**ISO IDMP Implementation Guide**

**Module 3**

**Data elements, structures and message specifications for unique identification and exchange of regulated information on substances**

Contents

[Table of Figures 2](#_Toc343650409)

[1 Introduction / Executive Summary 2](#_Toc343650410)

[2 Purpose 2](#_Toc343650411)

[3 Scope 2](#_Toc343650412)

[4 Business Case 2](#_Toc343650413)

[5 Application 2](#_Toc343650414)

[6 Background 2](#_Toc343650415)

[6.1 General Background and History 2](#_Toc343650416)

[6.2 Historical Perspective 2](#_Toc343650418)

[6.3 Markup Languages 2](#_Toc343650419)

[7 Representation of the IDMP 2](#_Toc343650420)

[7.1 Why Standardization of Identification of Medicinal Products is needed 2](#_Toc343650421)

[7.2 IDMP Scope 2](#_Toc343650422)

[8 Substance 2](#_Toc343650423)

[8.a Introduction 2](#_Toc343650424)

[8.b Defining Substances and Substance Identifiers (Substance IDs) 2](#_Toc343650425)

[8.c Requesting a Substance ID and providing information 2](#_Toc343650426)

[8.1 Substance ID 2](#_Toc343650427)

[8.2 Substance Name 2](#_Toc343650428)

[8.2.1 SubstanceName 2](#_Toc343650429)

[8.2.2 Substance Name Type 2](#_Toc343650430)

[8.2.3 Language 2](#_Toc343650431)

[8.2.4 Official Name 2](#_Toc343650432)

[8.2.5 Reference Source 2](#_Toc343650433)

[8.3 Substance Code 2](#_Toc343650434)

[8.3.1 Code 2](#_Toc343650435)

[8.3.2 Code System ID 2](#_Toc343650436)

[8.3.3 Code System 2](#_Toc343650437)

[8.3.4 Code System Status 2](#_Toc343650438)

[8.3.5 Code System Status Change Date 2](#_Toc343650439)

[8.3.6 Comment 2](#_Toc343650440)

[8.3.7 Reference Source 2](#_Toc343650441)

[8.4 Version 2](#_Toc343650442)

[8.4.1 Version Number 2](#_Toc343650443)

[8.4.2 Effective date 2](#_Toc343650444)

[8.4.3 Change Made 2](#_Toc343650445)

[8.5 Reference Information 2](#_Toc343650446)

[8.5.1 Comment 2](#_Toc343650447)

[8.5.2 Substance Classification 2](#_Toc343650448)

[8.5.3 Target 2](#_Toc343650449)

[8.5.4 Gene 2](#_Toc343650450)

[8.5.5 Gene Elements 2](#_Toc343650451)

[8.5.6 Substance Relationship 2](#_Toc343650452)

[8.6 Single Substance Type 2](#_Toc343650453)

[8.7 Structure 2](#_Toc343650454)

[8.7.1 Structural Representation Type 2](#_Toc343650455)

[8.7.2 Structural Representation 2](#_Toc343650456)

[8.7.3 Structural Representation Attachment 2](#_Toc343650457)

[8.7.4 Stereochemistry 2](#_Toc343650458)

[8.7.5 Optical Activity 2](#_Toc343650459)

[8.7.6 Molecular Formula 2](#_Toc343650460)

[8.7.7 Isotope 2](#_Toc343650461)

[8.8 Amount 2](#_Toc343650462)

[8.8.1 Average 2](#_Toc343650463)

[8.8.2 Low Limit 2](#_Toc343650464)

[8.8.3 High Limit 2](#_Toc343650465)

[8.8.4 Unit 2](#_Toc343650466)

[8.8.5 Non Numeric Value 2](#_Toc343650467)

[8.9 Property 2](#_Toc343650468)

[8.9.1 Property Type 2](#_Toc343650469)

[8.9.2 Property Name 2](#_Toc343650470)

[8.9.3 Substance ID 2](#_Toc343650471)

[8.9.4 Substance Name 2](#_Toc343650472)

[8.9.5 Amount 2](#_Toc343650473)

[8.10 Chemical Substance 2](#_Toc343650474)

[8.10.1 Substance Name 2](#_Toc343650475)

[8.10.2 Substance Code 2](#_Toc343650476)

[8.10.3 Reference Information 2](#_Toc343650477)

[8.10.4 Version 2](#_Toc343650478)

[8.10.5 Structure 2](#_Toc343650479)

[8.10.6 Stoichiometric 2](#_Toc343650480)

[8.10.7 Comment 2](#_Toc343650481)

[8.10.8 Reference Source 2](#_Toc343650482)

[8.10.9 Non- Stoichiometric 2](#_Toc343650483)

[8.11 Proteins/peptides 2](#_Toc343650484)

[8.11.1 Substance Name 2](#_Toc343650485)

[8.11.2 Substance Code 2](#_Toc343650486)

[8.11.3 Reference Information 2](#_Toc343650487)

[8.11.4 Version 2](#_Toc343650488)

[8.11.5 Reference Source 2](#_Toc343650489)

[8.11.6 Molecular Weight 2](#_Toc343650490)

[8.11.7 Structure 2](#_Toc343650491)

[8.11.8 Sequence Type 2](#_Toc343650492)

[8.11.9 Number of subunits 2](#_Toc343650493)

[8.11.10 Comment 2](#_Toc343650494)

[8.11.11 Protein Subunit 2](#_Toc343650495)

[8.11.12 Post-Translational Modifications 2](#_Toc343650496)

[8.11.13 Disulfide Linkage 2](#_Toc343650497)

[8.11.14 Glycosylation 2](#_Toc343650498)

[8.11.15 Modification 2](#_Toc343650499)

[8.12 Nucleic Acids 2](#_Toc343650500)

[8.12.1 Substance Name 2](#_Toc343650501)

[8.12.2 Substance Code 2](#_Toc343650502)

[8.12.3 Reference Information 2](#_Toc343650503)

[8.12.4 Version 2](#_Toc343650504)

[8.12.5 Structure 2](#_Toc343650505)

[8.12.6 Sequence Type 2](#_Toc343650506)

[8.12.7 Number of Subunits 2](#_Toc343650507)

[8.12.8 Area of hybridisation 2](#_Toc343650508)

[8.12.9 Comment 2](#_Toc343650509)

[8.12.10 Nucleic Acid Subunit 2](#_Toc343650510)

[8.12.11 Modification 2](#_Toc343650511)

[8.12.12 Molecular Weight 2](#_Toc343650512)

[8.12.13 Property 2](#_Toc343650513)

[8.12.14 Reference Source 2](#_Toc343650514)

[8.13 Polymers 2](#_Toc343650515)

[8.13.1 Substance Name 2](#_Toc343650516)

[8.13.2 Substance Code 2](#_Toc343650517)

[8.13.3 Reference Information 2](#_Toc343650518)

[8.13.4 Version 2](#_Toc343650519)

[8.13.5 Structure 2](#_Toc343650520)

[8.13.6 Polymer Class 2](#_Toc343650521)

[8.13.7 Polymer Geometry 2](#_Toc343650522)

[8.13.8 Copolymer Sequence type 2](#_Toc343650523)

[8.13.9 Comment 2](#_Toc343650524)

[8.13.10 Monomer Description 2](#_Toc343650525)

[8.13.11 Structural Repeat 2](#_Toc343650526)

[8.13.12 Molecular Weight 2](#_Toc343650527)

[8.13.13 Property 2](#_Toc343650528)

[8.13.14 Reference Source 2](#_Toc343650529)

[8.13.15 Source Material 2](#_Toc343650530)

[8.13.16 Modification 2](#_Toc343650531)

[8.14 Structurally-Diverse Substance 2](#_Toc343650532)

[8.14.1 Substance Name 2](#_Toc343650533)

[8.14.2 Substance Code 2](#_Toc343650534)

[8.14.3 Reference Information 2](#_Toc343650535)

[8.14.4 Version 2](#_Toc343650536)

[8.14.5 Source Material 2](#_Toc343650537)

[8.14.6 Comment 2](#_Toc343650538)

[8.14.7 Modification 2](#_Toc343650539)

[8.14.8 Source material 2](#_Toc343650540)

[8.14.9 Glycosylation 2](#_Toc343650541)

[8.14.10 Molecular Weight 2](#_Toc343650542)

[8.14.11 Property 2](#_Toc343650543)

[8.14.12 Reference Source 2](#_Toc343650544)

[8.15 Mixture Substance 2](#_Toc343650545)

[8.15.1 Mixture Type 2](#_Toc343650546)

[8.15.2 Mixture Constituent 2](#_Toc343650547)

[8.15.3 Source Material 2](#_Toc343650548)

[8.15.4 Modification 2](#_Toc343650549)

[8.15.5 Substance Name 2](#_Toc343650550)

[8.15.6 Substance Code 2](#_Toc343650551)

[8.15.7 Reference Information 2](#_Toc343650552)

[8.15.8 Version 2](#_Toc343650553)

[9 Specified Substance 2](#_Toc343650554)

[9.1 Group 1 Specified Substance 2](#_Toc343650555)

[9.1.1 Specified\_Substance\_Group1\_ID 2](#_Toc343650556)

[9.1.2 Specified\_Substance\_Group1\_Name 2](#_Toc343650557)

[9.1.3 Substance Name 2](#_Toc343650558)

[9.1.4 Substance Code 2](#_Toc343650559)

[9.1.5 Version 2](#_Toc343650560)

[9.1.6 Reference Information 2](#_Toc343650561)

[9.1.7 Property 2](#_Toc343650562)

[9.1.8 Reference Source 2](#_Toc343650563)

[9.1.9 Constituent 2](#_Toc343650564)

[9.1.10 Physical Form 2](#_Toc343650565)

[9.2 Group 2 Specified Substance 2](#_Toc343650566)

[9.2.1 Specified\_Substance\_Group2\_ID 2](#_Toc343650567)

[9.2.2 Specified\_Substance\_Group2\_Name 2](#_Toc343650568)

[9.2.3 Parent Substance ID 2](#_Toc343650569)

[9.2.4 Manufacturing 2](#_Toc343650570)

[9.2.5 Substance Name 2](#_Toc343650571)

[9.2.6 Substance Code 2](#_Toc343650572)

[9.2.7 Version 2](#_Toc343650573)

[9.2.8 Reference Information 2](#_Toc343650574)

[9.3 Group 3 Specified Substance 2](#_Toc343650575)

[9.3.1 Specified\_Substance\_Group3\_ID 2](#_Toc343650576)

[9.3.2 Specified\_Substance\_Group3\_Name 2](#_Toc343650577)

[9.3.3 Parent Substance ID 2](#_Toc343650578)

[9.3.4 Grade 2](#_Toc343650579)

[9.3.5 Substance Name 2](#_Toc343650580)

[9.3.6 Substance Code 2](#_Toc343650581)

[9.3.7 Version 2](#_Toc343650582)

[9.3.8 Reference Information 2](#_Toc343650583)

[9.4 Group 4 Specified Substance 2](#_Toc343650584)

[9.4.1 Specified\_Substance\_Group4\_ID 2](#_Toc343650585)

[9.4.2 Parent Substance ID 2](#_Toc343650586)

[9.4.3 Grade 2](#_Toc343650587)

[9.4.4 Physical Form 2](#_Toc343650588)

[9.4.5 Constituent 2](#_Toc343650589)

[9.4.6 Analytical Data 2](#_Toc343650590)

[9.4.7 Manufacturing 2](#_Toc343650591)

[9.4.8 Constituent 2](#_Toc343650592)

[9.4.9 Property 2](#_Toc343650593)

[9.4.10 Physical Form 2](#_Toc343650594)

[9.4.11 Substance Name 2](#_Toc343650595)

[9.4.12 Substance Code 2](#_Toc343650596)

[9.4.13 Version 2](#_Toc343650597)

[9.4.14 Reference Information 2](#_Toc343650598)

[10 Annex 2](#_Toc343650599)

[10.1 Complete Template エラー! ブックマークが定義されていません。](#_Toc343650600)

[10.2 Example AMOXICILLIN 2](#_Toc343650601)

Table of Figures

[Figure 1 – High-level Substance-Specified substance information model 2](#_Toc343720764)

[Figure 2 – High-level information model for Substances 2](#_Toc343720765)

[Figure 3 – Decision tree or process by which the type of substance is determined 2](#_Toc343720766)

[Figure 4 – Determine further the type of single substance. 2](#_Toc343720767)

[Figure 5 – Information model for Substance Names 2](#_Toc343720768)

[Figure 6 – Information model for Substance Codes 2](#_Toc343720769)

[Figure 7 – Information model for Substance Classification 2](#_Toc343720770)

[Figure 8 – Information model for Structure and Isotope 2](#_Toc343720771)

[Figure 9 – Information model for the Chemical Substance 2](#_Toc343720772)

[Figure 10 – Information model for the Protein Substance 2](#_Toc343720773)

[Figure 11– Information Model for Modifications 2](#_Toc343720774)

[Figure 12– Information Model for Nucleic Acid 2](#_Toc343720775)

[Figure 13– Information Model for Polymer Substance 2](#_Toc343720776)

[Figure 14– Information Model for Source Material 2](#_Toc343720777)

[Figure 15– Information Model for Structurally-Diverse Substance 2](#_Toc343720778)

[Figure 16– Information Model for Mixture Substance 2](#_Toc343720779)

[Figure 17– Information model for the Group 1 specified substance 2](#_Toc343720780)

[Figure 18– Information model for the Group 2 specified substance 2](#_Toc343720781)

[Figure 19– Information model for the Group 3 specified substance 2](#_Toc343720782)

[Figure 20– Information model for the Group 4 specified substance 2](#_Toc343720783)

[Figure 21– Information model for the Analytical Data 2](#_Toc343720784)

[Figure 22 – Information model for the Manufacturing Data 2](#_Toc343720785)

# Introduction / Executive Summary

This document is a guide for implementing ISO/FDIS 11238: Health Informatics - Identification of medicinal products Data elements and structures for the unique identification and exchange of regulated information on substances. This standard was developed in response to a worldwide demand for internationally harmonised specifications for medicinal products. It is one of a group of five standards which together provide the basis for the unique identification of medicinal products. The other standards in this group are:

* ISO/FDIS 11615 Health Informatics - Identification of medicinal products - Data elements and structures for the unique identification and exchange of regulated medicinal product information
* ISO/FDIS 11616 Health informatics - Identification of medicinal products - Data elements and structures for the unique identification and exchange of regulated pharmaceutical medicinal product
* ISO/FDIS 11239 Health Informatics — Identification of Medicinal Products — Data elements and structures for the unique identification and exchange of regulated information on pharmaceutical dose forms, units of presentation, routes of administration and packaging
* ISO/FDIS 11240 Health informatics —Identification of medicinal products - Data elements and structures for the unique identification and exchange of units of measurement

The standards for the Identification of Medicinal Products (IDMP) support the activities of medicines regulatory agencies worldwide by jurisdiction. These include a variety of regulatory activities related to development, registration and life cycle management of medicinal products as well as pharmacovigilance and risk management.

