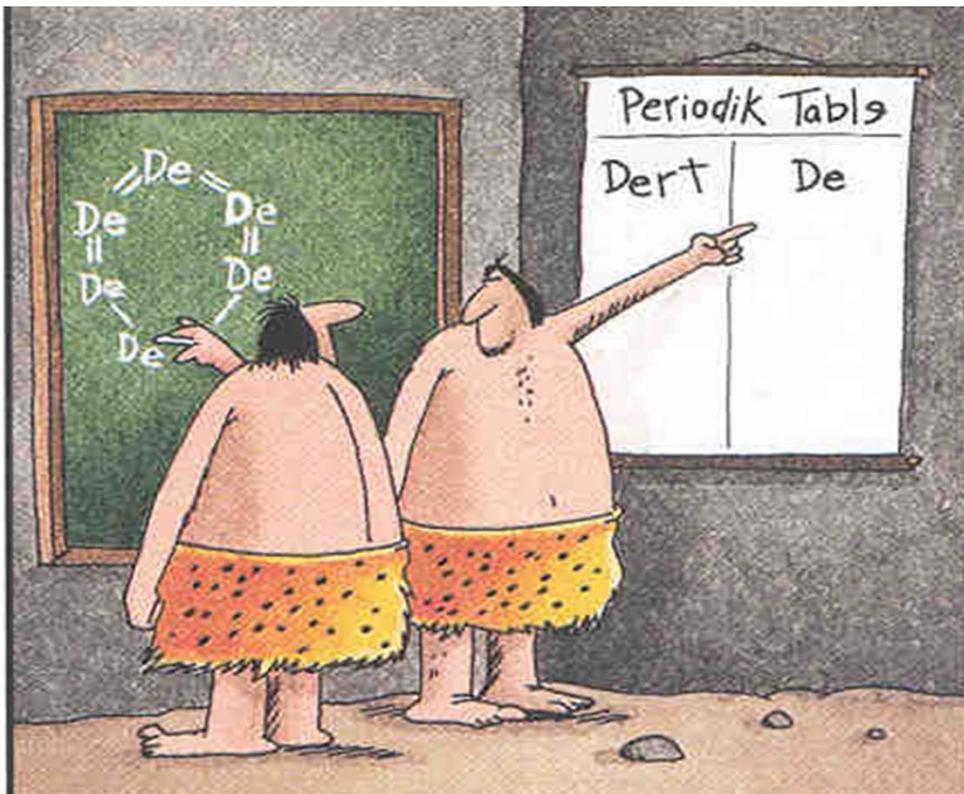
- Substances are defined based on **what they are** not how they are made or used
- Substances are defined independent of grade or level of purity



What is a Substance?

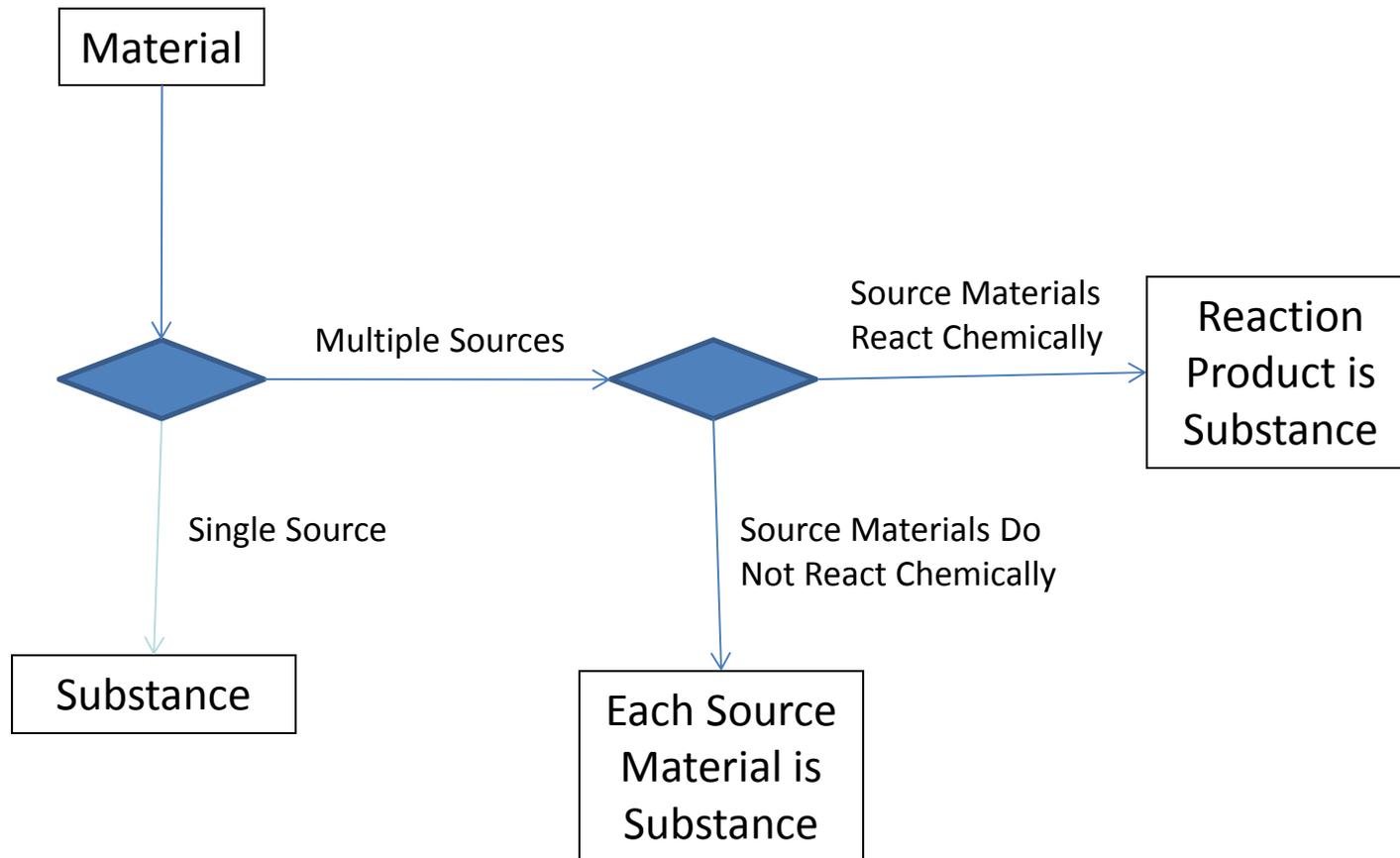
- **ARISTOTLE (Metaphysics)**...the generally recognizable substances... are the sensible substances, and sensible substances **all have matter**..., and in another sense the formula or form..., and thirdly the complex of matter and form, which alone is generated and destroyed, and is, without qualification, **capable of separate existence**

What is a Substance?



