FDA's GSRS System / Defining Excipients and the Inactive Ingredient Dictionary

Frank Switzer
FDA-SRS@fda.hhs.gov

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

Ground Rules

- Substances have matter and are capable of separate existence
- Substances are defined based on what they are, not how they are made or used

UNII Guiding Principles

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

Unique Ingredient Identifier

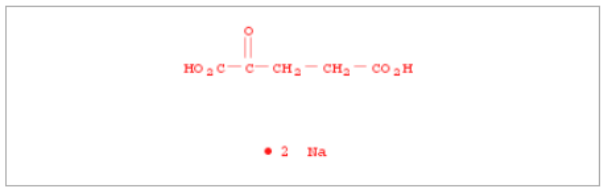
- The UNII consists of ten alphanumeric characters.
- Non-semantic non-chronological identifier
- The first nine alphanumeric characters are randomly generated.
- The tenth alphanumeric character is determined through a mathematical algorithm, and is appended to the first nine.
- $36^9 = 10^{13}$ potential identifiers
- **Nearly 100k public codes**

What About CAS RNs?

Sort by: Number of Commercial Sources (New) Display Options

0 of 14 Substances Selected

1. 305-72-6
(Component: 328-50-7)

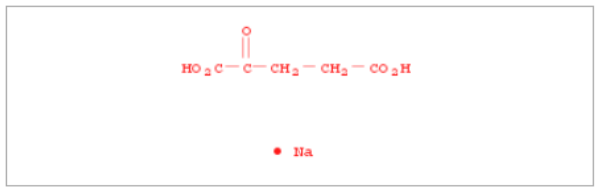


$C_5H_6O_5 \cdot 2\,Na$
Pentanedioic acid, 2-oxo-, sodium salt (1:2)

Regulatory Information
Spectra

DISODIUM OXOGLURATE
FLP7P4RM46

2. 22202-68-2
(Component: 328-50-7)

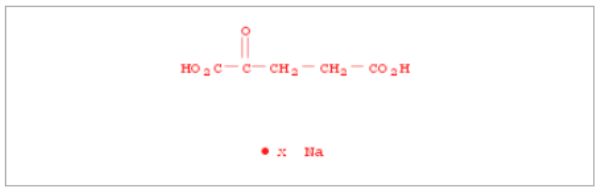


$C_5H_6O_5 \cdot Na$
Pentanedioic acid, 2-oxo-, sodium salt (1:1)

Regulatory Information
Spectra
Experimental Properties

MONOSODIUM OXOGLURATE
8GFV60F71R

3. 17091-15-5
(Component: 328-50-7)



$C_5H_6O_5 \cdot x\,Na$
Pentanedioic acid, 2-oxo-, sodium salt (1:?)