The standard allows for the establishment and maintenance of a system that is capable of assigning unique identifiers for all substances present in medicinal products or in packaging materials in which medicinal products are contained. The standard describes general rules for defining and distinguishing substances and a high level model for the structuring of information for substances. This implementation guide provides detailed explanation of each type or grouping of substance information, an element-by-element schema for the use of the standard and examples for a variety of substances. The implementation of the standard will result in strong non-semantic unique identifiers for every substance present in a medicinal product. The use of the identifier is essential for the description of substances in medicinal products on a global scale. The standard does not involve developing nomenclature for substances or specified substances but common and official substance names in current use can instead be mapped to each identifier.

# Purpose

The business objective of this implementation guide is to provide a means for exchanging regulatory substance information. To meet the primary objectives of the regulation of medicines and pharmacovigilance, it is necessary to reliably exchange medicinal product information in a robust and reliable manner.

In the context of exchange of regulatory information the purpose of this IDMP Implementation Guide is twofold:

* To specify data elements, structures and relationships between the data elements required to uniquely and with certainty identify the substances present in and in contact with medicinal products.
* To specify definitions of terms for all data elements required to uniquely and with certainty identify the substances present in and in contact with medicinal products.

# Scope

Materials used in medicinal products range from simple chemicals to gene-modified cells to animal tissues. To unambiguously define these substances is particularly challenging. This standard defines substances based on their scientific identity (i.e. what they are) rather than on their use or method of production. Molecular structure or other immutable properties such as taxonomic, anatomical and/or fractionation information are used to define substances. The standard contains five groups of elements that are sufficient to define all substances. Although it is certainly possible to define or classify substances in other ways, the standard uses a minimalist structured scientific concept approach focusing on the critical elements necessary to distinguish two substances from one another. Whenever substances are mixed together there are frequently interactions between substances but the standard has intentionally not included these supramolecular interactions at the substance level because of the variable nature and strength of such interactions. The standard also allows the capture of multiple terms which refer to a given substance and a variety of reference information that could be used to classify substances or relate one substance to another.

In addition to the substance level the standard also provides elements for the capture of further information on substances, such as grade, manufacturer, manufacturing specifications, and also to capture information on substances that are frequently combined together in commerce but are not a medicinal product per se. At the specified substance level, four groups of elements provide information essential to the tracking and description of substances in medicinal products.

The basic concepts in the regulatory and pharmaceutical standards development domain use a wide variety of terms in various contexts. The terms and definitions described in this standard shall be applied for the concepts which are required to uniquely identify, characterize and exchange information on regulated medicinal products.

The terms and definitions adopted in this standard are intended to facilitate the interpretation and application of legal and regulatory requirements, but they shall be without prejudice to any legally binding document. In case of doubt or potential conflict, the terms and definitions contained in legally binding documents shall prevail

# Business Case

The IDMP standards support the following interactions:

* Medicine Regulatory Agency to Medicine Regulatory Agency, e.g. European Medicines Agency to the US Food and Drug Administration (FDA) or vice versa
* Pharmaceutical company to Medicine Regulatory Agency, e.g. Pharma Company A to Health Canada
* Sponsor of clinical trial to Medicine Regulatory Agency, e.g. University X to Austrian Medicines Agency
* Medicine Regulatory Agency to other stakeholders, e.g. UK Medicines and Health Care Products Regulatory Agency (MHRA) to National Health Service (NHS)
* Interaction of Medicine Regulatory Agency with worldwide-maintained data sources e.g. Pharmaceutical and Medical Device Agency (PMDA) to the organization responsible for assigning substance identifiers

The necessary messaging specifications are included as an integral part of the IDMP standards to secure the interactions above. Unique identifiers produced in conformance with the IDMP standards will support applications where it is necessary to reliably identify and trace the use of medicinal products and the materials within medicinal products.

# Application

This IDMP implementation guide is built upon the following international standards:

• ISO/FDIS 11238: Health informatics — Identification of medicinal products — Data elements and structures for the unique identification and exchange of regulated information on substances

• HL7 Version 3 Standard, Common Product Model CMETs, DSTU Update Release 10

• HL7 Common Product Model, Substance Model (POCP\_MT080100) and Substance Specification Model (POCP\_MT090100)

# Background

## General Background and History

## Markup Languages

Extensible Markup Language (XML) is a set of rules for encoding documents in machine-readable form. The design goals of XML emphasize simplicity, generality, and usability over the Internet. It is a textual data format with strong support via Unicode for the languages of the world. Although the design of XML focuses on documents, it is widely used for the representation of arbitrary data structures, for example in web services.

# Representation of the IDMP

## Why Standardization of Identification of Medicinal Products is needed

Medicines regulatory authorities and pharmaceutical industry engage in an intensive information exchange during drug development, drug evaluation and approval phase and the post-authorization phase. The standardization of medicinal product information is regarded as one of the key elements of this information flow.

However, regulators in the ISOregions and observer countries have established their own procedures and applications with standards that differ in data format, content, language and applied terminology (e.g. terminology used for describing substances, routes of administration, pharmaceutical dose forms).

Due to the lack of a common and harmonized approach, both regulators and pharmaceutical industry are confronted with the following issues:

• No possibility to exchange medicinal product information between medicines regulatory authorities and pharmaceutical industry in a structured and efficient way;

• Difficulties in ensuring data consistency and in evaluating and comparing medicinal product -related information across the ISO regions due to the lack of harmonized definitions of terminologies and data sets. This currently impairs pharmacovigilance activities in particular;

• For the pharmaceutical industry, major administrative burdens and duplication of efforts requiring substantial human and financial resources to comply with and handle different regional requirements;

• Lack of consistency in the use of terminology in the health care community.

The objectives of the IDMP are to address the issues outlined above by developing harmonized standards that build on the regulatory and technical processes already established and to support the population and maintenance of existing systems/applications with fully reliable regulatory medicinal product information.

Harmonized standards will stimulate vendors to develop “off-the-shelf” tools (that are interoperable due to the standard itself). Harmonized standards will also help maximize forward compatibility of data and minimize the complexities of backward compatibility.

## IDMP Scope

IDMP is based on a group of five ISO standards, which include:

* ISO/FDIS 11615 Health Informatics — Identification of Medicinal Products — Data elements and structures for the unique identification and exchange of regulated medicinal product information
* ISO/FDIS 11616 Health informatics — Identification of medicinal products — Data elements and structures for the unique identification and exchange of regulated pharmaceutical product information
* ISO/FDIS 11238 Health informatics — Identification of medicinal products — Data elements and structures for the unique identification and exchange of regulated information on substances
* ISO/FDIS 11239 Health Informatics — Identification of medicinal products — Data elements and structures for the unique identification and exchange of regulated information on pharmaceutical dose forms, units of presentation, routes of administration and packaging
* ISO/FDIS 11240 Health informatics — Identification of medicinal products — Data elements and structures for the unique identification and exchange of units of measurement

Together, these five ISO standards define, characterise and uniquely identify regulated medicinal products for human use during their entire life cycle i.e. from development, to authorisation, post-marketing and renewal or withdrawal from the market, where applicable. More specifically:

The ISO/FDIS 11615 standard establishes definitions and concepts and describes data elements and their structural relationships, which are required for the detailed description and unique identification of medicinal products.

The ISO/FDIS 11616 standard defines data elements, structures, and relationships between data elements required for the exchange of regulated information to uniquely identify pharmaceutical medicinal products. The standard is applicable to regulatory and pharmacovigilance activities worldwide. It establishes definitions and concepts, and specifies data elements and their structural relationships for the detailed description and unique identification of pharmaceutical products.

The ISO/FDIS 11238 standard provides an information model to define and identify substances within medicinal products or used for medicinal purposes, including dietary supplements, food and cosmetics; the latter are also important in the context of pharmacovigilance, where food and cosmetics can interact with medicinal products.

The ISO/FDIS 11239 standard specifies:

* The data elements, structures, and relationships between the data elements required for the exchange of information that uniquely and with certainty identify pharmaceutical dose forms, units of presentation, routes of administration and packaging items (immediate containers, closures, administration devices and outer packaging) related to medicinal products;
* A mechanism for the translation of the terms from English into other languages, which is an integral part of the information exchange;
* A mechanism for the versioning of the concepts in order to track their evolution;
* Rules to allow regional medicines regulatory authorities to map existing regional terms to the terms created using this standard, in a harmonised and meaningful way.

The ISO/FDIS 11240 standard is limited to the representation of units of measurement for data interchange between computer applications. It:

* Specifies rules for the usage and coded representation of units of measurement for the purpose of exchanging information about quantitative medicinal product characteristics (e.g. strength) in the human medicine domain, that require units of measurement,
* Establishes requirements for units in order to provide traceability to international metrological standards,
* Provides rules for the standardized and machine-readable documentation of quantitative composition and strength of medicinal products, specifically in the context of medicinal product identification,
* Defines the requirements for the representation of units of measurement in coded form
* Provides structures and rules for mapping between different unit vocabularies and language translations to support the implementation of this standard, taking into account that existing systems, dictionaries and repositories use a variety of terms and codes for the representation of units.

Furthermore, to support the successful information exchange in relation to the unique identification and characterisation of medicinal products, the use of other normative HL7 messaging standards should be applied in the context of the ISO IDMP standards.

This document is a guide for implementing the “Health informatics — Identification of Medicinal Products — Data elements and structures for the unique identification and exchange of regulated information on Substances.”

This Implementation Guide is intended to assist reporters (including pharmaceutical companies, authorities and non-commercial sponsors) in constructing messages or transmitting information that allows substances to be defined and an unique ID assigned to such substances. It is not intended to be a guide for a maintenance authority responsible for the generation and storage of unique substance IDs and defining information on which the id is based.

This is an example table for class and elements description:

|  |  |
| --- | --- |
| **User Guidance** |  |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** |  |
| **OID** |  |
| **Value Allowed** |  |
| **Business Rule(s)** |  |
| **IDMP HL7 SPL Snippet** |  |

In contrast to other part of the guide, conformance will refer to whether an element is required for a given substance type or a specified substance group. Conformance is not meant to be applied globally.

# Substance

## 8.a Introduction

All medicinal products consist of substances; these substances can be active ingredients, excipients, or material in packaging materials. There are two fundamental levels of information described in the standard, a ***substance*** level and what is called a “***specified substance***” level. What the standard has strived for is a consistent, scientifically valid approach that is meaningful for defining and representing substances for regulatory purposes. For substances, the approach taken is to define substances based on a substance’s inherent attributes not on their use or method of manufacture. At the specified substance level, four separate groups of elements provide additional information that is particularly relevant for regulatory uses.

In order to decide to define or distinguish material either at a substance or specified substance level several attributes should be taken into consideration. There are certain information elements that are only associated with substances and others only associated with specified substances: For chemicals, molecular structure will only be captured at the substance level. For proteins, the amino-acid sequence, sites and type of glycosylation, and the presence and position of disulfide bonds will only be captured at the substance level. For nucleic acids, the sequence, type of sugar and linkage will only be captured at the substance level. For other polymers, the monomers used to synthesize the polymer, the structural repeating units, the molecular weight and/or a property related to molecular weight (i.e viscosity), the source of naturally derived polymers and any modifications that irreversibly alter the molecular structure of a polymer will be captured at the substance level. For structurally diverse material, taxonomic, anatomical and fractionation information, properties related to the underlying molecular structure of the material, and moid will be captured at the substance level. Mixtures consists of a simple combination of single substances that are either isolated together or the result of the same synthetic process. The biological source of the mixture is also captured at the substance level. Proportions are not captured at the substance level.

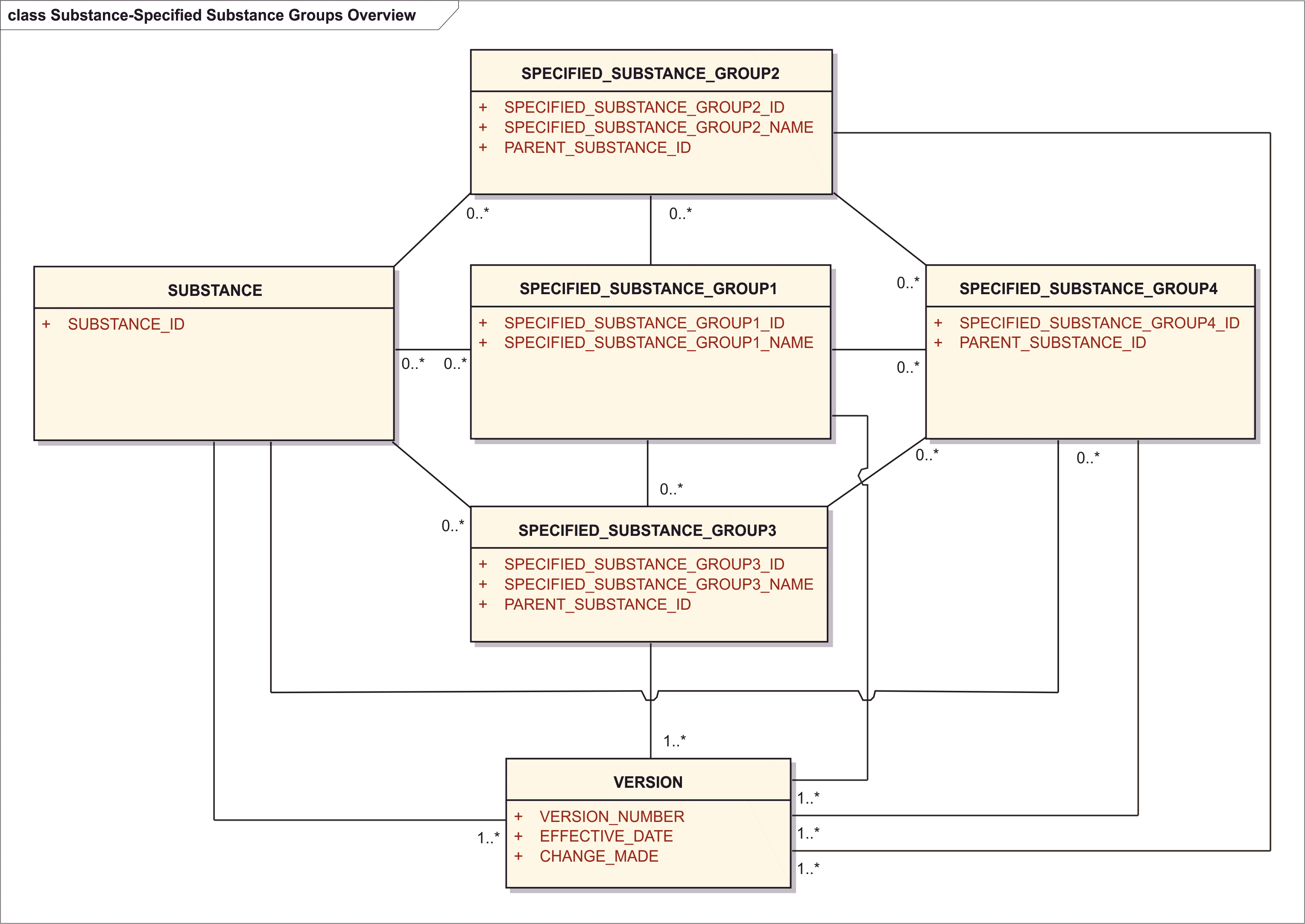


Figure 1 – High-level Substance-Specified substance information model

There are 4 groups of elements that are used to further define and specify substances. Specified substances will always map to a substance or substances. Material defined as a group 1 specified substance typically consist of material containing multiple substances mixed together to form an ingredient, extracts of material typically of biological origin, polymorphic forms of substances, or substances that exist in a particular physical state. Biological or physical properties that are essential to the use of a substance will also be captured at the group 1 specificied substance level.