**Early Chemists
describe the first
DIRT MOLECULE
(The Far Side by Gary
Larson)**

Is the Material (Ingredient) a Substance?

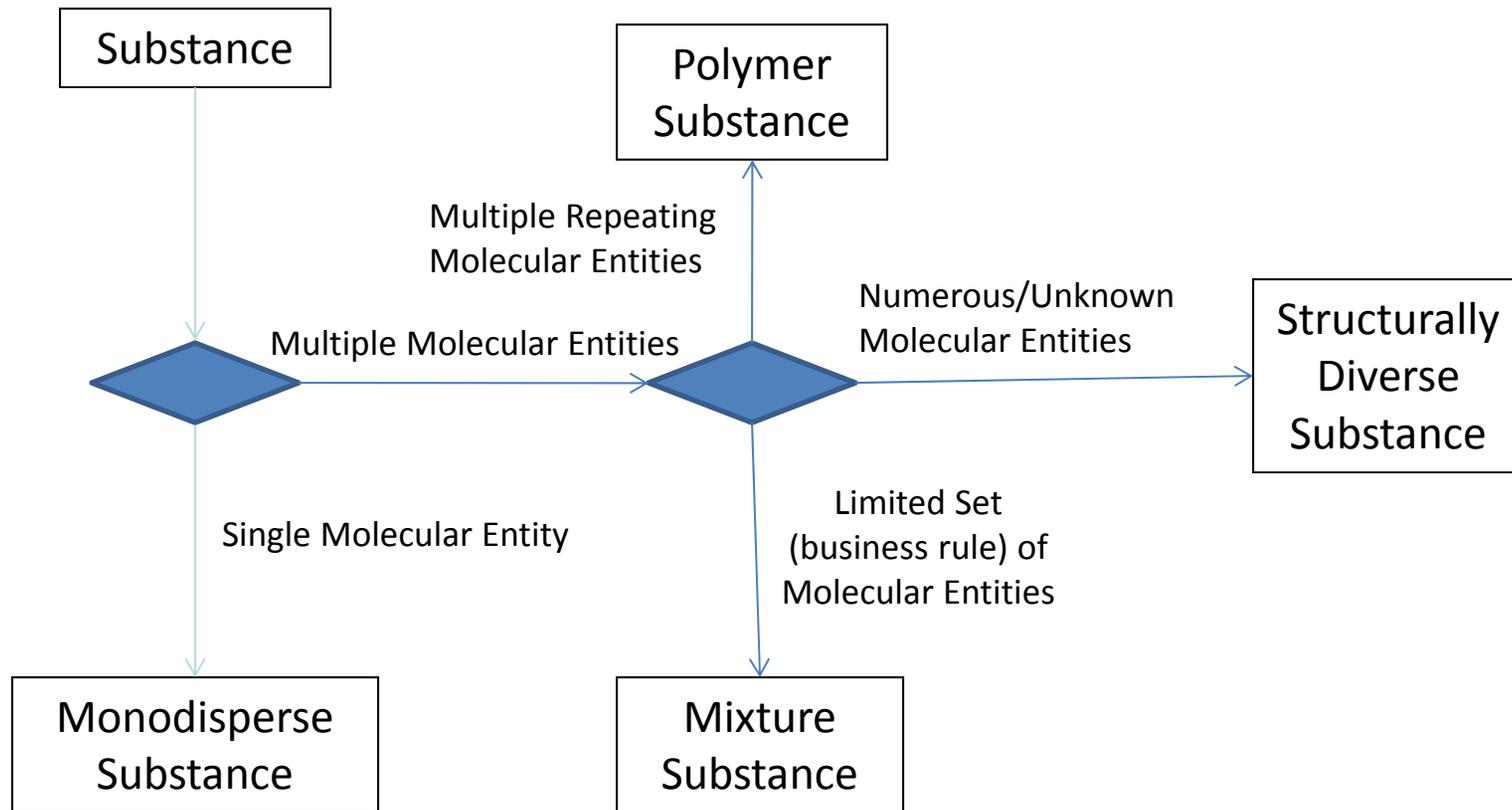




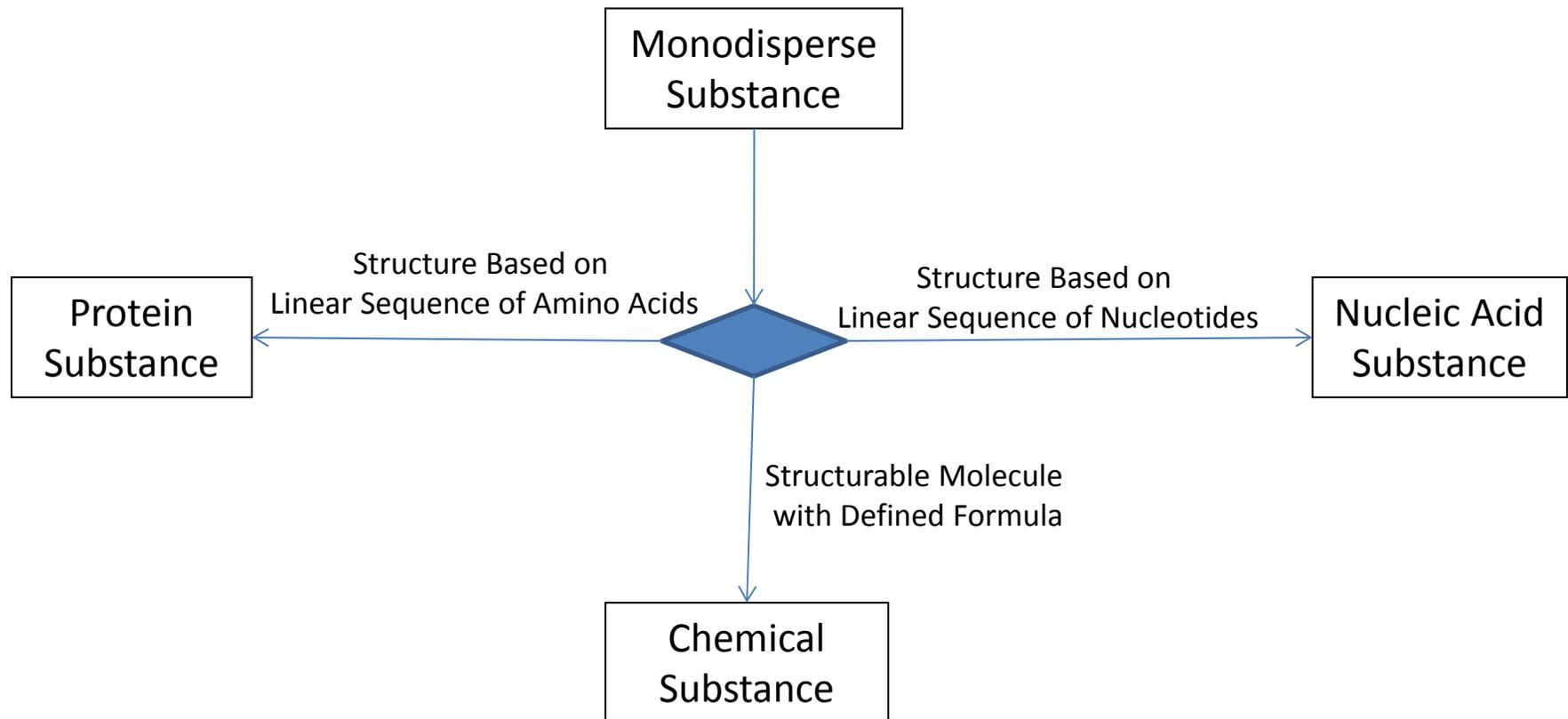
Substances

- Five groups of elements are used to describe single substances
 - **Monodisperse**
 - Chemicals
 - Proteins
 - Nucleic Acids
 - **Polydisperse**
 - Polymers (polysaccharides and synthetic polymers)
 - Structurally Diverse Substances
- Mixtures are comprised of combinations of single substances and source where relevant

Monodisperse, Polydisperse or Mixture Substance Type?



Which Monodisperse Substance Type?





Guiding Principles

- **Limited Ambiguity**
 - Uniqueness
 - Identity
 - Internal Consistency
 - Completeness
- **Confidentiality**
 - Single code to track an ingredient throughout the product lifecycle



Unique Ingredient Identifier

- The UNII consists of ten alphanumeric characters.
- Non-semantic non-chronological identifier
- The first nine alphanumeric characters are randomly generated.
- The tenth alphanumeric character is determined through a mathematical algorithm, and is appended to the first nine.
- $36^9 = 10^{14}$ potential identifiers
- **Over 70k public codes**



SRS @ NLM

Created for SPL Listing

- <http://fdasis.nlm.nih.gov/srs/srs.jsp>

The screenshot shows the Substance Registration System (SRS) interface. The top navigation bar includes the FDA logo and the text "U.S. Food and Drug Administration Protecting and Promoting Your Health". Below this, the page title is "Substance Registration System - Unique Ingr". The main content area shows a search bar with "FISH OIL" entered. Below the search bar, there is a "Go back to previous page." link. The search results section displays the following information:

Preferred Substance Name:	FISH OIL	[show more names]
UNII:	XGF7L72M0F	

Below the search results, there is a "Resources" section with a link to the "NCI Thesaurus".

On the right side, a browser window titled "Synonyms and Mappings - Mozilla Firefox" is open, displaying a list of synonyms for "FISH OIL":

- FISH OIL [INCI]
- FISH OIL [VANDF]
- FISH OIL [WHO-DD]
- FISH OILS
- OMEGAVEN



What About CAS RNs?

- 0 to many RNs for substances –**not an identity standard**
- CAS has no consistent way to capture polydispersity
- CAS RNs are copyrighted

InChI, InChIKey, SMILES, UUID (GUID)?



Data Elements and Standards for Drug Dictionaries

- [ICH M5](#) (2003-2013, Step 2 [published](#) 2005)
 - ISO IDMP (2005-present, 5 standards [published](#) 2012, implementation guides currently being balloted)
 - Substances – implemented as [GInAS](#)
 - Medicinal Products
 - Pharmaceutical Products – clinical drugs
 - Dosage Forms, Units of Presentation and Routes of Administration
 - Units of Measure ([UCUM](#))



Data Elements and Standards for Drug Dictionaries

- ▶ [SPL](#) and [UNII](#) developed in US through [HL7](#) (2005-present)
- ▶ 110k+ NDC's using about 10k UNII's
- ▶ Europe is now trying to catch up



DailyMed



I am a **Healthcare Professional** or from the **Pharmaceutical Industry**

I am a **Consumer or Patient**

ALL DRUGS

HUMAN DRUGS

ANIMAL DRUGS

SEARCH



Enter drug, NDC code, drug class, or Set ID



MORE WAYS TO SEARCH:

ADVANCED SEARCH

BROWSE DRUG CLASSES

LABEL ARCHIVES

TABLET/CAPSULE ID TOOL

This website contains **68147** drug listings as submitted to the **Food and Drug Administration (FDA)**.
At the present time, this Web site does not contain a complete listing of labels for approved prescription drugs.