Regulatory Information

Ambiguous
NO UNII

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

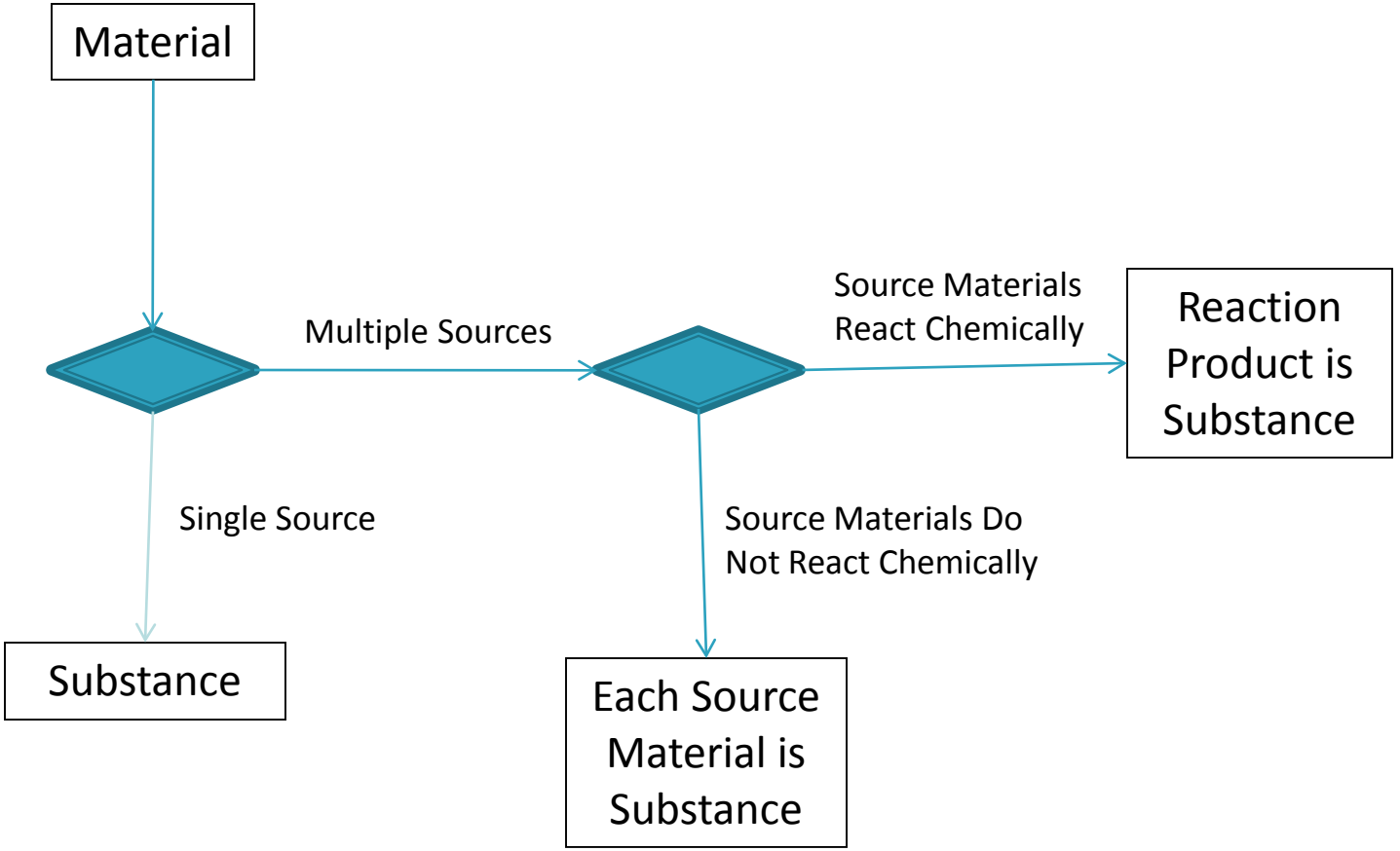
Why Aggregate/Curate?

- 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 Do GSRS Sources Provide?

Information Source	INN/USAN	CAS	Applications	Labeling	GSRS (UNII)
Focus	Name	Primary Literature	Product Quality	Use	Identity
Benefits	Required Standardized Name with description	Primary Literature Public Information	Info directly from Product Sponsor	Public domain Info	Clear Description Seeks Objective Truth
Limitations	Limited Scientific Analysis	Info may be Incomplete obiguous	Info is proprietary Info may not be available Info may be Incomplete / Ambiguous Data format problems	Errors may have propagated from previous sources	Single Standardized Description (GInAS will allow alternate descriptions and reference document storage)