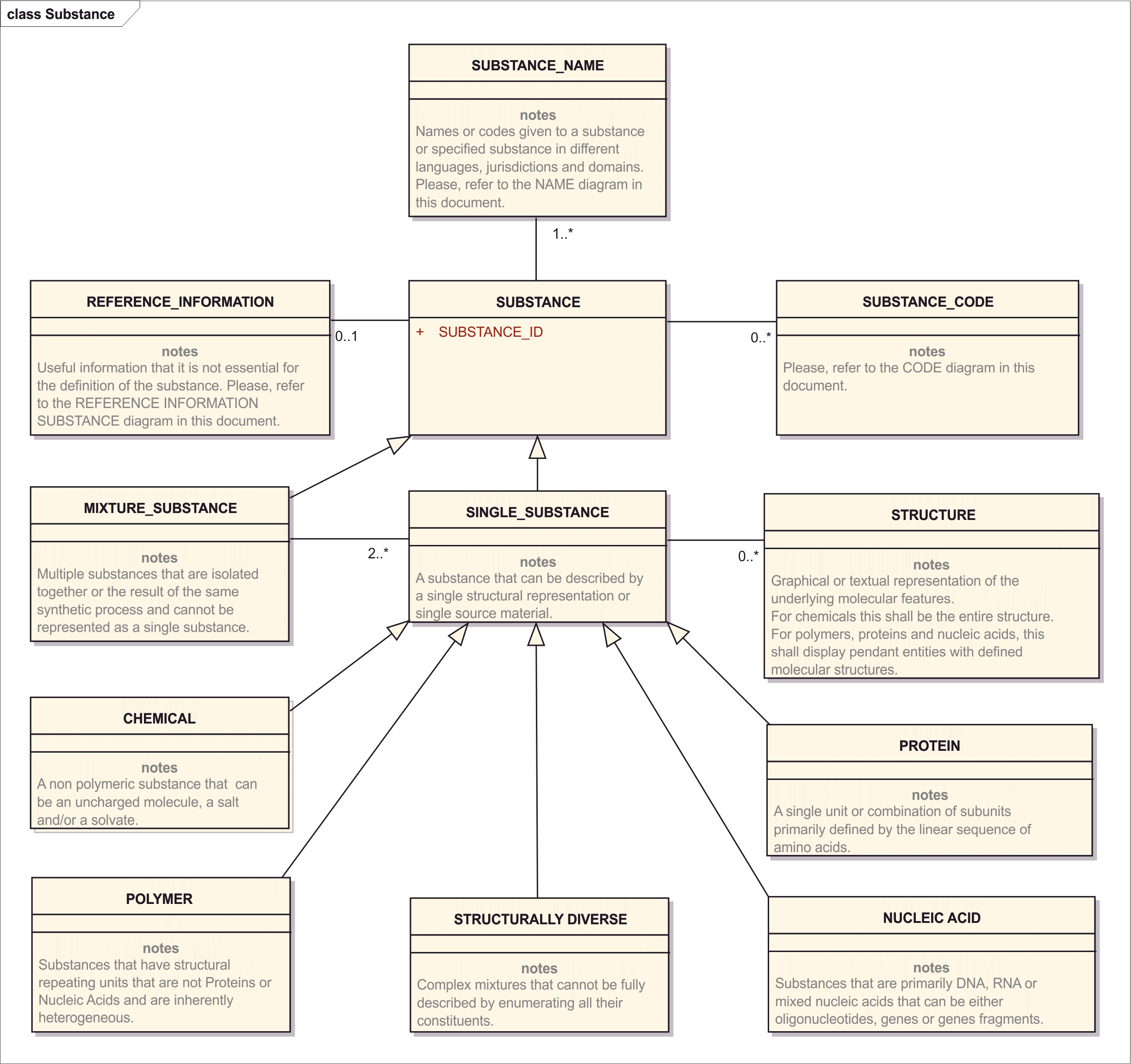


Figure 2 – High-level information model for Substances

## 8.b Defining Substances and Substance Identifiers (Substance IDs)

A substance is any matter that has a discrete existence, irrespective of origin, which may be biological or chemical. Substances can be single well-defined chemical entities containing a definite molecular structure, synthetic (e.g., isomeric mixtures) or naturally-occurring (e.g., conjugated estrogens) mixtures of chemicals, or materials derived from plants, animals, microorganisms or inorganic matrices that are not definable by a single or limited number of molecular structures. Substances may be active moieties, salts, solvates, stoichiometric complexes or mixtures of compounds that are either isolated or synthesized together. Materials that are combined from multiple sources to form a product are not considered substances.

A substance is defined by what it is, and is, not by how it is made or how it is used. Substance definitions are based on the immutable properties of a given material. These properties include the molecular structure, or structures of a given material, taxonomic, anatomical or fractionation information on material that cannot be represented by molecular structures. Purity, physical form, and method of production are not considered when defining substances.

For the purposes of definition, all **single substances** are of one of the following types:

* (simple) chemical
* protein
* nucleic acid
* polymer
* structurally-diverse

Each of these different types of substances has distinct elements essential for definition. Substances can also be a **mixture of single substances**.

The primary goal of the ISO IDMP - *Data Elements and Structures for the unique identification and exchange of regulated information on substances* is to define unambiguously all substances present in regulated products. Once a substance has been defined, the maintenance organization will assign a unique identifier that is permanently associated with that definition. This document describes the necessary information for this registration process and in addition reference information that can be associated with a substance.

In order for the maintenance organization to generate a Substance ID, a certain amount of information must be provided so that the substance can be identified unambiguously.

The assignment of the same Substance ID to two materials does not indicate that these materials are equivalent either biologically or pharmaceutically, but rather only indicates that the material is defined in an equivalent manner with identical defining elements.

Some materials will not have a Substance ID because they are manufactured materials containing multiple substances rather than individual substances.

Substances that are mixed together to form a product, even if there are physical interactions between the substances, will not receive a Substance ID but will be described as a specified substance.

*Example: FD&C Blue No. 1 Aluminum Lake is not defined as a mixture substance but rather as two separate substances FD&C Blue No. 1 and aluminum oxide. Also, simethicone is not a substance because it contains two substances, namely dimethicone and silicon dioxide. These materials would be described as group 1 specified substances.*

Conversely, if new covalent bonds and/or simple salts or solvates are formed, separate Substance IDs for each molecular entity are assigned.

*Examples: fluticasone propionate would have a separate ID from fluticasone; atorvastatin calcium trihydrate would have a separate ID from atorvastatin calcium anhydrous.*

## 8.c Requesting a Substance ID and providing information

Before requesting a Substance ID, the Substance ID list on the public website should be checked to determine if a Substance ID has already been generated for the relevant substance.

When selecting a Substance ID, always choose the Substance ID that is the most specific.

*Example: if the substance is polyethylene glycol 8000, use the Substance ID for the substance rather than the more generic polyethylene glycol Substance ID.*

If the appropriate Substance ID is not available on this list, a Substance ID request should be sent to the maintenance organization.

In submitting information for the assignment of a Substance ID any information considered confidential should be so indicated. Any confidential information linked to the substance will only be released with the permission of the submitter or if it is unambiguously found in the public domain associated with the substance.

*Example: If a chemical structure or protein sequence is found in an INN, USAN, or JAN definitions that information can be released to the public even it was marked confidential when the substance ID was generated.*

Because of the diversity of materials used in medicinal products there are many fields in the standard that are used to define substances. In most instances, however, only a few fields are necessary to unambiguously define a substance. For most simple chemicals a complete molecular structure is sufficient to assign an ID.

At least one name or company code should be associated with each request for a substance id. To facilitate mapping and to limit ambiguity a company should submit all codes and common names that are known to have be associtiated with a given substance. In addition to a request for substance IDs, additional information can be associated with the substance and this information will be the maintained. Once an ID is assigned it will be permanently associated with the substance.

<document>  
 <id root="***Document Id***"/>  
 <code code="64124-1" codeSystem="2.16.840.1.113883.6.1"  
 displayName="Indexing - Substance"/>  
 <effectiveTime value="***Effective Date***"/>  
 <setId root="***Document Set Id***"/>  
 <versionNumber value="***Version Number***"/>  
 <author>  
 <node>***Changes Made***</node>  
 <!-- ***author information*** -->  
 </author>  
 <component>  
 <structuredBody>  
 <component>  
 <section>  
 <subject>  
 <identifiedSubstance><!-- **Substance** -->  
 <id extension="*Substance Id*" root="***Substance Id System (OID)***"/>  
 <identifiedSubstance>  
 <code code="***Substance Id*** "   
 codeSystem="***Substance Id System (OID)***"/>  
 <name>Substance Name</name>  
 <!-- ***Substance Definition*** -->  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification><!-- **Specified Substance** -->  
 <code code="***Substance Specification Code***"   
 codeSystem="***Code System (OID)***"  
 displayName="***Substance Specification Name***">

## Substance ID

Every substance will be identified by an ID. The ID should be unique to each substance, non-semantic, non-chronological, and of fixed length with an integrity check. Once an ID is assigned to a given physical or conceptual material the ID will not change. Changes or additions to defining information will result in a new version of the the substance definition but would not result in a new substance id. Although an effort should always be made to obtain complete information prior to assignment there will be instances when additional defininational information on a given substance is transmitted after assignment. There will also be instances when incorrect information may be transmitted and used to define a substance. In these cases the information will be changed or added but the ID will not change. In the rare cases when the same substance is found to have two IDs one of the IDs will remain a primary ID and the use of the other ID will be deprecated.

Both substances and specified substances will be identified with a substance ID. Substance IDs will be released to the public if the substance is in a licensed medicinal product. The IDs for substances in the investigational stage will only be released if an official name exists or a company code is associated with defining information is found in a single reference that is in the public domain.

|  |  |
| --- | --- |
| **User Guidance** | ID to be used in all electronic submissions to identify a substance. Generated when sufficient information is available to unambiguously define a substance. ID will be permanently associated with a given substance and each substance at the substance level shall have one and only one ID.  NOTE: If a unique “Substance ID” has been assigned, this “Substance ID” shall be specified based on the Substance Name controlled vocabulary. In the absence of a unique “Substance ID”, e. g. for the initial submission of the substance, this data element is not required. |
| **Example(s)** | UNII Code |
| **Conformance** | REQUIRED for Substance Definition, NOT SPECIFIED for Substance ID Request |
| **Data Type** | II / CV |
| **OID** | II.root / CV.codeSystem are 2.16.840.1.113883.4.9 |
| **Value Allowed** |  |
| **Business Rule(s)** | All substances will be identified by a single ID.  NOTE: The ID will only be released to the public if the defining information is in the public domain or if a company that provides the defining informations requests public release or releases the code in public marketing materials. Defining information found in patents will usually not be sufficient to release the ID to the public. Even if an ID can be released there may be elements that will not be released to the general public. An ID will always be released to an organization that requests an ID and supplies information to define.a substance. A Flag to control the release of the ID to the general public is part of the Reference Source information (see Section 8.2.5). |

<document>  
 <component>  
 <structuredBody>  
 <component>  
 <section>  
 <subject>  
 <identifiedSubstance><!-- **Substance** -->  
 <id extension="*Substance Id*" root="***Substance Id System (OID)***"/>  
 <identifiedSubstance>  
 <code code="***Substance Id*** "   
 codeSystem="***Substance Id System (OID)***"/>

Note: the Substance ID occurs twice in the SPL, once as the id/@extension attribute of the identifiedSubstance Role and once as the code/@code attribute of the identifiedSubstance Entity. The id/@root and code/@codeSystem are the same OID “2.16.840.1.113883.4.9”

## Substance Type

Material will either be described as a substance one of four groups of specified substances. Substances are one of five types of single substances or as mixture. Mixtures consist of a combination of single substances where the source material is also captured but proportions are not.

The first step in defining a substance is to determine whether the substance can be defined as a **single substance** or a **mixture of single substances** which is referred to as a **mixture substance**. A mixture substance is a substance that cannot be defined using a single set of substance elements or structural representations. Racemic mixtures, most polymers and structurally diverse substances although inherently mixtures can be defined using a single set of elements and therefore are not describes as mixtures.

Mixtures are substances like Gentamicin or Gentamicin C which consists of a mixture of three related individual aminoglycosides, Gentamicin C1A, Gentamicin C1, and Gentamicin C2 which are typically isolated together from a bacterial culture. Corn Starch would also be a mixture substance of two distinct polymers amylose and amylopectin with corn as the source of the substance.

Most other polymer will be described as single substances with the definition based on the structural repeating unit and a representation of the average molecular weight of the polymer.

The figures below represent decision tree or process by which the type of substance is determined.

A substance will either be a mixture or a single substance.



Figure 3 – Decision tree or process by which the type of substance is determined

If a substance is monodisperse the following tree would be used to further determine the type of single substance.



Figure 4 – Determine further the type of single substance.

Single substances shall be defined by at least one of the following single substance classes; CHEMICAL, PROTEIN, NUCLEIC ACID, POLYMER, or STRUCTURALLY DIVERSE. Choice of substance class is described in the following figure and the one above.

All single substance will be defined as one and only one of the five classes listed in figure below. Although some single substances may be made up of material of more than one class a single call will be chosen to define each substance. Pegolylated Proteins which contain a polymeric form of ethylene glycol attached to a protein would be described as a protein with a polymeric fragment modification. All single substance should be able to describe by the elements listed below.

In order to define a substance that is not a mixture a single type of substance should be chosen. If multiple types of substances exist within the same material a decision should first be made as to the type of substance upon which the primary definition will be based. If the substance can be defined as a single entity that type of entity will be chosen. If not the decision should first be based on the portion of the material provides the predominant functional activity of the material.

*Examples: A well defined small molecule chemical that is covalently attached to a polymer, protein or nucleic acid will be described as a polymer or protein. The small molecule will either be an end group of a polymer or part of a fragment modification of the protein or peptide. Pegoylated proteins will be described as proteins and not polymers. The polyethylene glycol substitutient with the linker will be captured as a fragment modification of the protein. Polysaccharide antigens conjugated to a protein will captured as the same type of substance as the antigen, a polymer if the structure of the polysaccharide is known or structurally diverse material if it is not.*

|  |  |
| --- | --- |
| **User Guidance** | All substances will be defined as a single type of substance. Monodisperse substances are typically material that can be defined as a single molecular entity. Monodisperse substances are either chemicals, proteins, or a nucleic acids. Polydisperse material can be a mixture, polymer or structurally diverse material. Material that cannot be defined |
| **Example(s)** | CHEMICAL |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **OID** |  |
| **Value Allowed** | CHEMICAL, PROTEIN, NUCLEIC ACID, POLYMER, STRUCTURALLY DIVERSE, MIXTURE |
| **Business Rule(s)** | Only one type |
| **IDMP SPL Snippet** |  |

Representation in HL7 is as a substance classification, the “IDMP Substance Type” may be assigned like any other substance classification code. The fact that this is the substance type specification becomes evident from recognizing that the assigned classification code is a concept considered “substance type”.