SHARE

DailyMed

INDICATIONS AND USAGE

AMBIEN, a gamma-aminobutyric acid (GABA) A agonist, is indicated for the short-term treatment of insomnia characterized by difficulties with sleep initiation. AMBIEN has been shown to decrease sleep latency for up to 35 days in controlled clinical studies. [\(1\)](#)

DOSAGE AND ADMINISTRATION

- Use the lowest dose effective for the patient [\(2.1\)](#)
- Recommended initial dose is 5 mg for women and 5 or 10 mg for men, immediately before bedtime with at least 7–8 hours remaining before the planned time of awakening [\(2.1\)](#)
- Geriatric patients and patients with hepatic impairment: Recommended dose is 5 mg for men and women [\(2.2\)](#)
- Lower doses of CNS depressants may be necessary when taken concomitantly with AMBIEN [\(2.3\)](#)
- The effect of AMBIEN may be slowed if taken with or immediately after a meal [\(2.4\)](#)

DOSAGE FORMS AND STRENGTHS

5 mg and 10 mg tablets. Tablets not scored. [\(3\)](#)

AMBIEN

zolpidem tartrate tablet, film coated

PRODUCT INFORMATION

Product Type	HUMAN PRESCRIPTION DRUG LABEL	Item Code (Source)	NDC:0024-5401
Route of Administration	ORAL	DEA Schedule	CIV

ACTIVE INGREDIENT/ACTIVE MOIETY

Ingredient Name	Basis of Strength	Strength
zolpidem tartrate (zolpidem)	zolpidem tartrate	5 mg

INACTIVE INGREDIENTS

Ingredient Name	Strength
hypromelloses Imprecise Ingredient	
lactose	
magnesium stearate	Missing Strength
cellulose, microcrystalline	
polyethylene glycols Imprecise Ingredient	
titanium dioxide	
FD&C Red No. 40 Color formulation component	
polysorbate 80	



Why Aggregate/Curate?

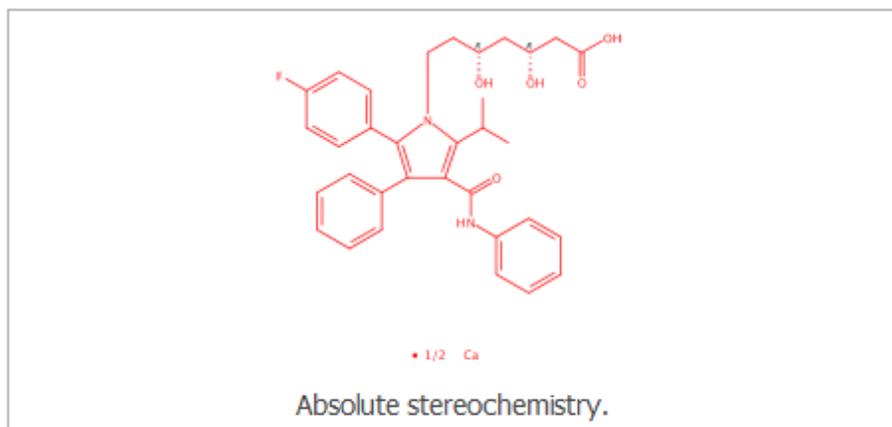
- Data sources can be incomplete/ambiguous/contradictory
- To provide a set of substance master data
- To facilitate interoperability
 - Richer data facilitates communication
 - Data must be useful both to humans and systems

What is the active ingredient in Lipitor?

1. **134523-03-8** 🔍

(Component: 134523-00-5)

~1625 📄 🧪 ~154 🧪



$C_{33}H_{35}FN_2O_5 \cdot \frac{1}{2}Ca$

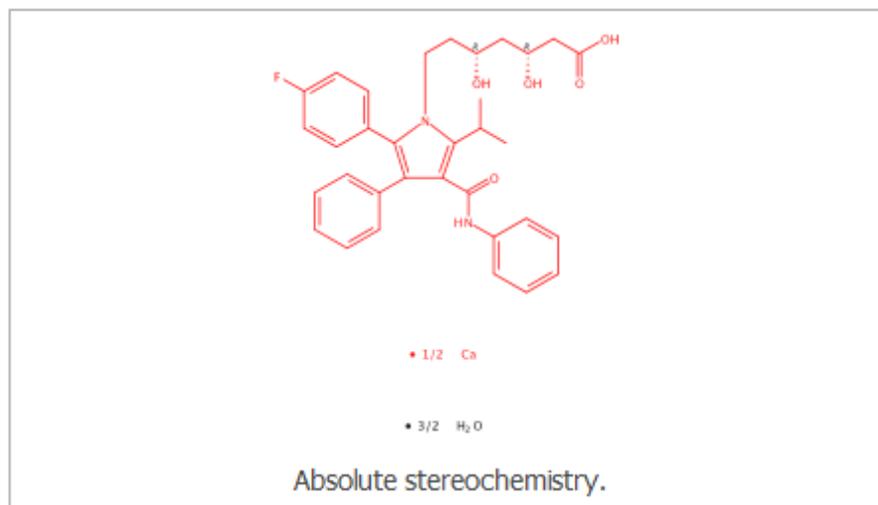
1*H*-Pyrrole-1-heptanoic acid, 2-(4-fluorophenyl)- β,δ -dihydroxy-5-(1-methylethyl)-3-phenyl-4-[(phenylamino)carbonyl]-, calcium salt (2:1), ($\beta R,\delta R$)-

[Regulatory Information](#)
[Experimental Properties](#)

2. **344423-98-9** 🔍

(Component: 134523-00-5)

~47 📄 🧪 ~27 🧪



$C_{33}H_{35}FN_2O_5 \cdot \frac{1}{2}Ca \cdot \frac{3}{2}H_2O$

1*H*-Pyrrole-1-heptanoic acid, 2-(4-fluorophenyl)- β,δ -dihydroxy-5-(1-methylethyl)-3-phenyl-4-[(phenylamino)carbonyl]-, calcium salt, hydrate (2:1:3), ($\beta R,\delta R$)-



Identity Problems

atorvastatin calcium

methylnaltrexone bromide

levothyroxine sodium

ferric citrate (USAN)