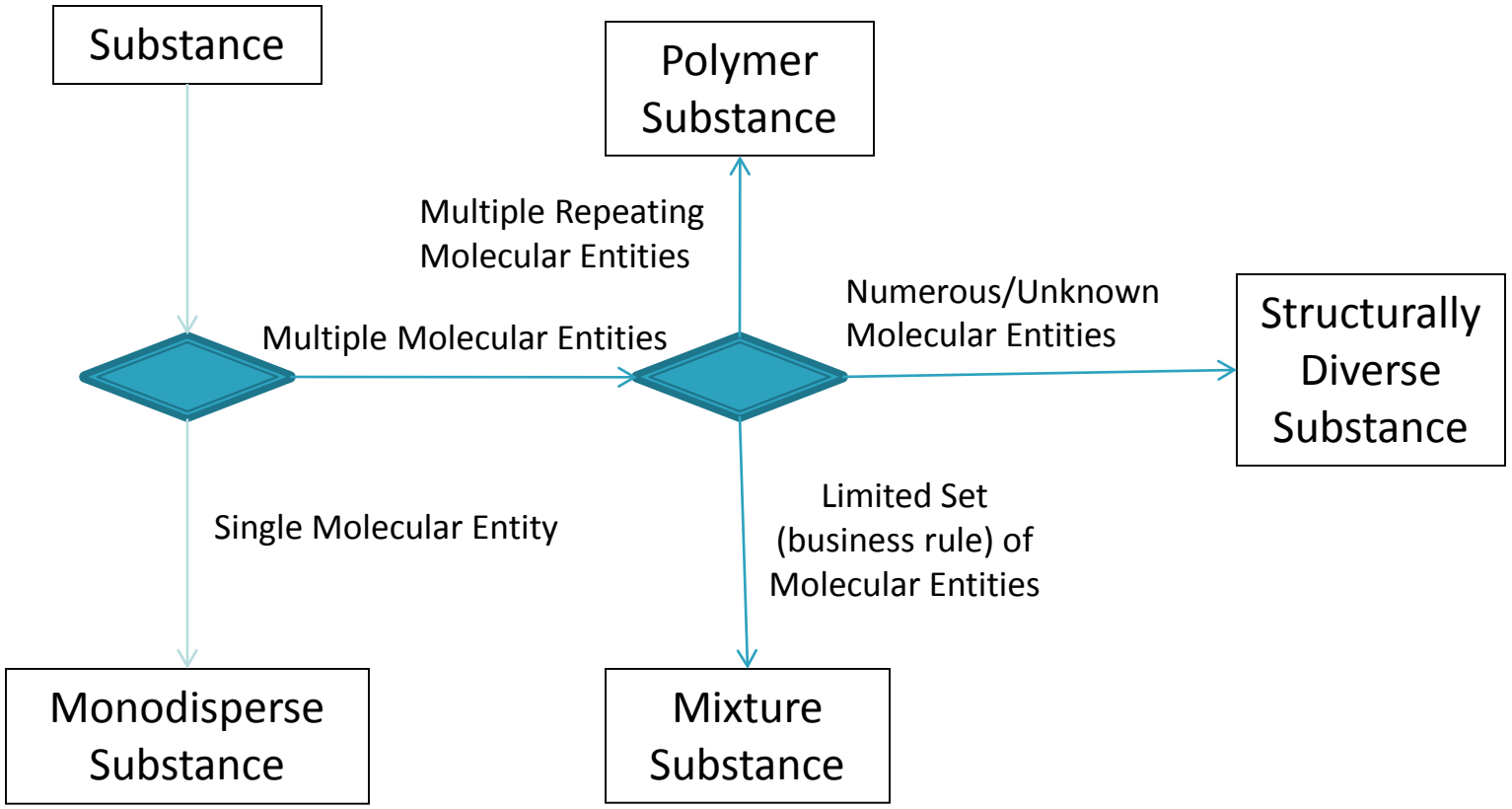
Is the Material (Ingredient) a Substance?