<identifiedSubstance>  
 <identifiedSubstance>

<asSpecializedKind>  
 <definingMaterialKind>  
 <code code="***Substance Type*** *Code*"   
 codeSystem="*Substance Type Code System (OID)*"/>  
 </definingMaterialKind>

## Substance Name (repeat as necessary)

Although the standard does not provide any guidance on substance nomenclature, it does provide a structure for the capture of names and codes that are often used to identify a substance. For the purpose of the standard there are basically four types of names, Official names, other names, brand names and company codes. Official names are typically nonproprietary names used in a given jurisdiction and domain to identify a substance. The domain, jurisdiction, authority that assigned the name (USAN, INN, JAN etc) and the language are also captured. Other names are names associated with a given substance that are typically common or chemical names. Brand names are typically names by which a company identifies a given substance typically for marketing purposes. Company codes are treated as names. They are assigned by a given company to substances in clinical or preclinical development.

At least one name or company code will be associated with each substance.

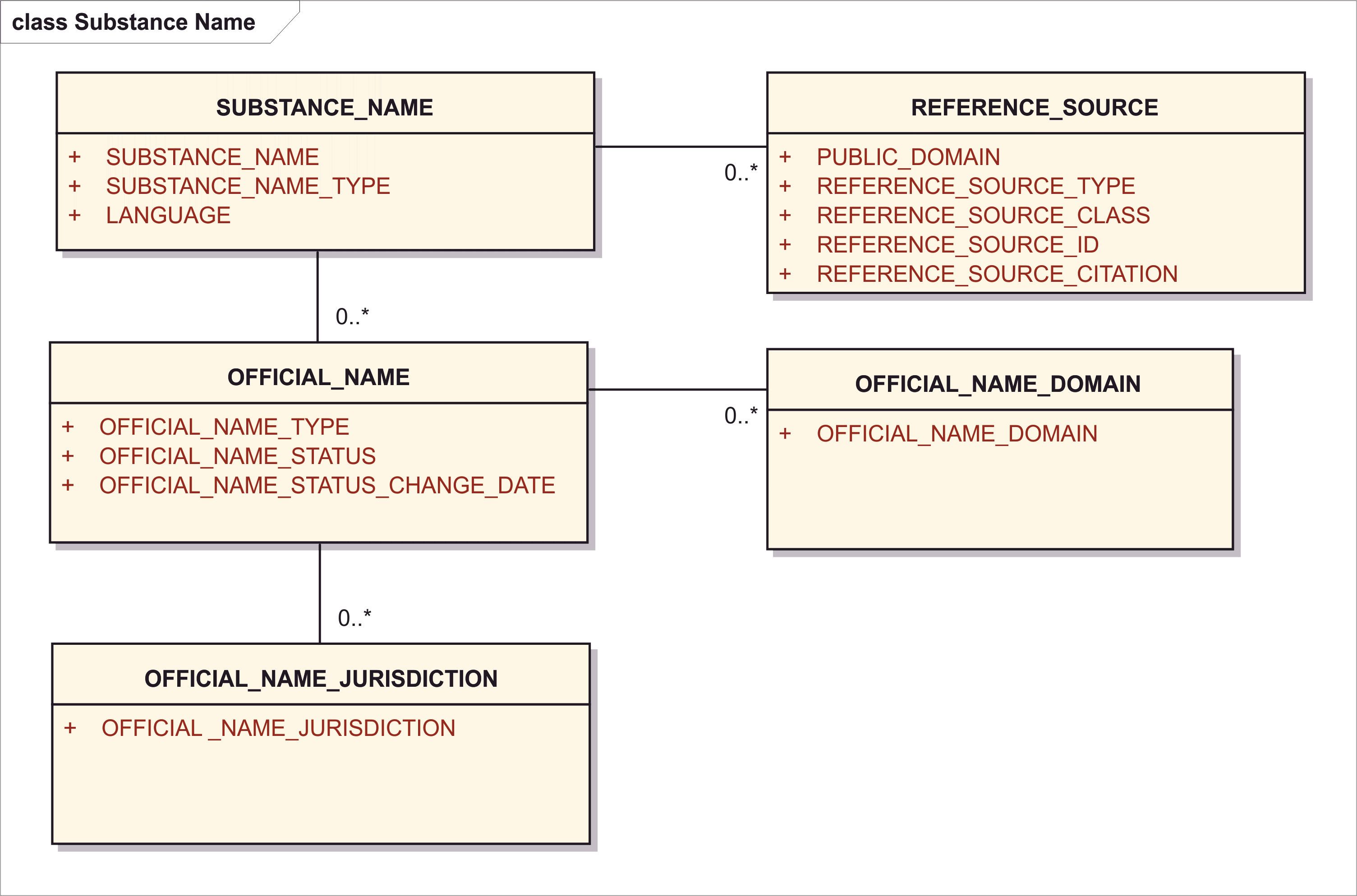


Figure 5 – Information model for Substance Names

<identifiedSubstance><!-- **Substance** -->  
 <identifiedSubstance>  
 <name>***Substance Name*** *(Preferred Name)*</name>  
 <asNamedEntity>  
 <code code="***Substance Name Type*** *(Code)*"  
 codeSystem="***Substance Name Type*** *Code System (OID)*"  
 displayName="***Substance Name Type*** *(Display Name)*"/>  
 <name xml:lang="***Language*** *Code*">***Substance Name***</name>  
 <statusCode code="***Official Name Status***"/>  
 <effectiveTime>  
 <low value="***Official Name Status Change Date*** *(if new status = active)*"/>  
 <high value="***Official Name Status Change Date*** *(if new status = terminated)*"/>  
 </effectiveTime>  
 <assigningTerritory>  
 <code code="***Official Name* *Jurisdiction*** *(Country) Code*"   
 codeSystem="2.16.840.1.113883.5.28"/>  
 <assigningTerritory>  
 <subjectOf>  
 <document>  
 <id extension="***Reference Source*** *(Document)* ***Id***"   
 root="*Reference Source Type Document Id System (OID)*"/>  
 <code code="***Reference Source Type*** *(Code)*"   
 codeSystem="*Reference Source Type Code Sytem (OID)*"  
 displayName="*Reference Source Type Display Name*"/>  
 <text>  
 <reference value="*Reference Source Type Document Id (URL)*"/>  
 </text>  
 <bibliographicDesignationText>  
 ***Reference Source Citation***</bibliographicDesignationText>  
 </document>  
 </subjectOf>  
 <subjectOf>  
 <policy>  
 <code code="***Official Name Domain*** *(Code)*" codeSystem="*Code System (OID)*"   
 displayName="*Display Name*"/>  
 </policy>  
 </subjectOf>  
 </asNamedEntity>

Note: The SPL <identifiedSubstance> Entity has a direct <name> tag used by the document author to devise on suggested preferred name. This selection of a preferred name may be entirely subjective and expresses only the choice of the author. All names with the detailed name metadata are provided also in the <asNamedEntity> element and all of its child elements.

### SubstanceName

|  |  |
| --- | --- |
| **User Guidance** | Name or company code associated with the Substance. |
| **Example(s)** | “Amoxicillin” |
| **Conformance** | REQUIRED |
| **Data Type** | ST |
| **Business Rule(s)** | Mandatory all substances shall have at least one name or company code associated with the substance. Free text, multiple names allowed, each in its own <asNamedEntity> element. |

<identifiedSubstance><!-- **Substance** -->  
 <identifiedSubstance>  
 <asNamedEntity>  
 <name>***Substance Name***</name>

### Substance Name Type

|  |  |
| --- | --- |
| **User Guidance** | Each name shall be associated with a type. Official names are names assigned by a national or international nomenclature or regulatory authority and are intended to be uniquely associated with substance. Other names could be common names or chemical names. IUPAC and CAS based names are considered Other Names and not Official Names. Brand Names are names a company assigns to the substance for use in commerce. A company code typically consists of letters and numbers and is used within a company as a codename for an investigational substance. |
| **Example(s)** | Official name, other name, brand name, company code |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **OID** | @codeSystem = 2.16.840.1.113883.3.26.1.1 |
| **Value Allowed** | To be defined by maintenance. |
| **Business Rule(s)** | A name shall have one and only one name type. |

<identifiedSubstance><!-- Substance -->  
 <identifiedSubstance>  
 <asNamedEntity>  
 <code code="***Substance Name Type*** *(Code)*"  
 codeSystem="*Substance Name Type**Code System (OID)*"  
 displayName="*Substance Name Type**(Display Name)*"/>

### Language

|  |  |
| --- | --- |
| **User Guidance** | If the name is language dependent, that language shall be specified. Company codes are not language dependent and no language should be specified for them. The language used to provide the structure substance information shall be specified according to the XML standard’s xml:lang attribute specification, which is based on ISO 639-2 alpha-2 codes |
| **Example(s)** | en – English, de – German, fr –French |
| **Conformance** | CONDITIONAL – if the name is actually language dependent, an IUPAC systematic name may not be language dependent, a company code is not language dependent. |
| **Data Type** | CD |
| **Business Rule(s)** | Multiple codes assigned to the same language are to be considered synonyms and only one of them is specified.  While xml:lang allows country-specification of language, e.g., en\_US vs. en\_UK or pt\_PT vs. pt\_BR, such specificity is to be avoided except if the country variant of the language exceptionally does give rise to an alternative spelling. In that case, only the exceptional spelling needs to be tagged with the country-specific language code, not the regular spelling for that language in all other countries.  If the same name spelling exists for multiple languages of interest, the entire name is repeated each time with its own language code. |

<identifiedSubstance><!-- **Substance** -->  
 <identifiedSubstance>  
 <asNamedEntity>  
 <name xml:lang="***Language*** *Code*">*Substance Name*</name>

### Official Name (repeat as necessary)

Each official name shall have one or more Official Name Types. The Official Name Type reflects the organization that assigns or recognizes the name associated with the substance. These names are typically non-proprietary names that are used in the labelling of pharmaceuticals. The domains and jurisdictions in which the Official Name is used are also captured. It is anticipated that the maintenance organization will track and maintain this information.

#### Official name Type

|  |  |
| --- | --- |
| **User Guidance** | Designation of which authority assigned the Official Name. All official names need to have at least one such designation. The name type is the name of the authority or authorities that have assigned or have adopted the name. |
| **Example(s)** | BAN, COSING, EP, FCC, INCI, INN, JAN, JECFA, MARTINDALE, USAN, USP. |
| **Conformance** | REQUIRED – all official names require a designation of naming authority |
| **Data Type** | II |
| **OID** | II.root for document ids of publications considered “official name type” TBD. As shown in the examples, this is envisioned as a separate identifier system. |
| **Value Allowed** | All values in the name assigning authority identifier system. |
| **Business Rule(s)** | All official names must have a naming authority designated. If multiple authorities have assigned the same name, a separate <asNamedEntity> is specified for each authority that needs to be listed. |

Any authority which assigns official names will do so in a publication, and as such, referencing such well known publication under a well known document id is sufficient indication for the name being “official” and sufficient representation for the “official name type”.

<identifiedSubstance><!-- Substance -->  
 <identifiedSubstance>  
 <asNamedEntity>  
 <name>*Substance Name*</name>  
 <subjectOf>  
 <document>  
 <id extension="*Reference Source (Document) Id /* ***Official Name Type***"   
 codeSystem="*Official Name Document Id System (OID)*"/>

Example:

<identifiedSubstance><!-- Substance -->  
 <identifiedSubstance>  
 <asNamedEntity>  
 <name>*Substance Name*</name>  
 <subjectOf>  
 <document>  
 <id extension="USP"   
 codeSystem="*Official Name Document Id System (OID)*"/>

#### Official Name Status

|  |  |
| --- | --- |
| **User Guidance** | The status of the official name. The status of the official name is associated with the name type. It is possible that an official name could be official in one authority and superseded in another. Each official name type will have a status. |
| **Example(s)** | Current, Alternate, Superceded, Proposed |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Value Allowed** | “active” – Current, “superceded” – Terminated, “pending” – Proposed. |
| **Business Rule(s)** | Each official name shall have a single status.  The fact that a name is alternate is evident by having more than one name assigned to the same substance by the same authority for the same language, domain, etc. |

<identifiedSubstance><!-- Substance -->  
 <identifiedSubstance>  
 <asNamedEntity>  
 <name>*Substance Name*</name>  
 <statusCode code="***Official Name Status***"/>

#### Official Name Status Change Date

|  |  |
| --- | --- |
| **User Guidance** | Date of official name change if known. The change date of the status of the official name shall be specified according to the ISO 8601 date format (the variant without any delimiters). |
| **Example(s)** | 20110601 |
| **Conformance** | OPTIONAL |
| **Data Type** | TS |
| **Business Rule(s)** | If official name status changes the date of the official change will be captured if known. |

The HL7 representation provides this as an effective time range beginning when the name became active and ending when the name status changes to superceded / terminated.

<identifiedSubstance><!-- Substance -->  
 <identifiedSubstance>  
 <asNamedEntity>  
 <name>*Substance Name*</name>  
 <effectiveTime>  
 <low value="***Official Name Status Change Date*** *(if new status = active)*"/>  
 <high value="***Official Name Status Change Date*** *(if new status = terminated)*"/>  
 </effectiveTime>  
 </asNamedEntity>

#### Official Name Domain (repeat as necessary)

|  |  |
| --- | --- |
| **User Guidance** | Specifies when and for what purpose the official name is to be used. All names can have at least one domain. This is useful to differentiate different names for the same substance as used for a drug active ingredient as opposed to a food color additive. |
| **Example(s)** | BIOLOGIC (FDA), BIOLOGIC (JAPAN), COSMETIC (FDA), DRUG (FDA), DRUG (JAPAN); FOOD (FDA), MEDICINE (EU) |
| **Conformance** | CONDITIONAL – if different names exist that are not appropriate to be used across all domains. |
| **Data Type** | CD |
| **OID** | codeSystem OID TBD |
| **Value Allowed** |  |
| **Business Rule(s)** | Used if different names exist that are not appropriate to be used across all circumstances. |

<identifiedSubstance><!-- Substance -->  
 <identifiedSubstance>  
 <asNamedEntity>  
 <name>*Substance Name*</name>  
 <subjectOf>  
 <policy>  
 <code code="***Official Name Domain*** *(Code)*" codeSystem="*Code System (OID)*"   
 displayName="*Display Name*"/>  
 </policy>  
 </subjectOf>  
 </asNamedEntity>

Note: in HL7 SPL the “domain” is modelled as a name usage “policy”. Any code can be provided which specifies in what circumstances the respective name is appropriate.