- Multiple established names for the same substance
- Missing links

Sitagliptin (INN)

sitagliptinum
sitagliptin

(3*R*)-3-amino-1-[3-(trifluoromethyl)-5,6,7,8-tetrahydro-5*H*-[1,2,4]-triazolo[4,3-*a*]pyrazin-7-yl]-4-(2,4,5-trifluorophenyl)butan-1-one
antidiabetic

sitagliptine

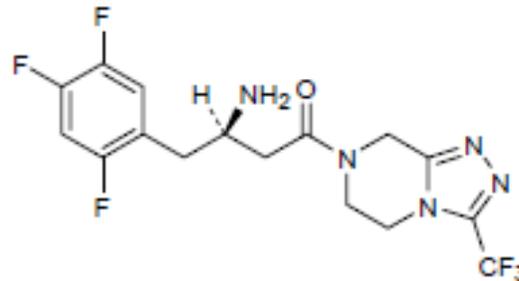
7-[(3*R*)-3-amino-4-(2,4,5-trifluorophényl)butanoyl]-3-(trifluorométhyl)-5,6,7,8-tétrahydro-1,2,4-triazolo[4,3-*a*]pyrazine
antidiabétique

sitagliptina

7-[(3*R*)-3-amino-4-(2,4,5-trifluorofenil)butanoil]-3-(trifluorometil)-5,6,7,8-tetrahidro-1,2,4-triazolo[4,3-*a*]pirazina
hipoglucemiante

C₁₅H₁₅F₆N₅O

486460-32-6



INN avoids salts, solvates AND Clinical Trial codes

Sitagliptin phosphate (USAN)

USAN

SITAGLIPTIN PHOSPHATE

PRONUNCIATION

sit a glip' tin

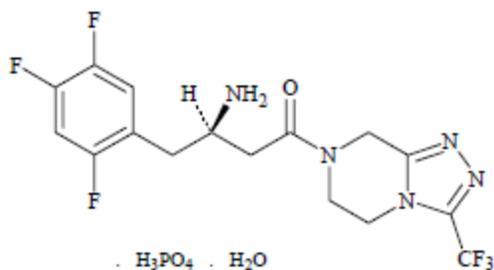
THERAPEUTIC CLAIM

Treatment of type 2 diabetes mellitus and related disorders

CHEMICAL NAMES

- 1,2,4-triazolo[4,3-a]pyrazine, 7-[(3*R*)-3-amino-1-oxo-4-(2,4,5-trifluorophenyl)butyl]-5,6,7,8-tetrahydro-3-(trifluoromethyl)-, phosphate (1:1) monohydrate
- 7-[(3*R*)-3-amino-4-(2,4,5-trifluorophenyl)butanoyl]-3-(trifluoromethyl)-5,6,7,8-tetrahydro-1,2,4-triazolo[4,3-a]pyrazine monophosphate monohydrate

STRUCTURAL FORMULA



MOLECULAR FORMULA

$\text{C}_{16}\text{H}_{17}\text{F}_6\text{N}_5\text{O} \cdot \text{H}_3\text{O}_4\text{P} \cdot \text{H}_2\text{O}$

MOLECULAR WEIGHT

523.3

TRADEMARK

Unknown as yet

MANUFACTURER

Merck & Co., Inc.

CAS REGISTRY NUMBER

654671-77-9

Sitagliptin hydrochloride (no official definition)

Get References Get Reactions Get Commercial Sources Tools Create Keep Me Posted Alert Send to SciPlanner

Sort by: CAS Registry Number Display Options

0 of 3 Substances Selected

1. **1000153-09-2**

(Component: 486460-32-6)

~12

• x HCl

Rotation (-), Absolute stereochemistry.

C₁₆ H₁₅ F₆ N₅ O · x Cl H
 1-Butanone, 3-amino-1-[5,6-dihydro-3-(trifluoromethyl)-1,2,4-triazolo[4,3-*a*]pyrazin-7(8*H*)-yl]-4-(2,4,5-trifluorophenyl)-, hydrochloride (1:?), (3*R*)-

2. **862156-92-1**

(Component: 486460-32-6)

~4

• HCl
• H₂O

Rotation (-), Absolute stereochemistry.

C₁₆ H₁₅ F₆ N₅ O · Cl H · H₂ O
 1-Butanone, 3-amino-1-[5,6-dihydro-3-(trifluoromethyl)-1,2,4-triazolo[4,3-*a*]pyrazin-7(8*H*)-yl]-4-(2,4,5-trifluorophenyl)-, hydrochloride, hydrate (1:1:1), (3*R*)-

3. **486459-71-6**

(Component: 486460-32-6)

~18

• HCl

Rotation (-), Absolute stereochemistry.

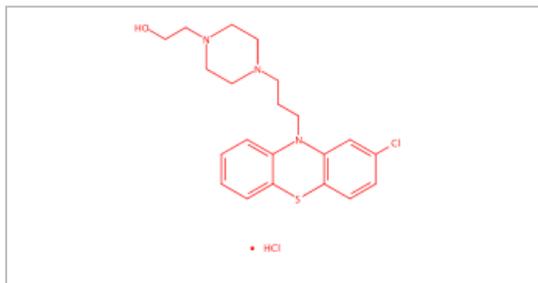
C₁₆ H₁₅ F₆ N₅ O · Cl H
 1-Butanone, 3-amino-1-[5,6-dihydro-3-(trifluoromethyl)-1,2,4-triazolo[4,3-*a*]pyrazin-7(8*H*)-yl]-4-(2,4,5-trifluorophenyl)-, hydrochloride (1:1), (3*R*)-

Perphenazine hydrochloride (no official definition)

7. 3111-71-5

(Component: 58-39-9)