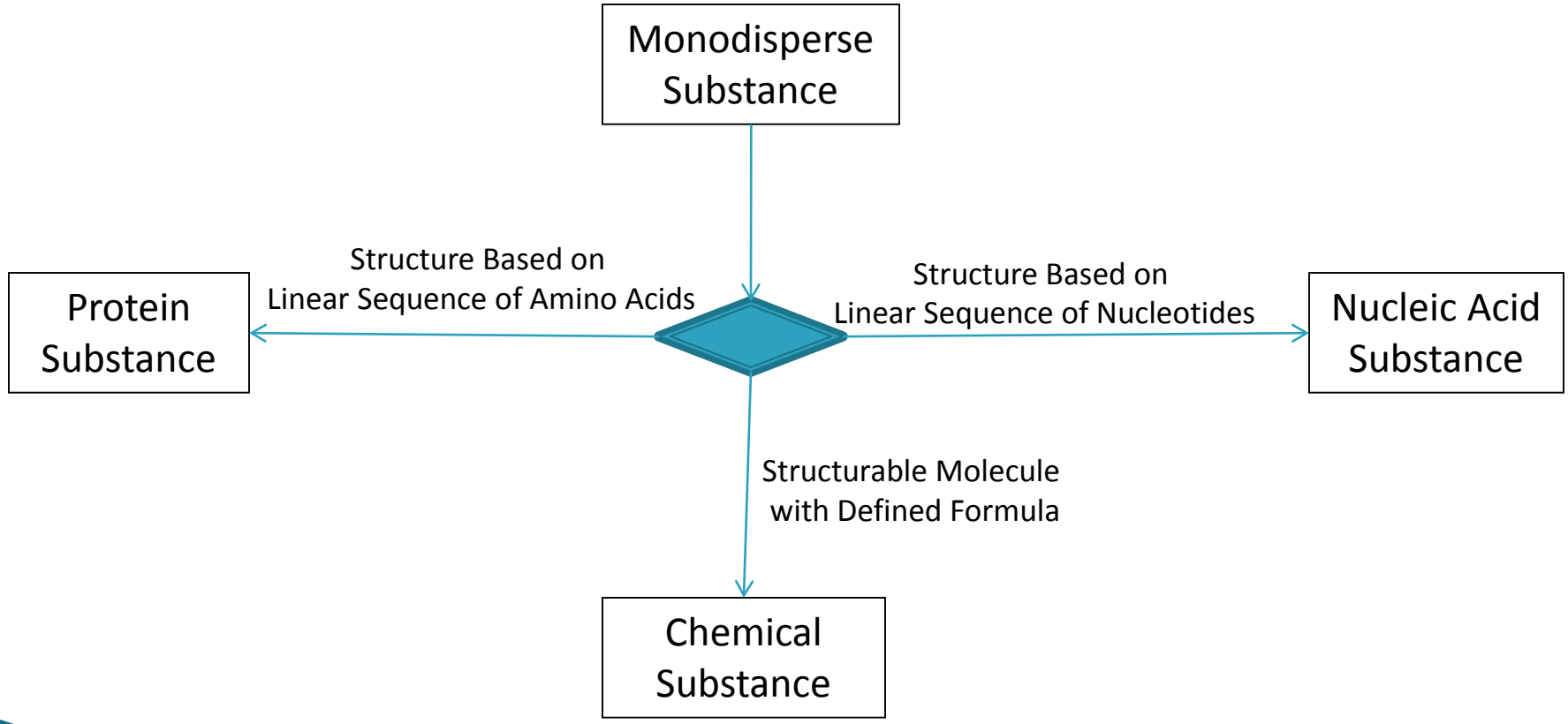
Substances

- 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?



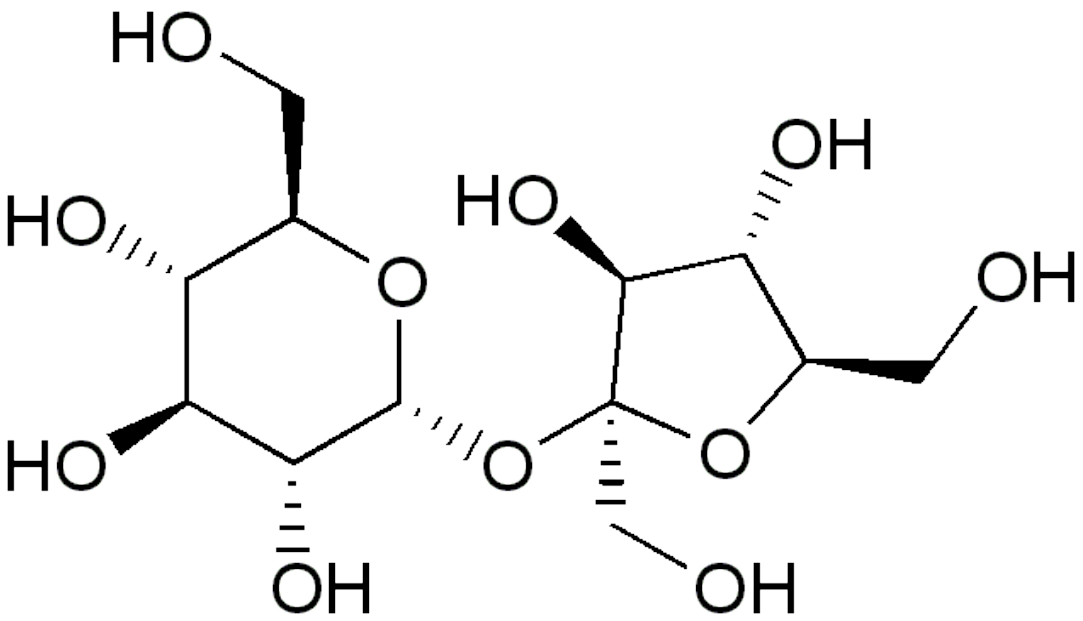
Which Monodisperse Substance Type?