#### Official Name Jurisdiction (repeat as necessary)

|  |  |
| --- | --- |
| **User Guidance** | The jurisdiction of the official name shall be provided based on ISO 3166-1 alpha-2 codes with two exceptions: EL (not GR) is used to represent Greece, and UK (not GB) is used to represent the United Kingdom.  For the European Union EU shall be used. For pharmaceutically active substances at least one jurisdiction should be associated with the name. For inactive substances or excipients the jurisdiction may not be apparent or captured. |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **OID** | codeSystem = 2.16.840.1.113883.5.28 |
| **Business Rule(s)** | All the jurisdictions in which the Official Name is to be used should be indicated. |

Note: in HL7 SPL not alpha-2 but alpha-3 version of ISO 3166-1 are used.

<identifiedSubstance><!-- Substance -->  
 <identifiedSubstance>  
 <asNamedEntity>  
 <name>*Substance Name*</name>  
 <assigningTerritory>  
 <code code="***Official Name* *Jurisdiction*** *(Country) Code*"   
 codeSystem="2.16.840.1.113883.5.28"/>  
 <assigningTerritory>

Note: this is the exact same structure as the ISO 11615 7.2.1.6.1 Country and 7.2.1.6.2 Jurisdiction where the Medicinal Product Medicinal Product Name is applicable. The same comments hold: if a name is used in more than one jurisdictions which need to be specifically mentioned, it is done by repeating the name. If jurisdictions are multi-national entities (e.g., EU, ASEAN) these multi-national treaty organizations can be referred to by a single code even if it is not ISO 3166-1 country codes.

### Reference Source (repeat as necessary)

|  |  |
| --- | --- |
| **User Guidance** | All names should have at least one reference source. For official names the reference sources is the publication (compendium) of that official naming authority. Multiple references sources can also exist. The same structure can be used for all reference sources. |
| **Conformance** | REQUIRED |

<identifiedSubstance><!-- Substance -->  
 <identifiedSubstance>  
 <name>*Substance Name (Preferred Name)*</name>  
 <asNamedEntity>  
 <subjectOf>  
 <document>  
 <id extension="***Reference Source*** *(Document)* ***Id***"   
 root="*Reference Source Document Id System (OID)*"/>  
 <code code="***Reference Source Type*** *(Code)*"   
 codeSystem="*Reference Source Type Code Sytem (OID)*"  
 displayName="*Reference Source Type Display Name*"/>  
 <text>  
 <reference value="*Reference Source Type Document Id (URL)*"/>  
 </text>  
 <bibliographicDesignationText>  
 ***Reference Source Citation***</bibliographicDesignationText>  
 </document>  
 </subjectOf>

Note: reference sources may be represented by a well known document id.

#### Public Domain

|  |  |
| --- | --- |
| **User Guidance** | Indication of whether the source is in the public domain. Company codes will only be released if the defining information is associated with the code in a public source. |
| **Example(s)** | Yes, No |
| **Conformance** | REQUIRED |
| **Data Type** | BL |
| **Business Rule(s)** | If any reference source for a given name is in the public domain the name could be release provided the substance ID is in the public domain. |

Note: whether a source is in the public domain depends on the source. Notably even official name sources are not “in the public domain” as they are usually all released under Copyright.

What the “public domain” flag in the reference source is really used for is described in Section 8.1, where it defines if the “substance id is released to the public” or not. Therefore, what this serves to is as a confidentiality flag for the entire substance:

<identifiedSubstance>  
 <id extension="*Substance Id*" root="*Substance Id System (OID)*"/>  
 <confidentialityCode code="B"/>  
 <identifiedSubstance>  
 ...

#### Reference Source Type

|  |  |
| --- | --- |
| **User Guidance** | The reference source type in which the name was actually found. The use of primary sources is encouraged. |
| **Example(s)** | BP, CHEMID (NLM), EP, IND, INN, ITIS, JAN, JOURNAL, JP, KEGG, MARTINDALE, NDA, PERSONAL CARE PRODUCTS COUNCIL (PCPC), PUBCHEM, USAN, USP, WEB PAGE |
| **Conformance** | REQUIRED |
| **Data Type** | II |
| **OID** | II.root for document ids of publications considered “official name type” TBD. As shown in the examples, this is envisioned as a separate identifier system. |
| **Business Rule(s)** | Mandatory all names should have at least one source. |

<identifiedSubstance><!-- Substance -->  
 <identifiedSubstance>  
 <name>*Substance Name (Preferred Name)*</name>  
 <asNamedEntity>  
 <subjectOf>  
 <document>  
 <id extension="***Reference Source Type*** *(Document)* ***Id***"   
 root="*Reference Source Type Document Id System (OID)*"/>  
 <code code="***Reference Source Type*** *(Code)*"   
 codeSystem="*Reference Source Type Code Sytem (OID)*"  
 displayName="*Reference Source Type Display Name*"/>

Note: as apparent from the examples above, reference source “type” contains specific publications or compendia (e.g. USP, USAN, MARTINDALE). These are really one identified reference source, not a type that could be ascribed to various reference sources. So, when the reference source “type” is one of those individual publications, compendia, or databases it is represented as Reference Source Id.

When the Reference Source Type really is a Type which can be ascribed to multiple reference sources, then the Reference Source Type Code proper is used.

When the Reference Source Type is a databases (e.g., KEGG, CHEMID), and the Reference Source Id can identify one particular record in the database, then the nature of this database is encoded in the Reference Source Id’s root (*Reference Source Type Document Id System OID)*

#### Reference Source Class

|  |  |
| --- | --- |
| **User Guidance** | The reference source class in which the name was found. Useful for classification of sources. |
| **Example(s)** | Regulatory submission |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **OID** |  |
| **Value Allowed** | CLINICAL TRIAL APPLICATION, LITERATURE, OFFICIAL NAME SOURCE, REGULATORY SUBMISSION, WEB. |
| **Business Rule(s)** | Need not be entered the value is dependent on the Reference Source Type. |

#### Reference Source ID

|  |  |
| --- | --- |
| **User Guidance** | An ID associated with reference source. IND number, INN list; PMID. |
| **Conformance** | OPTIONAL |
| **Data Type** | II |
| **Business Rule(s)** | A number associated with a given source. |

<identifiedSubstance><!-- Substance -->  
 <identifiedSubstance>  
 <name>*Substance Name (Preferred Name)*</name>  
 <asNamedEntity>  
 <subjectOf>  
 <document>  
 <id extension="***Reference Source*** *(Document)* ***Id***"   
 root="*Reference Source Type Document Id System (OID)*"/>  
 <code code="*Reference Source Type (Code)*"   
 codeSystem="*Reference Source Type Code Sytem (OID)*"  
 displayName="*Reference Source Type Display Name*"/>  
 <text>  
 <reference value="*Reference Source Document Id (URL)*"/>  
 </text>

Note: when the Reference Source Id is a URL of a Web reference, the URL is placed in the text/reference/@value attribute, not the id/@extension.

#### Reference Source Citation

|  |  |
| --- | --- |
| **User Guidance** | The reference source or citation in which the name was found. Reference could be citation or a URL; the textual description of the reference source within the literature or regulatory document shall be described. |
| **Example(s)** | Ph.E. 7.1, p.34  Literature reference: Encycl. 3: 429 (1792) |
| **Conformance** | CONDITIONAL |
| **Data Type** | ST |
| **Business Rule(s)** | Mandatory all names should have at least one source. Standardized format for sources should be developed. Free Text. |

<identifiedSubstance><!-- Substance -->  
 <identifiedSubstance>  
 <name>*Substance Name (Preferred Name)*</name>  
 <asNamedEntity>  
 <subjectOf>  
 <document>  
 <bibliographicDesignationText>  
 ***Reference Source Citation***</bibliographicDesignationText>  
 <text>  
 <reference value="*Reference Source Document Id (URL)*"/>  
 </text>

Note: the HL7 SPL Reference Source Citation contains a bibliography entry. A web URL is placed in the text/reference/@value

## Substance Code (repeat as necessary)

This section of reference information allows the capture of codes that are typically associated with a substance. The capture of these codes typically facilitates the defining and mapping of substances and linking of substances to a variety of information sources. Codes should be communicated if known but are not necessary or mandatory. The maintenance organization should have the ability to capture, store and distribute codes as necessary. All the codes that are captured should be associated with a publicly recognized code system and map directly to a given substance. It should be noted that company codes are not captured in this section. The maintenance organization should make an effort to associate commonly used public codes with a given substance.

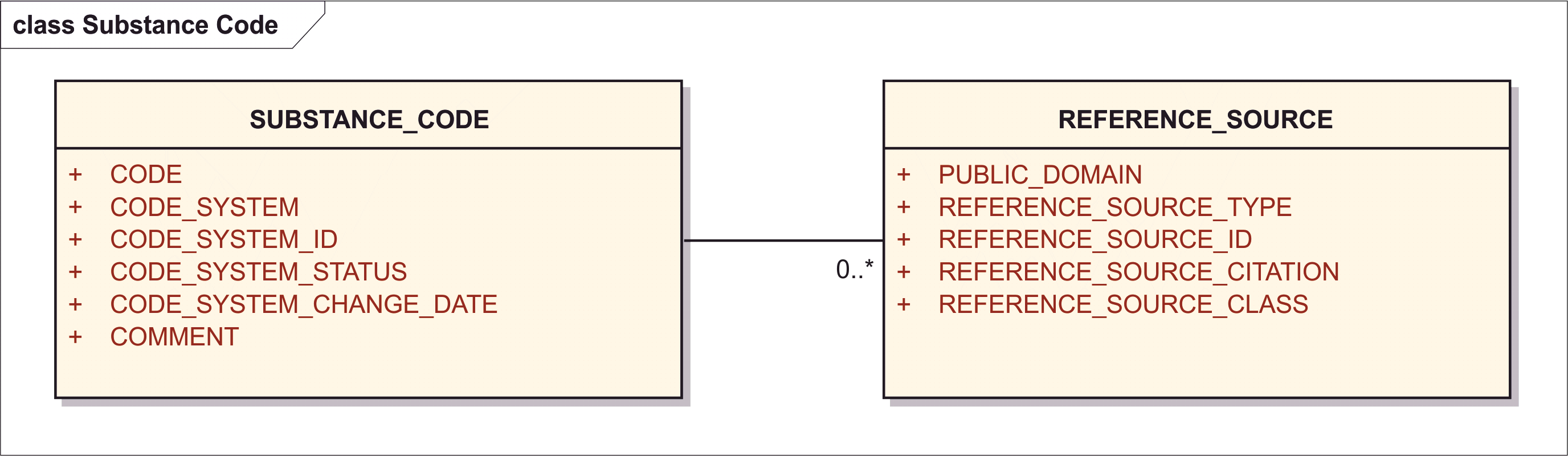


Figure 6 – Information model for Substance Codes

<identifiedSubstance>  
 <identifiedSubstance>

<asEquivalentSubstance>  
 <statusCode code="***Code System Status***"/>  
 <effectiveTime>  
 <low value="***Code System Change Date*** *(if new status = active)*"/>  
 <high value="***Code System Change Date*** *(if new status = terminated)*"/>  
 </effectiveTime>  
 <definingMaterialKind>  
 <code code="***Substance Code***" codeSystem="*Substance* ***Code System Id*** *(OID)*"/>  
 </definingMaterialKind>  
 <subjectOf>  
 <document>  
 <id extension="***Reference Source*** *(Document)* ***Id***"   
 root="*Reference Source Type Document Id System (OID)*"/>  
 <code code="***Reference Source Type*** *(Code)*"   
 codeSystem="*Reference Source Type Code Sytem (OID)*"  
 displayName="*Reference Source Type Display Name*"/>  
 <text>  
 <reference value="*Reference Source Type Document Id (URL)*"/>  
 </text>  
 <bibliographicDesignationText>  
 ***Reference Source Citation***</bibliographicDesignationText>  
 </document>  
 </subjectOf>  
 </asEquivalentSubstance>

### Code

|  |  |
| --- | --- |
| **User Guidance** | The specific code (the value) with regard to a referenced code system.  NOTE: InChI is not considered a code and shall be specified within the structure description. |
| **Conformance** | CONDITIONAL |
| **Data Type** | ST |
| **Business Rule(s)** | Only codes from recognized code systems will be captured. The code should specifically link to a substance but need not uniquely link to substance. |

<asEquivalentSubstance>  
 <definingMaterialKind>  
 <code code="***Substance Code***" codeSystem="*Substance Code System Id (OID)*"/>

### Code System ID

|  |  |
| --- | --- |
| **User Guidance** | The OID of the code system |
| **Conformance** | REQUIRED |
| **Data Type** | OID |
| **Business Rule(s)** | Conditionally mandatory all codes must be linked to a code system. |

<asEquivalentSubstance>  
 <definingMaterialKind>  
 <code code="*Substance Code*" codeSystem="*Substance* ***Code System Id*** *(OID)*"/>

### Code System

|  |  |
| --- | --- |
| **User Guidance** | Name of the code system |
| **Example(s)** | CAS, EINECS, NSC, ASK, ATC, NDF-RT; RX-CUI ETC. For every Code there must be a code system. |
| **Conformance** | OPTIONAL – DISCOURAGED. Codes Systems must always be specified by their unique OID. Code System Name is superfluous and potentially confusing because people tend to try to convey something different in it that is not implied by the OID. |

<asEquivalentSubstance>  
 <definingMaterialKind>  
 <code code="*Substance Code*" codeSystem="*Substance* ***Code System Id*** *(OID)*"  
 codeSystemName="***Code System Name***"/>

### Code System Status

|  |  |
| --- | --- |
| **User Guidance** | The status of the code assignment |
| **Example(s)** | current, proposed, alternate, superceded |
| **Conformance** | REQUIRED with default |
| **Data Type** | CD |
| **Value Allowed** | “active” – Current, “superceded” – Terminated, “pending” – Proposed |
| **Business Rule(s)** | Conditionally mandatory all codes have a status, but the default is “current”. |

<asEquivalentSubstance>  
 <statusCode code="***Code System Status***"/>

Note: this is not the status of the code in the system, but the status of assignment of this code to the substance.

### Code System Status Change Date

|  |  |
| --- | --- |
| **User Guidance** | The date at which the code status is changed by the maintenance organization. |
| **Example(s)** | 20110601 |
| **Conformance** | OPTIONAL |
| **Data Type** | TS.DATE |
| **Business Rule(s)** | This applies only when the code is changed, so it is not applicable for initial submission.  Since codes are not official the date captured is more of a bookkeeping requirement. |

The HL7 representation provides this as an effective time range beginning when the name became active and ending when the name status changes to superceded / terminated.