~9



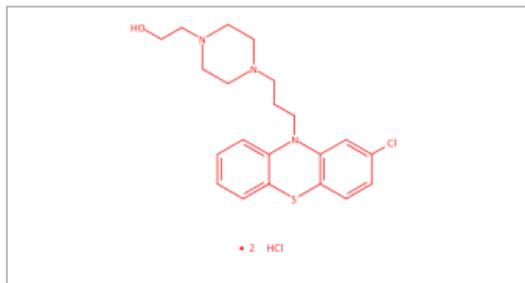
$C_{21}H_{26}ClN_3OS \cdot ClH$

1-Piperazineethanol, 4-[3-(2-chloro-10H-phenothiazin-10-yl)propyl]-, hydrochloride (1:1)

8. 2015-28-3

(Component: 58-39-9)

~39



$C_{21}H_{26}ClN_3OS \cdot 2ClH$

1-Piperazineethanol, 4-[3-(2-chloro-10H-phenothiazin-10-yl)propyl]-, hydrochloride (1:2)

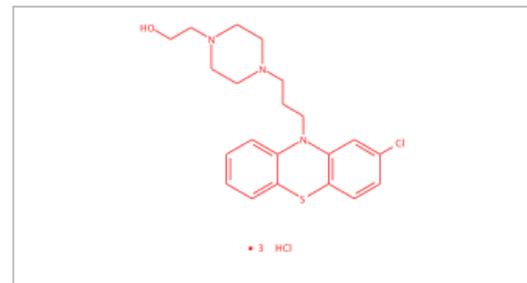
Key Physical Properties

Regulatory Information
Experimental Properties

9. 130-69-8

(Component: 58-39-9)

~7



$C_{21}H_{26}ClN_3OS \cdot 3ClH$

1-Piperazineethanol, 4-[3-(2-chloro-10H-phenothiazin-10-yl)propyl]-, hydrochloride (1:3)



(Polydisperse) Polymers

Hypromellose Example

USP 33: Hypromellose is a methyl and hydroxypropyl mixed ether of cellulose. It contains, calculated on the dried basis, methoxy ($-OCH_3$: 31.03) and hydroxypropoxy ($-OC_3H_6OH$: 75.09) groups conforming to the limits for the types of Hypromellose (hydroxypropyl methylcellulose) set forth in the accompanying table.

Substitution Type	Methoxy (percent)		Hydroxypropoxy (percent)	
	Min.	Max.	Min.	Max.
1828		16.5 20.0		23.0 32.0
2208		19.0 24.0		4.0 12.0
2906		27.0 30.0		4.0 7.5
2910		28.0 30.0		7.0 12.0



(Polydisperse) Polymers

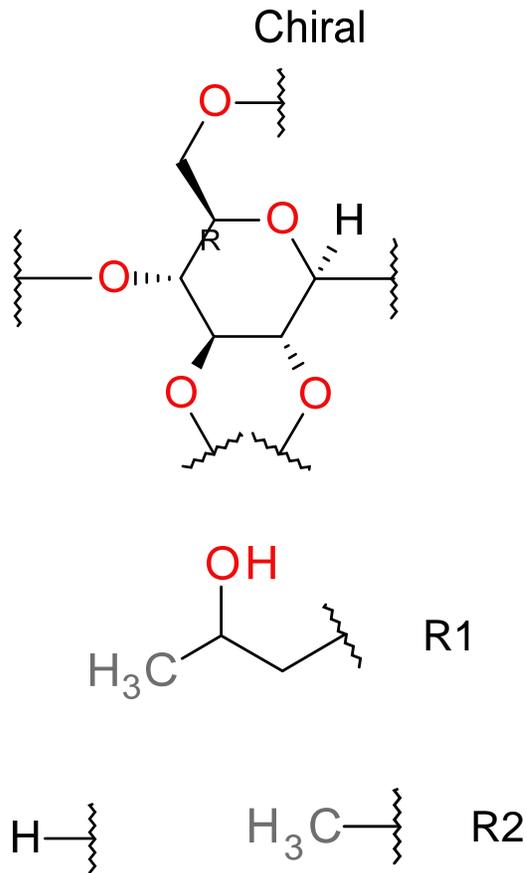
Hypromellose Example

USP 33: Labeling—Label it to indicate its substitution type and its nominal viscosity value in milli-Pascal per second (mPa·s).

Hypromellose is the INN and BAN

CAS 9004-65-3

Hypromellose



```

<POLYMER_TYPE>HOMOPOLYMER
<NUMBER_OF_SRU>1
<ORIENTATION_OF_POLYMERIZATION>HEAD-TAIL
<R_ID>R1
<LIMIT_TYPE>WEIGHT
<AVERAGE>10
<LOW_LIMIT>7
<HIGH_LIMIT>12
<R_ID>R2
<LIMIT_TYPE>WEIGHT
<AVERAGE>29
<LOW_LIMIT>28
<HIGH_LIMIT>30
<TYPE_MW>NUMBER
<MW_AVERAGE>8000
<LOW_LIMIT_MW/>
<HIGH_LIMIT_MW/>
<PHYSICAL_PROPERTY_TYPE>VISCOSITY
<AVERAGE>3
<LOW_LIMIT>2.4
<HIGH_LIMIT>3.6
<UNITS>MPA.S
    
```



Hypromello

HYPROMELLOSE 2910 (3 MPA.S)
HYPROMELLOSE 2910 (5 MPA.S)
HYPROMELLOSE 2910 (6 MPA.S)
HYPROMELLOSE 2910 (15 MPA.S)
HYPROMELLOSE 2910 (50 MPA.S)
HYPROMELLOSE 2910 (4000 MPA.S)
HYPROMELLOSE 2910 (15000 MPA.S)
HYPROMELLOSE 2906 (50 MPA.S)
HYPROMELLOSE 2906 (4000 MPA.S)
HYPROMELLOSE 2208 (3 MPA.S)
HYPROMELLOSE 2208 (100 MPA.S)
HYPROMELLOSE 2208 (4000 MPA.S)
HYPROMELLOSE 2208 (15000 MPA.S)
HYPROMELLOSE 2208 (100000 MPA.S)

0VUT3PMY82
R75537T0T4
0WZ8WG20P6
36SFW2JZ0W
1IVH67816N
RN3152OP35
288VBX44JC
612E703ZUQ
5EYA69XGAT
9H4L916OBU
B1QE5P712K
39J80LT57T
Z78RG6M2N2
VM7F0B23ZI

Fish oil

From Wikipedia, the free encyclopedia

Fish oil is oil derived from the tissues of oily fish. Fish oils contain the omega-3 fatty acids eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), precursors of certain eicosanoids that are known to reduce inflammation in the body,^{[1][2]} and have other health benefits.