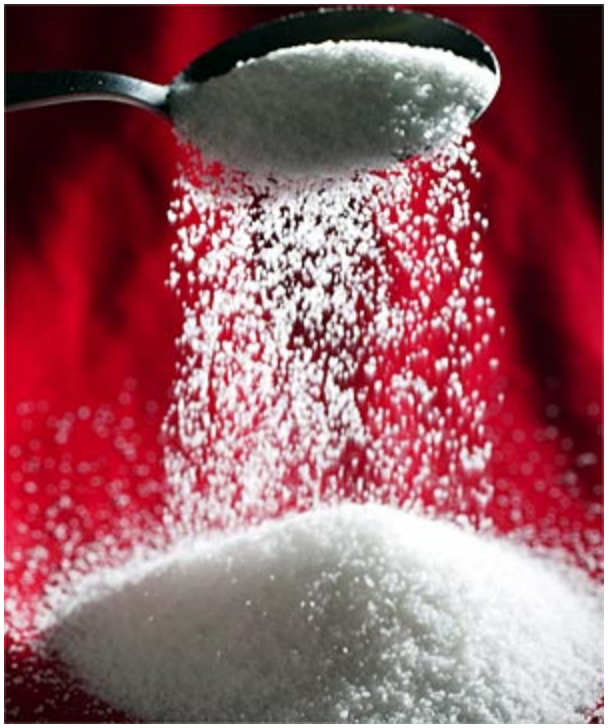


Inactive Ingredients Database

- Excipients in FDA approved NDA/ANDA formulations
- Varying Granularity - Formulations/Non-Specific Terms/Co-Processed Excipients
- Many “Ingredients” Map only as Concepts in GSRS
- UNII Mapping Provides More Standardized Names – IID does not carry synonyms

Sucrose UNII-C151H8M554





(2R,3R,4S,5S,6R)-2-[(2S,3S,4S,5R)-3,4-dihydroxy-2,5-bis(hydroxymethyl)tetrahydrofuran-2-yl]oxy-6-(hydroxymethyl)tetrahydropyran-3,4,5-triol

(Polydisperse) Polymers

Hypromellose Example

USP 33: Hypromellose is a methyl and hydroxypropyl mixed ether of cellulose. It contains, calculated on the dried basis, methoxy (–OCH3: 31.03) and hydroxypropoxy (–OC3H6OH: 75.09) groups conforming to the limits for the types of Hypromellose (hydroxypropyl methylcellulose) set forth in the accompanying table.

	Methoxy (percent)		Hydroxypropoxy (percent)	
Substitution Type	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

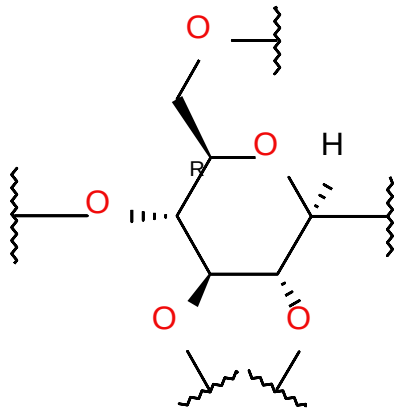
USP 33: Labeling—Label it to indicate its substitution type and its nominal viscosity value in milli-Pascal per second (mPa·s).

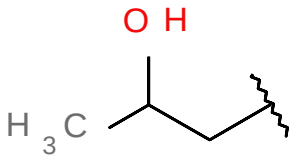
Hypromellose is the INN and BAN

CAS *9004-65-3*

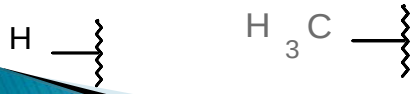
Hypromellose

Chiral





R1



R2

<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

Hypromellose

HYPROMELLOSE 2910 (3 MPA.S)	0VUT3PMY82
HYPROMELLOSE 2910 (5 MPA.S)	R75537T0T4
HYPROMELLOSE 2910 (6 MPA.S)	0WZ8WG20P6
HYPROMELLOSE 2910 (15 MPA.S)	36SFW2JZ0W
HYPROMELLOSE 2910 (50 MPA.S)	1IVH67816N
HYPROMELLOSE 2910 (4000 MPA.S)	RN3152OP35
HYPROMELLOSE 2910 (15000 MPA.S)	288VBX44JC
HYPROMELLOSE 2906 (50 MPA.S)	612E703ZUQ
HYPROMELLOSE 2906 (4000 MPA.S)	5EYA69XGAT
HYPROMELLOSE 2208 (3 MPA.S)	9H4L916OBU
HYPROMELLOSE 2208 (100 MPA.S)	B1QE5P712K
HYPROMELLOSE 2208 (4000 MPA.S)	39J80LT57T
HYPROMELLOSE 2208 (15000 MPA.S)	Z78RG6M2N2
HYPROMELLOSE 2208 (100000 MPA.S)	VM7F0B23ZI

Peanuts

- <SUBSTANCE_NAME> **PEANUT**
- <SUBSTANCE_ID> **QE1QX6B99R**
- <SOURCE_TYPE> **PLANT**
- <FAMILY> **FABACEAE**
- <GENUS> **ARACHIS**
- <SPECIES> **HYPOGAEA**
- <PART>**SEED**