<asEquivalentSubstance>  
 <effectiveTime>  
 <low value="***Code System Change Date*** *(if new status = active)*"/>  
 <high value="***Code System Change Date*** *(if new status = terminated)*"/>  
 </effectiveTime>

### Comment

|  |  |
| --- | --- |
| **User Guidance** | Any comment can be provided in this field, if necessary. |
| **Conformance** | OPTIONAL – DISCOURAGED |
| **Data Type** | ST |

Note: commenting on code assignments should not be specially necessary. If it is, such comment can always be packaged into a characteristic:

<asEquivalentSubstance>  
 <subjectOf>  
 <characteristic>  
 <text>***Comment*** *Text*</text>  
 </characteristic>  
 </subjectOf>

### Reference Source (repeat as necessary)

The reference source and related information will be captured according to Section 8.3.5. Reference sources are required when a code is entered.

## Version (repeat as necessary)

This section shall provide information on the version of the substance. Where no “Version Number” and “Effective Date” have been assigned by an authority source, the version number shall be set as 0 with the date set as the date of initial submission of the substance.

<document>  
 <id root="***Document Id***"/>  
 <code code="64124-1" codeSystem="2.16.840.1.113883.6.1"  
 displayName="Indexing - Substance"/>  
 <effectiveTime value="***Effective Date***"/>  
 <setId root="***Document Set Id***"/>  
 <versionNumber value="***Version Number***"/>  
 <author>  
 <noteText>***Changes Made***</noteText>  
 <!-- ***author information*** -->  
 </author>  
 <component>  
 <structuredBody>  
 <component>  
 <section>  
 <subject>  
 <identifiedSubstance><!-- **Substance** -->

### Version Number

|  |  |
| --- | --- |
| **User Guidance** | The number of the version of the substance shall be provided. |
| **Example(s)** | 1, 2, 3, 99 |
| **Conformance** | REQUIRED |
| **Data Type** | INT |
| **Business Rule(s)** | If same document set exists, version must be greater than the last submitted version |

<document>  
 <id root="***Document Id***"/>  
 <setId root="***Document Set Id***"/>  
 <versionNumber value="***Version Number***"/>

Each document version has a globally unique identifier. In addition a *version number* which is relative to the document version-set.

NOTE: the colloquial notion of a single document which has multiple versions each different from the other does not exist in HL7. Each modification of a document is itself a new document. However, in HL7 the *document set*, or, for additional clarity the “*document version set*” is recognized in HL7 and reflected in a “set id”, which is a globally unique identifier for the entire set of versions of what might colloquially be called “one document”. Thus, in summary, any change to a document produces a new document with a new document id, but with the same set id as the previous document and an increased version number.

### Effective date

|  |  |
| --- | --- |
| **User Guidance** | The date when the substance was effective shall be provided in line with the ISO 8601 date format. This shall be defined when the substance is generated or modified. |
| **Example(s)** | 20110219 |
| **Conformance** | OPTIONAL |
| **Data Type** | TS |
| **Business Rule(s)** |  |

<document>  
 <effectiveTime value="***Effective Date***"/>

### Change Made

|  |  |
| --- | --- |
| **User Guidance** | The description of the updates or changes of the substance shall be specified. The field will be left empty when first insert of substance. |
| **Conformance** | OPTIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<document>  
 <author>  
 <noteText>***Changes Made***</noteText>

## Reference Information (repeat as necessary)

The following information is captured for reference purposes only. The information is typically not used in defining a substance but is important information that can be submitted and retained by the maintenance organization.

### Comment

|  |  |
| --- | --- |
| **User Guidance** | Any comment can be provided in this field, if necessary. |
| **Conformance** | OPTIONAL – DISCOURAGED |
| **Data Type** | ST |

Note: general commenting in a field should not be specially necessary. If it is, such comment can always be packaged into a characteristic:

<identifiedSubstance>  
 <identifiedSubstance .../>  
 <subjectOf>  
 <characteristic>  
 <text>***Comment*** *Text*</text>  
 </characteristic>  
 </subjectOf>

### Substance Classification (repeat as necessary)

There are a number of systems developed to classify substances. These systems are useful for grouping compounds that either have similar chemical structures or similar effects when administered in a medicinal product. The standard should have the ability to capture a variety of classification systems. Classification systems are typically based on molecular structure, chemical properties, pharmacological effects, mechanism of action, therapeutic targets or indication.

The standard shall have the ability to capture multiple classifications and variable levels of classification. Although most classification will be associated with an external classification system the maintenance organization will have the ability to classify substances as needed.

The following information shall be provided:

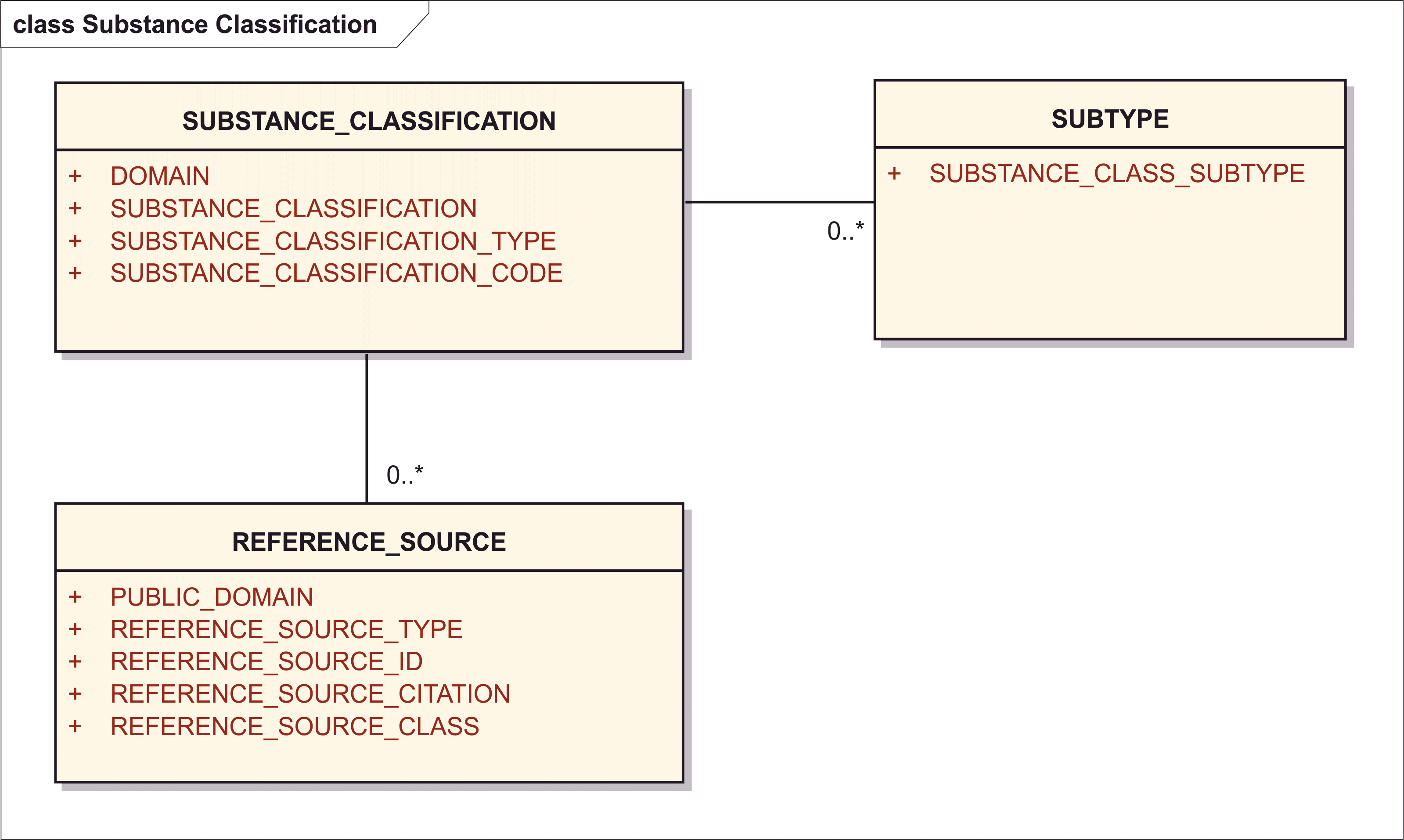


Figure 7 – Information model for Substance Classification

<identifiedSubstance>  
 <identifiedSubstance>

<asSpecializedKind>  
 <definingMaterialKind>  
 <code code="***Substance Classification*** *Code*"   
 codeSystem="*Substance Classification System (OID)*"/>  
 </definingMaterialKind>  
 <subjectOf>  
 <policy>  
 <code code="*Substance Classification* ***Domain*** *(Code)*"   
 codeSystem="*Code System (OID)*"   
 displayName="*Display Name*"/>  
 </policy>  
 </subjectOf>  
 <subjectOf>  
 <document>  
 <id extension="***Reference Source*** *(Document)* ***Id***"   
 root="*Reference Source Type Document Id System (OID)*"/>  
 <code code="***Reference Source Type*** *(Code)*"   
 codeSystem="*Reference Source Type Code Sytem (OID)*"  
 displayName="*Reference Source Type Display Name*"/>  
 <text>  
 <reference value="*Reference Source Document Id (URL)*"/>  
 </text>  
 <bibliographicDesignationText>  
 ***Reference Source Citation***</bibliographicDesignationText>  
 </document>  
 </subjectOf>  
 </asSpecializedKind>

#### Domain

|  |  |
| --- | --- |
| **User Guidance** | The domain of the substance classification shall be provided. |
| **Example(s)** | Human pharmaceuticals, human vaccine, animal drug, animal vaccine, food, food additive, colorant, pesticide; tobacco additive; flavour, excipient |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>

<asSpecializedKind>  
 <definingMaterialKind .../>  
 <subjectOf>  
 <policy>  
 <code code="*Substance Classification* ***Domain*** *(Code)*"   
 codeSystem="*Code System (OID)*"   
 displayName="*Display Name*"/>

#### Substance Classification (Code System)

|  |  |
| --- | --- |
| **User Guidance** | Substance Classification terms based on a controlled vocabulary that is controlled by the system or type of substance classification. |
| **Example(s)** | Single, muxture |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Value Allowed** | ATC, NDF-RT, NCI, AHPA, MESH |
| **Business Rule(s)** | The system should be capable of capturing a variety of classification system and each system should be identified. All active ingredients are typically classified in multiple systems. The maintenance organization may also create its own classification system. |

<identifiedSubstance>  
 <identifiedSubstance>

<asSpecializedKind>  
 <definingMaterialKind>  
 <code code="*Substance Classification Code*"   
 codeSystem="***Substance Classification*** *Code System**(OID)*"/>

#### Substance Classification Code

|  |  |
| --- | --- |
| **User Guidance** | Code from the classifying code system. |
| **Example(s)** | R03CC – Adrenergics for Systemic Use (ATC), N0000009922 - Adrenergic beta2-Agonists (NDF-RT) |
| **Conformance** | OPTIONAL |
| **Data Type** | II |
| **Business Rule(s)** | The classification code will be part of an internationally recognized classification system. |

<identifiedSubstance>  
 <identifiedSubstance>

<asSpecializedKind>  
 <definingMaterialKind>  
 <code code="***Substance Classification******Code***"   
 codeSystem="*Substance Classification System (OID)*"/>

#### Substance Classification – Type

|  |  |
| --- | --- |
| **User Guidance** |  |
| **Example(s)** |  |
| **Conformance** | IMPLICIT – by Substance Classification – Subtype |
| **Data Type** | CD |
| **Business Rule(s)** |  |

#### Substance Classification – Subtype (repeat as necessary)

|  |  |
| --- | --- |
| **User Guidance** | Many classification systems will have multiple subtypes of classification this field will attempt to capture this information. |
| **Example(s)** | NDF-RT Mechanism of Action (MOA) or Physiological Effect. |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<asSpecializedKind>  
 <code code="***Substance Classification - Subtype*** *Code*"   
 codeSystem="*Substance Type Code System (OID)*"/>

#### Reference Source (repeat as necessary)

The reference source and related information will be captured according to Section 8.2.5. Reference sources are required for all classifications. The source should preferably be a primary source.

### Target

The Target section allows the capture of information related to the targets that a substance interacts with. These targets can therapeutic, metabolic, interactions postulated to result in toxicity or interactions of unknown effect.

Ideally the target molecular entity is a substance with its own substance id, in this case the format is as follows:

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>  
 <interactsIn>  
 <interaction>  
 <code code="***Interaction Type*** *Code*" codeSystem="*Code System (OID)*"/>  
 <interactor>  
 <functionCode code="***Target Class*** *Code*"  
 codeSystem="*Target Class Code System (OID)*"   
 displayName="*Target Class Name*"/>  
 <identifiedSubstance>  
 <id extension="***Target Substance Id***"   
 root="*Substance Id System (OID)*">  
 <identifiedSubstance>  
 <code code="***Target Id***"   
 codeSystem="*Target Id System (OID)*"/>  
 <name>***Target Name*** *(one preferred/primary name)*</name>  
 <asNamedEntity>  
 <name>***Target Name***</name>  
 <!-- full detail about the name -->  
 </asNamedEntity>  
 </identifiedSubstance>  
 </identifiedSubject>  
 </interactor>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Target Type’*" codeSystem="*Code Sytem (OID)*"/>  
 <value code="***Target Type*** *Code*" xsi:type="CE"  
 codeSystem="*Code Sytem (OID)*"/>

The following information shall be provided or captured where applicable:

#### Target (Gene) ID

|  |  |
| --- | --- |
| **User Guidance** | ID of the target molecular entity upon which the substance acts. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | One Target Id should be specified. The target should be a molecular entity or a family of molecular entities. The target is usually a protein entity that drug actually interacts with. Both therapeutic and metabolic targets and targets associated with toxicity should be captured. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>  
 <interactsIn>  
 <interaction>  
 <interactor>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Target Id***"   
 codeSystem="*Target Id System (OID)*"/>

When the Target has a proper Substance Id of its own, one can make a simple reference to that Substance Id only without the need to specify any further detail about it.

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>  
 <interactsIn>  
 <interaction>  
 <interactor>  
 <identifiedSubstance>  
 <id extension="***Target Substance Id***"   
 root="*Substance Id System (OID)*">

#### Target (Gene) Name

|  |  |
| --- | --- |
| **User Guidance** | Name of the target molecular entity upon which the substance acts. |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | A variety of names can be captured. A single name will be a preferred name and associated with the target id. Each target should be associated with a target name. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>  
 <interactsIn>  
 <interaction>  
 <interactor>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <name>***Target Name*** *(one preferred/primary name)*</name>  
 <asNamedEntity>  
 <name>***Target Name***</name>  
 <!-- full detail about the name -->  
 </asNamedEntity>

At least one preferred name should be specified and for most cases that is sufficient.

<interactor>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <name>***Target Name*** *(one preferred/primary name)*</name>

When multiple names are required, with all the metadata available about substance named, that full repeated <asNamedEntity> structure may be used as described in 8.2Secgtion 8.2 上の.