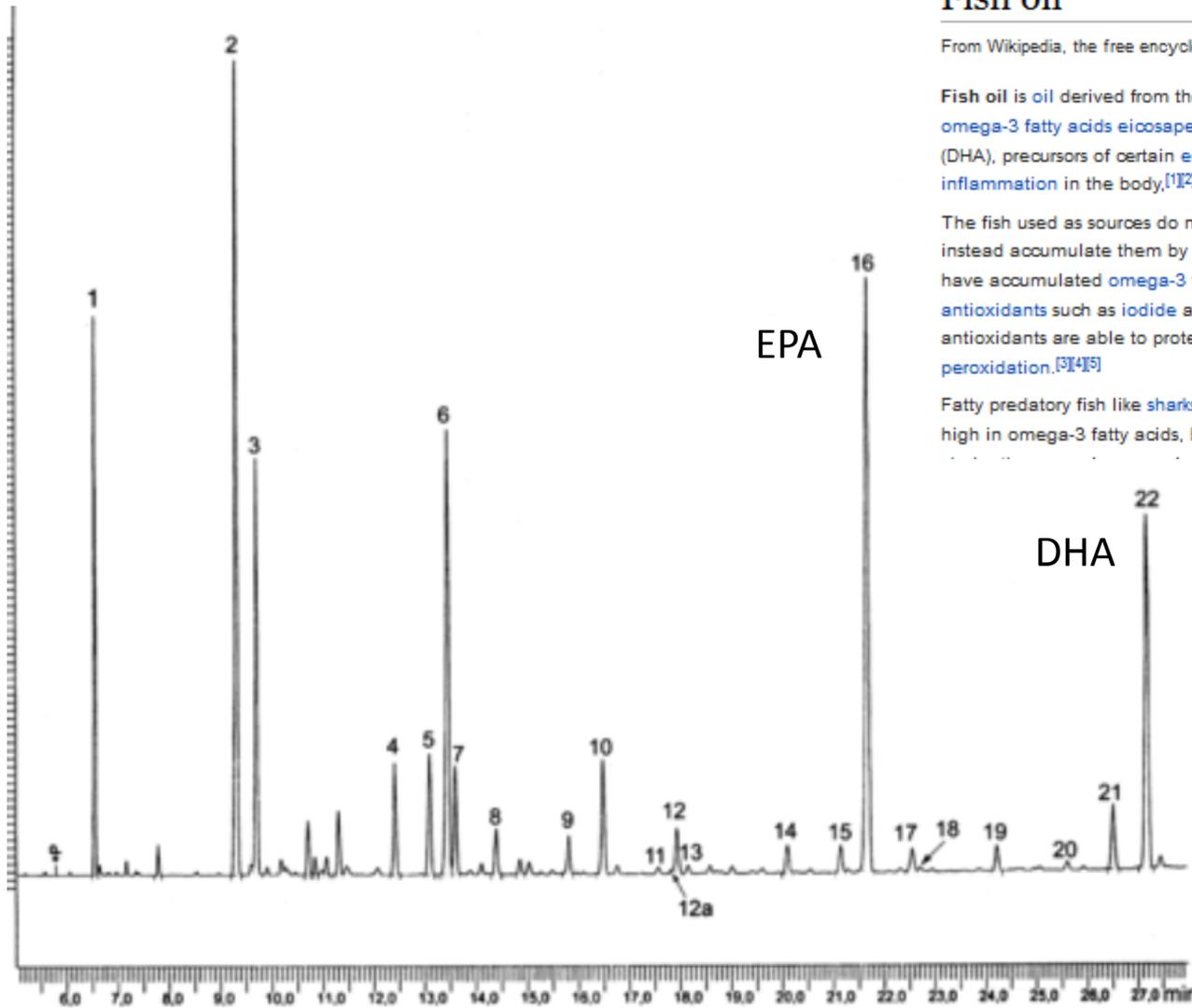
The fish used as sources do not actually produce omega-3 fatty acids, but instead accumulate them by consuming either microalgae or prey fish that have accumulated omega-3 fatty acids, together with a high quantity of antioxidants such as iodide and selenium, from microalgae, where these antioxidants are able to protect the fragile polyunsaturated lipids from peroxidation.^{[3][4][5]}

Fatty predatory fish like sharks, swordfish, tilefish, and albacore tuna may be high in omega-3 fatty acids, but due to their position at the top of the food



A typical fish oil softgel

Fish Oil Concentrate	2400 mg
Total Omega-3 Fatty Acids	720 mg
Omega-3 EPA (Eicosapentaenoic Acid)	360 mg
Omega-3 DHA (Docosahexaenoic Acid)	240 mg
Omega-3 Other	120 mg



1. C14:0	6. C18:1 n-9	11. C20:0	15. C20:4 n-3	20. C22:5 n-6
2. C16:0	7. C18:1 n-7	12. C20:1 n-9	16. C20:5 n-3	21. C22:5 n-3
3. C16:1 n-7	8. C18:2 n-6	12a. C20:1 n-11	17. C22:1 n-11	22. C22:6 n-3
4. C16:4 n-1	9. C18:3 n-3	13. C20:1 n-7	18. C22:1 n-9	
5. C18:0	10. C18:4 n-3	14. C20:4 n-6	19. C21:5 n-3	



Figure 1912-2. - Chromatogram for the assay of total omega-3 acids in fish oil rich in omega-3 acids

Garcinia Cambogia

WIKIPEDIA
The Free Encyclopedia

Wikipedia

Article Talk

Read View source View history

Search

Create account Log in

We need your opinion
How will this site thrive for the next billion internet users?