<asNamedEntity>  
 <name>***Target Name***</name>  
 <!-- full detail about the name -->  
 </asNamedEntity>

#### Interaction Type

|  |  |
| --- | --- |
| **User Guidance** | Type of interaction between the substance and the target. |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Value Allowed** | AGONIST, INDUCER, INHIBITOR, PARTIAL AGONIST, SUBSTRATE, PRODUCT |
| **Business Rule(s)** | The interaction type will be captured for all substrate-target interactions. A substance can have multiple types of interactions. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>  
 <interactsIn>  
 <interaction>  
 <code code="***Interaction Type*** *Code*" codeSystem="*Code System (OID)*"/>  
 <interactor>  
 <identifiedSubstance>  
 <id extension="***Target Substance Id***"   
 root="*Substance Id System (OID)*">  
 <identifiedSubstance>  
 <code code="***Target Id***"   
 codeSystem="*Target Id System (OID)*">  
 <name>***Target Name*** *(one preferred/primary name)*</name>  
 <asNamedEntity>  
 <name>***Target Name***</name>  
 <!-- full detail about the name -->  
 </asNamedEntity>  
 </identifiedSubstance>  
 </identifiedSubject>  
 </interactor>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Target Type’*" codeSystem="*Code Sytem (OID)*"/>  
 <value code="***Target Type*** *Code*" xsi:type="CE"  
 codeSystem="*Code Sytem (OID)*"/>

#### Target Organism Type

|  |  |
| --- | --- |
| **User Guidance** | The organism type for which the active substance is targeted. |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Value Allowed** | HUMAN, VIRAL, BACTERIAL, FUNGAL, INSECT, HELMINTH, MALARIAL |
| **Business Rule(s)** | At least one organism type should associated. |

The requirement for Organism Id is to go together with Gene Id as the Genome Id. For example, in Entrez Gene the gene id number or the gene symbol is not unique across all genomes but only within one genome / organism, so the same Entrez Gene Id number means different things depending on which genome. This is less an information requirement than a technical / terminology representation requirement. If each gene (or indeed mature target protein, which is often a complex and not derived from a single gene) were to be represented by a single terminologic concept, there would be no need to specify the Target Organism. The lack of proper terminology code assignments by many Bioinformatics resources can be addressed in further implementation guidance. For example, for NCBI Entrez Gene Ids one would put the NCBI Organism Id as the last component in the codeSyste OID. For example, if 1.2.3.99.100 was an OID dedicated to NCBI Gene Id representations, one would put the NCBI organism/genome/taxonomy id 9606 for Homo sapiens at the end of that OID: 1.2.3.99.100.9606. This way all NCBI gene ids are unique across all genomes.

If the target is an organism as a whole, such as antibiotics on a bacteria when its target protein is not specified (even though it could and should be specified in most cases, e.g., for penicillin antibiotics it is murin synthase, not the entire bacterium), in any such case where the entire organism is considered the target, that organism itself has an id and becomes the Target Id itself.

#### Target Type

|  |  |
| --- | --- |
| **User Guidance** | Type is a limited form of classification associated with a target of an active substance. |
| **Example(s)** | THERAPEUTIC, METABOLIC, TOXIC |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Each target should be associated with a target type when the target can be classified, Some targets can be both toxic and therapeutic the choice is dependent on the dose that is of a given substance. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>  
 <interactsIn>  
 <interaction>  
 <interactor>  
 </interactor>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Target Type’*" codeSystem="*Code Sytem (OID)*"/>  
 <value code="***Target Type*** *Code*" xsi:type="CE"  
 codeSystem="*Code Sytem (OID)*"/>

#### Reference Source

The reference source and related information will be captured according to Section 8.2.5. Reference sources are required when a target is entered.

NOTE: This reference source is for the source that references the association of substance with target not the source of the target name or pathway.

### Gene

The Gene section will only be captured for proteins and nucleic acids. It does not apply for chemicals and specified substances. It is only used AS reference information and refers to the gene that is the origin of the substance (protein).

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Transcription’*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Gene Id***"   
 codeSystem="*Gene Id System (OID)*">  
 <name>***Gene Name***</name>  
 </identifiedSubstance>  
 </identifiedSubject>  
 </interactor>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Gene Sequence Origin’*"   
 codeSystem="*Code Sytem (OID)*"/>  
 <value code="***Gene Sequence Origin*** *Code*" xsi:type="CE"  
 codeSystem="*Code Sytem (OID)*"/>

When the gene section applies, the following information can be provided:

#### Gene Sequence Origin

|  |  |
| --- | --- |
| **User Guidance** | Common name for the organism from which the final gene sequence originated.  NOTE: it applies primarily to proteins. |
| **Example(s)** | Human, bovine, olive |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Typically the GENE ID from the NCBI Gene database. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Transcription’*"   
 codeSystem="*Code System (OID)*"/>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Gene Sequence Origin’*"   
 codeSystem="*Code Sytem (OID)*"/>  
 <value code="***Gene Sequence Origin*** *Code*" xsi:type="CE"  
 codeSystem="*Code Sytem (OID)*"/>

#### Gene ID

|  |  |
| --- | --- |
| **User Guidance** | ID associated with the gene of origin |
| **Example(s)** | YP\_299723.1 |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Typically the GENE ID from the NCBI Gene database. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Transcription’*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Gene Id***"   
 codeSystem="*Gene Id System (OID)*">

#### Gene Name

|  |  |
| --- | --- |
| **User Guidance** | Complete name given for a gene |
| **Example(s)** | hIL-10 gene |
| **Conformance** | CONDITIONAL |
| **Data Type** | ST |
| **Business Rule(s)** | Every gene which has an ID should also have a name |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Transcription’*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <name>*Gene Name*</name>

#### Reference Source

The reference source and related information will be captured according to Section 8.2.5. The GENE database from NCBI

### Gene Elements

Gene Elements will be captured for genes that are used in gene therapy.

When the gene element section applies, the following information shall be provided:

For HL7 SPL representation a Gene Element is simply a Substance, either alone:

<identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Gene Element*** *Substance* ***Id***"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>***Gene Element Name***</mame>  
 <asSpecializedKind>  
 <definingMaterialKind>  
 <code code="***Gene Element Type*** *Code*"   
 codeSystem="*Gene Element Type Code System (OID)*"/>

or as a part of a complex Substance containing multiple Gene Elements a Vector, etc.

<moiety>  
 <partMoiety>  
 <code code="***Gene Element*** *Substance* ***Id***"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>***Gene Element Name***</mame>  
 <asSpecializedKind>  
 <definingMaterialKind>  
 <code code="***Gene Element Type*** *Code*"   
 codeSystem="*Gene Element Type Code System (OID)*"/>

#### Gene Element Type

|  |  |
| --- | --- |
| **User Guidance** | Type of the gene element |
| **Example(s)** | enhancer, promoter, silencer, terminator, repressor |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Should be captured for all genes that are transduced into cells and are intended to be expressed. |

<identifiedSubstance>  
 <identifiedSubstance>  
 <asSpecializedKind>  
 <definingMaterialKind>  
 <code code="***Gene Element Type*** *Code*"   
 codeSystem="*Gene Element Type Code System (OID)*"/>

#### Gene Element ID

|  |  |
| --- | --- |
| **User Guidance** | Unique identifier for gene element. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | A unique identifier will be associated with each gene element. |

<identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Gene Element*** *Substance* ***Id***"   
 codeSystem="*Substance Id System (OID)*"/>

#### Gene Element Name

|  |  |
| --- | --- |
| **User Guidance** | Specific gene element name. |
| **Example(s)** | SV40 Enhancer |
| **Conformance** | CONDITIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <name>***Gene Element Name***</mame>

#### Reference Source

The reference source and related information will be captured according to Section 8.2.5. One or more reference sources could be cited.

### Substance Relationship

Relationships between substances such as the relationship between a salt form and its active moiety or the parent substance should be captured. These relationships are often essential to the description of medicinal products, the basis of strength and the classification of substances. These relationships are often obvious and rules will be developed for specifying substances involved in each type of relationship. For example the active moiety of all sodium salts would be the free acid, conversely the active moiety of a hydrochloride salt would be a free base.

#### Relationship

|  |  |
| --- | --- |
| **User Guidance** | Relationship between substances.  NOTE: Target and gene elements are not to be specified in this section. |
| **Example(s)** | parent, active moiety, salt, radiolabeled, prodrug, metabolite, enantiomer |
| **Conformance** | Conditional by Region |
| **Data Type** | CD |
| **Business Rule(s)** |  |

The relationship is not a mere code but the actual connection between the two substances which may be different for different kinds of relationships:

Prodrug, Metabolite are related through one or more chemical reactions (which may all be summarized in a single complex reaction, pathway). We need to distinguish between reactions which lead to the current substance (e.g., prodrug metabolized to this substance) or lead to the other substance (e.g., this substance, which might be a prodrug, metabolized to other substance(s)).

Leading to the current substance (e.g., prodrug metabolized to this substance):

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>

<interactsIn typeCode="PRD">  
 <interaction>  
 <code code="*Code meaning ‘Metabolization’*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <identifiedSubstance>  
 <id extension="*Related Substance Id*"   
 root="*Substance Id System (OID)*">

Leading to the other substance (e.g., this substance, which might be a prodrug, metabolized to other substance(s)):

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>

<interactsIn typeCode="CSM">  
 <interaction>  
 <code code="*Code meaning ‘Metabolization’*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="PRD">  
 <identifiedSubstance>  
 <id extension="*Related Substance Id*"   
 root="*Substance Id System (OID)*">

Active moiety is specified as a moiety of the substance:

<identifiedSubstance>  
 <identifiedSubstance>  
 <!-- Substance Definition -->  
 <moiety typeCode="ACTM">  
 <partMoiety>  
 <code code="*Related Substance Id*"   
 codeSystem="*Substance Id System (OID)*"/>

#### Interaction Type

|  |  |
| --- | --- |
| **User Guidance** | Type of interaction between the substance and the target. |
| **Example(s)** | Covalent, agonist, inducer, inhibitor, partial agonist, substrate, product |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | The interaction type will be captured for all substrate-target interactions. A substance can have multiple types of interactions. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>

<interactsIn typeCode="CSM/PRD/DIR">  
 <interaction>  
 <code code="*Interaction Type Code*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM/PRD/DIR">

The choice of interactsIn/@typeCode is as follows:

* Covalent bond – not an interaction but a structure
* Agonist, Inducer, Inhibitor – DIR (direct target) of DIR (direct target)
* Substrate – CSM (consumable) of DIR (direct target)
* Product – PRD (product) of DIR (direct target)

#### Substance ID

|  |  |
| --- | --- |
| **User Guidance** | Substance ID associated with related substance |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | All related substances should have a substance id. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>

<interactsIn>  
 <interaction>  
 <interactor>  
 <identifiedSubstance>  
 <id extension="*Substance Id*"   
 root="*Substance Id System (OID)*">

#### Substance Name

|  |  |
| --- | --- |
| **User Guidance** | Preferred term or primary name of related substance. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | The preferred term of the related substance is captured; mandatory when there is a “Substance Relationship”. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>

<interactsIn>  
 <interaction>  
 <interactor>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <name**>***Substance Name</name>*

#### Amount

For some substance relationships it may be desirable to capture a quantitative or semi-quantitative amount that may further define the relationship (i.e. major metabolite, degradent etc.). Amount shall be captured according to Section 8.8.

Substance Relationships which refer to reactions (e.g., metabolism, prodrug) are quantified by specifying the quantity of the reactants:

<interactsIn>  
 <quantity><!—Quantity of This Substance --></quantity>  
 <interaction>  
 <interactor>  
 <quantity><!-- Quantity of Related Substance --></quantity>

Substance Relationships which are structural (e.g., moiety) are quantified in the moiety relationship quantity as a ratio:

<identifiedSubstance>  
 <moiety>  
 <quantity>  
 <numerator><!-- Quantity of Related Substance --></numerator>  
 <denominator><!-- Quantity of This Substance --></denominator>  
 </quantity>

#### Reference Source

The reference source will be captured according to Section 8.2.5 when necessary.

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Definition --></identifiedSubstance>

<interactsIn>  
 <interaction>  
 <subjectOf>  
 <document>  
 <!-- **Reference Source** -->

<identifiedSubstance>  
 <identifiedSubstance>  
 <!-- Substance Definition -->  
 <moiety typeCode="ACTM">  
 <subjectOf>  
 <document>  
 <!-- **Reference Source** -->

8.x Incomplete definition

If some required information is missing in the substance definition, the definition should be marked as incomplete.

<subjectOf>

     <characteristic>

      <code code="*Code* for ‘*Incomplete Definition’*" codeSystem=" *Code System (OID)*"  displayName="incomplete definition"/>

     </characteristic>

</subjectOf>

## Single Substance Type

## Structure (repeat as necessary)

Structural information is an essential element for chemical substances and for other types of substances that have structurally definable modifications. The representation should contain structural information in one or more of the standardized formats as indicated below. A graphical image of the molecular structure should also be provided if available. A complete representation of the structure should be provided and is usually sufficient to define chemical substances.

Although most chemical substances can only be represented unambiguously by a single structural representation, there are also many substances that can be represented by multiple structural representations or as a mixture or single substance. For the initial implementation and the transmittal of data for substance definition an approach that limits ambiguity with the least number of structural representions is recommended. There are several examples below that illustrate process in choosing a defining structural representation. A more detailed structure representation guide will provided as a separate document. Below are several examples that illustrate the current approach to choosing defining structural representations.

Salts: Amine salts: with the exception of ammonium and quarternium salts shall be represented as a free base.

Example: [DIFLOXACIN HYDROCHLORIDE](http://srsid.fda.gov:9618/ePS/ViewUpdateIngredient.do?calledFrom=complexSearchSubstance&BDNUM=0145879AA)



Metal and ammonium salt salts: All metal and ammonium salts shall be represented in a charged configuration with all equivalent functional groups ionized. Charge balance will be achieved when necessary by adding hydrogen ions (protons).

Example: MONOSODIUM CITRATE



Racemic mixtures will be represented as 1:1 mixtures of enantiomers.

Substances can exists in intraconverting forms or with a different defined structure depending on the physical state. In cases where the actual structure is a single configuration in a solid physical state that structure should be captured. When a substance exists in a liquid state and can intraconvert to a variety of structures the least ambigious single structure should be chosen

Example DEXTROSE (GLUCOSE): Exists as three separate substances in the crystalline state. For a crystalline substance one of the following representations should be chosen.



In the liquid state dextrose and most other monosaccharides exist as a mixtures of interconverting substances as illustrated in the five structures below.