Garcinia gummi-gutta

From Wikipedia, the free encyclopedia

Gamboge redirects here; for the dye made from various tree species of the family Clusiaceae see Gamboge

Garcinia gummi-gutta is a tropical species of *Garcinia* native to Indonesia. Common names include *garcinia cambogia* (a former scientific name), as well as *brindieberry*,^[R] *Malabar tamarind*,^[R] and *kudam pull* (pot tamarind).^[R] This fruit looks like a small pumpkin and is green to pale yellow in color.

Although it has received considerable media attention purporting its effects on **weight loss**, no clinical evidence supports this claim.^[R]

Contents (hide)

- Cultivation
- Uses
 - Cooking
 - Weight loss
- References
- External resources

Cultivation

Garcinia gummi-gutta is grown for its fruit in Southeast Asia, coastal Karnataka/Kerala, India, and west and central Africa. It thrives in moist forests.

G. gummi-gutta is one of several closely related *Garcinia* species from the plant family Clusiaceae.^[R] With thin skin and deep vertical lobes, the fruit of *G. gummi-gutta* and related species range from about the size of an orange to that of a grapefruit; *G. gummi-gutta* looks more like a small yellowish, greenish, or sometimes reddish pumpkin.^[T] The color can vary considerably. When the rinds are dried and cured in preparation for storage and extraction, they are dark brown or black in color.

Along the west coast of South India, *G. gummi-gutta* is popularly termed "Malabar tamarind", and shares culinary uses with the tamarind (*Tamarindus indica*). The latter is a small and the former a quite large evergreen tree. *G. gummi-gutta* is also called *goreika* or, in some areas, simply *kratchie pull* (souring fruit). It is called *uppage* in Kannada language and fruits are collected and dried for selling to dealers in Birsil, Karnataka.^[R]

Uses

Cooking

G. gummi-gutta is used in cooking, including in the preparation of curries. The fruit rind and extracts of *Garcinia* species are called for in many traditional recipes,^[R] and various species of *Garcinia* are used similarly in food preparation in Assam (India), Thailand, Malaysia, Burma, and other Southeast Asian countries. In the Indian Ayurvedic medicine, "sour" flavors are said to activate digestion. The extract and rind of *G. gummi-gutta* is a curry condiment in India.^[citation needed]^[citation needed] It is an essential souring ingredient in the southern Thai variant of *kaeng som*, a sour curry.^[citation needed]

G. gummi-gutta is used commercially in fish curing, especially in Sri Lanka and South India.

The trees can be found in forested areas and also are protected in plantations otherwise given over to pepper, spice, and coffee production.

Weight loss

In late 2012, a United States television personality, Dr. Oz, promoted *garcinia cambogia* extract as "an exciting breakthrough in natural weight loss"^[R] Dr. Oz's previous endorsements have often led to a substantial increase in consumer interest in the promoted products. Scientific evidence is lacking and clinical trials do not support claims that *garcinia cambogia* is an effective weight-loss aid.^{[1][12][13][14]} A meta-analysis of several clinical trials found no compelling evidence for short-term weight loss.^[15] Further, **side effects** — namely **hepatotoxicity** (chemical-driven liver damage) — led to one preparation being withdrawn from the market.^{[16][17]}

A 1998 randomized, controlled trial looked at the effects of **hydroxylictric acid**, the purported active component in *G. gummi-gutta*, as a potential antiobesity agent in 136 people. The conclusion from this trial was that *garcinia cambogia* failed to produce significant weight loss and fat mass loss beyond that observed with placebo.^[18]

Garcinia gummi-gutta



Scientific classification

Kingdom: Plantae
(unranked): Angiosperms
(unranked): Eudicots
(unranked): Rosids
Order: Malpighiales
Family: Clusiaceae
Genus: *Garcinia*
Species: *G. gummi-gutta*

Binomial name

Garcinia gummi-gutta
(L.) Rob.

Synonyms^[T]

- Cambogia binucata* Walp.
- Cambogia gemmi-gutta* L.
- Cambogia solitaria* Stokes
- Garcinia affinis* Wight & Arn.
- Garcinia cambogia* (Gaertn.)

See:

- Garcinia sulcata* Stokes



Garcinia gummi-gutta tree in Kerala, India

Medicinal Plants Name Service

<http://apps.kew.org/mpns-portal/>

Accepted scientific name:

***Garcinia gummi-gutta* (L.) Roxb., Hort. Bengal.: 42 (1814).**

Taxonomic source: [World Checklist - unpublished](#)

Family: Clusiaceae

Non-scientific names and plant parts

Scientific synonyms

Published in medicinal plant references as

Further information

Non-scientific names for this plant and parts used:

Non-scientific name:	Class of name:	Trade forms:	Plant parts:	Medicinal plant reference:
brindall berry	Other			Herbs of Commerce (McGuffin et al., 2000)
garcinia	Other			Herbs of Commerce (McGuffin et al., 2000)
Garcinia cambogia	Other	dried pericarp of the fruits	fruit	U.S. Pharmacopoeia USP 37 (2013)
Malabar tamarind	Other			Herbs of Commerce (McGuffin et al., 2000)
Powdered Garcinia Hydroxycitrate Extract	Other	powdered extract	fruit	U.S. Pharmacopoeia USP 37 (2013)



Need for Specified Substance

- Need to tie material to a manufacturer and a process
- Need to tie material to a specific grade
- Need to obtain specification information
- Need to obtain information about processing materials
- Need to establish and monitor the supply chain



Specified Substance

- To be implemented with the GInAS Application
- An explicit grouping of elements and concepts put forward in ISO IDMP
 - Group-1 Multiple substance materials (Coatings, Colorants, Flavorants); Physical Form; Extracts
 - Group-2 Manufacturer and minimal manufacturing information
 - Group-3 Grade of material (USP, EP, technical, standardized etc.)
 - Group-4 Detailed manufacturing information, impurities, degradents etc.



Why IDMP?

- No other approach can meet the regulatory need for substance/ingredient identity codes
- Precise at the molecular level
- Consistent and standardized dataset
- UNIs are free to use, highly searchable and possess internal validation