In this case the linear single representive structure is chosen to represent dextrose in the liquid or amorphous state. If pyranose for furanose forms were chosen the only single representative structure would have on site of stereochemical ambigiuty

*DEXTROSE (liquid) structural representation:*



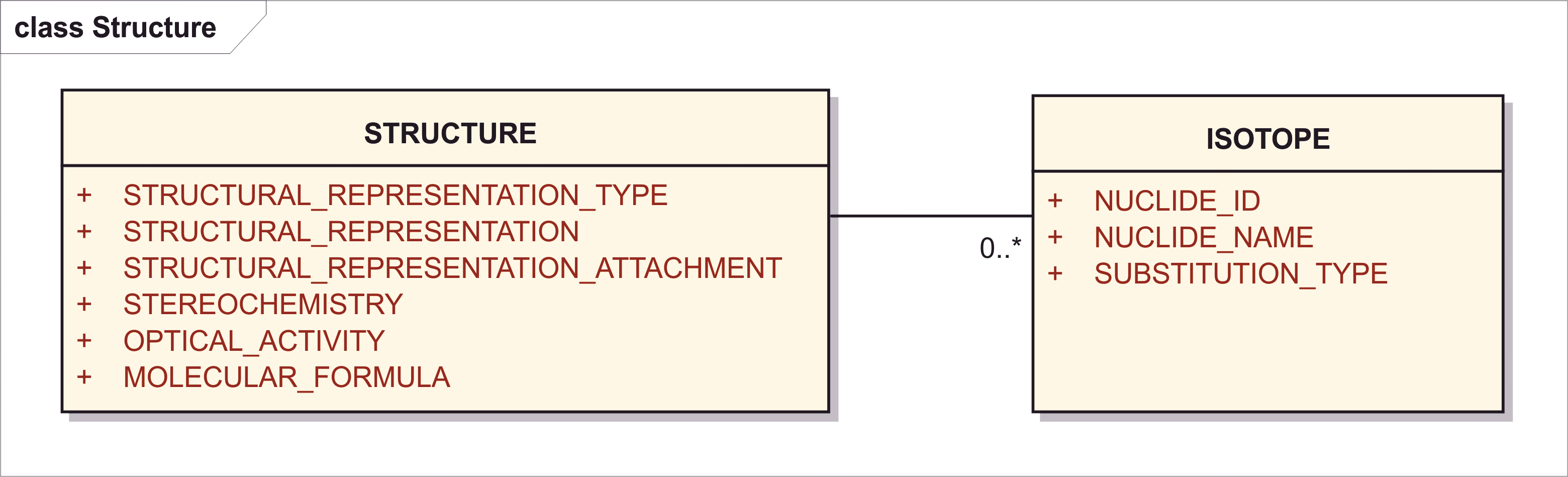


Figure 8 – Information model for Structure and Isotope

A structural representation is an essential element for defining chemical substance and is sufficient for defining most chemical substances.

Four common structural representation types are listed below.Each of the types can unambiguously describe substances with a definite molecular structure. Stereochemistry should be completely defined in each structural representation. Only MOLFILES and CDX have the ability to describe polymeric repeating units and molecular fragments. There are a number of tools available, both free and commercial, that allow the creation of molecular structural representations. Multiple representations can be submitted.

The structure is associated with one moiety, where in HL7 the notion of “moiety” is used as defined by the IUPAC Gold Book, which is the authority of international chemical nomenclature, and can be paraphrased as any sub-structure of the compound. We may also paraphrase “moiety” for the purpose presented here as “structural unit”. As a matter of convention, every structural unit shall be presented in one moiety.

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Moiety Details --></partMoiety>

<subjectOf>  
 <characteristic>  
 <code code="*Structure Type**Code*" codeSystem="*Code System (OID)*"  
 displayName="Structure Type Name"/>  
 <value xsi:type="ED" mediaType="***Structural Representation Type***">  
 <!-- Structural Representation -->  
 </value>  
 </characteristic>  
 </subjectOf>

In simple chemicals that are defined based on a simple structure in a single structural unit, there would be no need to insert a <moiety> element to represent that structural unit separate from the substance itself. However, the advantage of this convention is that it leads to a uniformity of representation. All structurally defined substances have one or more structural units (moieties.)

Note: the difference between MOLFILE V2000 and V3000 is represented inside the data itself. Structure information that is textual in nature should be encapsulated right inside the <value> tag. When binary images are provided, this can be done using the external image file reference:

A pictorial or graphical image of the molecular can also be attached and will assist in the unambiguous definition of a substance.

<value>  
 <reference value="*Structural Representation Attachment File Name*"/>

### Structural Representation Type

|  |  |
| --- | --- |
| **User Guidance** | Field indicates the type of structure. A FULL structure is a structure in which the complete connectivity of the substance is known and the substance is monodisperse. A PARTIAL structure is a structure in which either complete connectivity or stereochemistry is not defined. A REPRESENTATIVE structure is used when connectivity the underlying material is diverse and a single structure is needed to represent the underlying material |
| **Example(s)** | Full, partial, representative, fragment, ionic moiety |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | A structure can only be of one type. Field is mandatory when structural elements are present |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Moiety Details --></partMoiety>

<subjectOf>  
 <characteristic>  
 <code code="*Structure Type**Code*" codeSystem="*Code System (OID)*"  
 displayName="*Structure Type Name*"/>  
 <value xsi:type="ED" mediaType="***Structural Representation Type***">  
 <!-- Structural Representation -->  
 </value>  
 </characteristic>  
 </subjectOf>

Structural Representation Type is specified as a MIME Media Type code:

|  |  |
| --- | --- |
| Type | MIME Media Type |
| InCHI | application/x-inchi |
| MOLFILE | application/x-mdl-molfile |
| SMILES | application/x-smiles |
| CHIME | application/x-mdl-chime |
| Amino Acid Letter Sequence | application/x-aa-seq |
| DNA Sequence | application/x-dna-seq |
| RNA Sequence | application/x-rna-seq |

### Structural Representation

|  |  |
| --- | --- |
| **User Guidance** | The structural representation shall be provided for text strings, the following formats are acceptable InChI, SMILES, MOL FILE, CDX. |
| **Example(s)** |  |
| **Conformance** | REQUIRED for InChI /OPTIONAL for other formats |
| **Data Type** | ST |
| **Value Allowed** | INCHI, MOLFILE SMILES, CDX |
| **Business Rule(s)** | Each structural representation will have one and only one type. |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Moiety Details --></partMoiety>  
 <subjectOf>  
 <characteristic>  
 <code code="*Structure Type Code*" codeSystem="*Code System (OID)*"  
 displayName="*Structure Type Name*"/>  
 <value xsi:type="ED" mediaType="*Structural Representation Type*">  
 <!-- **Structural Representation** -->  
 </value>

### Structural Representation Attachment

|  |  |
| --- | --- |
| **User Guidance** | An attached file that may contain the structural representation shall be provided for complex data formats, e.g. CDX, MOLFILE. |
| **Example(s)** |  |
| **Conformance** | OPTIONAL |
| **Data Type** | ED |
| **Business Rule(s)** | A graphical representation of the substance can be provided. |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Moiety Details --></partMoiety>  
 <subjectOf>  
 <characteristic>  
 <code code="*Structure Type Code*" codeSystem="*Code System (OID)*"  
 displayName="*Structure Type Name*"/>  
 <value xsi:type="ED" mediaType="*Structural Representation Type*">  
 <reference value="***Structural Representation Attachment File Name***"  
 </value>

### Stereochemistry

|  |  |
| --- | --- |
| **User Guidance** | The stereochemistry of the substance shall be indicated in the structure if it can be. Special cases of stereochemistry that can’t be indicated in the structure shall be described based on a controlled vocabulary. |
| **Example(s)** | AXIAL R SQUARE PLANAR 1, TETRAHEDRAL, OCTAHEDRAL 12 |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Value Allowed** | AXIAL R, AXIAL S, SQUARE PLANAR 1, SQUARE PLANAR 2, SQUARE PLANAR 3, SQUARE PLANAR 4, TETRAHEDRAL, OCTAHEDRAL 12, OCTAHEDRAL 22, OCTAHEDRAL 21 |
| **Business Rule(s)** | Mixtures of stereoisomers shall be represented explicitly as a mixture of substances with absolute stereochemistry. In case the absolute stereochemistry is unknown the substance definition should be marked as “Incomplete”. |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Moiety Details --></partMoiety>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Strereochemistry’*"   
 codeSystem="*Code System (OID)*"  
 displayName="stereochemistry"/>  
 <value code="***Stereo Chemistry*** *Code*" xsi:type="CV"  
 codeSystem="*Stereo Chemistry Code System (OID)*"  
 displayName="*Stereo Chemistry Name*"/>

### Optical Activity

|  |  |
| --- | --- |
| **User Guidance** | The optical activity shall be described based on a controlled vocabulary. |
| **Example(s)** | (+/-), (+), (-) |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Will only be captured for substances that have at least one moiety with stereochemistry defined as chiral. Can be a defining element when the absolute stereo configuration is not fully defined. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Parts --></identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Optical Activity*"   
 codeSystem="*Code System (OID)*"  
 displayName="optical activity"/>  
 <value code="***Optical Activity*** *Code*" xsi:type="CV"  
 codeSystem="*Optical Activity Code System (OID)*"  
 displayName="*Optical Activity Name*"/>

### Molecular Formula

|  |  |
| --- | --- |
| **User Guidance** | Derived from structural representation for defined stoichiometry molecule. Specified according to the Hill system, i.e., first C, then H, then alphabetical.  NOTE: The molecular formula shall be specified with all the characters in capital letters. |
| **Example(s)** | C2H6O |
| **Conformance** | CONDITIONAL |
| **Data Type** | ST |
| **Business Rule(s)** | * Calculated from structure * Validity check for identity and structural representation |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Parts --></identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Molecular Formula’*"   
 codeSystem="*Code System (OID)*"  
 displayName="stereochemistry"/>  
 <value code="***Molecular Formula***" xsi:type="CV"  
 codeSystem="*Molecular Formula Code System (OID)*"/>

Note: molecular formula is written as a sequence of pairs of atom symbol and stoichiometric number without any text formatting or special characters. This formal notation is essentially a code system.

### Isotope (repeat as necessary)

Applicable for single substances that contain a radionuclide or a non-natural isotopic ratio, e.g. C-13 enriched material). All radionuclide and non-natural isotopes will also be represented in the structure representation.

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety>  
 <bond>  
 <code code="*Substitution Type ‘Isotope Substitution’ or specialized*"  
 codeSystem="*Code System (OID)*"/>  
 <position value="*Substitution Type if specified atom position)*"/>  
 <distalMoiety>  
 <code code="***Nuclide Id***" codeSystem="*Nuclide Id System (OID)*"/>

#### Nuclide ID

|  |  |
| --- | --- |
| **User Guidance** | A substance ID for each non-natural or radioisotope will be created. The ID will be linked to a single atom zero valence element. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | II |
| **Business Rule(s)** | Mandatory when substance contains non-natural or radioisotope. |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety>  
 <bond>  
 <code code="*Substitution Type ‘Isotope Substitution’ or specialized*"  
 codeSystem="*Code System (OID)*"/>  
 <distalMoiety>  
 <code code="***Nuclide Id***" codeSystem="*Nuclide Id System (OID)*"/>

#### Nuclide Name

|  |  |
| --- | --- |
| **User Guidance** | The name for each isotope shall be provided, e.g. CARBON C-13 |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | ST |
| **Business Rule(s)** | Mandatory when substance contains non-natural or radioisotope. |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety>  
 <bond>  
 <code code="*Substitution Type ‘Isotope Substitution’ or specialized*"  
 codeSystem="*Code System (OID)*"/>  
 <position value="*Substitution Type if specified atom position)*"/>  
 <distalMoiety>  
 <name>***Nuclide Name***</name>

#### Substitution Type

Applicable for single substances that contain a radionuclide or a non-natural isotopic ratio, e.g. C-13 enriched material). All radionuclide and non-natural isotopes will also be represented in the structure representation.

|  |  |
| --- | --- |
| **User Guidance** | The type of isotopic substitution present in a single substance:  - specific if the site of attachment/substitution indicated in structure);  - non-specific if nuclide distributed throughout molecule or substance;  - unknown if site unknown.  Substitution refers to the relationship between the nuclide and the rest of the substance. |
| **Example(s)** | specific; non-specific; unknown |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Mandatory when substance contains non-natural or radioisotope. |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety>  
 <bond>  
 <code code="***Substitution Type*** *‘Isotope Substitution’ or specialized*"  
 codeSystem="*Code System (OID)*"/>

## Amount

The amount section serves to provide the quantitative or qualitative values that are assoiciated with a variety of elements. The same format will be used for all quantitative and qualitative values.

The information shall be provided by means of the following data elements:

In HL7 the quantification may be attached to the specific element to which it applies.

* Amount of a substitution (bond) or moiety is represented as a ratio of amount of content or substituent or moiety in the numerator over amount of the larger molecular unit in the denominator.
* Amount of reactants in a reaction (interactors in an interaction) is represented as quantities on the Participation element called <interactsIn>, <interactor>, <productOf>. These amounts are simple quantities, not ratios.
* Amounts specified as Characteristic values.
* Amount of moiety in a substance, e.g., frequently used when specifying mixtures.

<identifiedSubstance>  
 <moiety>  
 <quantity>  
 <numerator *Amount of Moiety* />  
 <denominator *Amount of Entire Substance* />  
 </quantity>  
 <partMoiety>

Amount of substituent in a substitution or amounts of bonding of different structural units.

<partMoiety>  
 <bond>  
 <quantity>  
 <numerator *Amount of Substituent* />  
 <denominator *Amount of This Structural Unit* />  
 </quantity>  
 <distalMoiety>

Amount of interactands of an interaction

<identifiedSubstance>  
 <interactsIn>  
 <quantity *Amount of This Substance* />  
 <interaction>  
 <interactor>  
 <quantity *Amount of Other Reactant* />

Amount of reactants of a derivation process

<identifiedSubstance>  
 <productOf>  
 <quantity *Amount of This Substance* />  
 <derivationProcess>  
 <interactor>  
 <quantity *Amount of Other Reactant* />

Characteristic of type physical quantity (PQ):

<subjectOf>  
 <characteristic>  
 <code code="*Characteristic Code*" codeSystem="*Code System (OID)*"/>  
 <value xsi:type="PQ" value="*quantity* *value*" unit="*quantity* *unit*">

Characteristic of type number (REAL):

<subjectOf>  
 <characteristic>  
 <code code="*Characteristic Code*" codeSystem="*Code System (OID)*"/>  
 <value xsi:type="REAL" value="*quantity* *value*"/>

Characteristic of type integer number (INT):

<subjectOf>  
 <characteristic>  
 <code code="*Characteristic Code*" codeSystem="*Code System (OID)*"/>  
 <value xsi:type="INT" value="*quantity* *value*"/>

Characteristic of type uncertain range of physical quantity (URG<PQ>):

<subjectOf>  
 <characteristic>  
 <code code="*Characteristic Code*" codeSystem="*Code System (OID)*"/>  
 <value xsi:type="URG\_PQ">  
 <low value="*quantity value low boundary*" unit="*quantity unit*"/>  
 <high value="*quantity value high boundary*" unit="*quantity unit*"/>  
 </value>

Characteristic of coded type (CV):

<subjectOf>  
 <characteristic>  
 <code code="*Characteristic Code*" codeSystem="*Code System (OID)*"/>  
 <value xsi:type="CV" code="*value* *code*"  
 codeSystem="*value code system OID*"  
 displayName="*value code display name*">

### Average

|  |  |
| --- | --- |
| **User Guidance** | Used to capture quantitative values for a variety of elements. If only limits are given the arithmetic mean would be the average. If only a single definite value for a given element is given it would be captured in this field. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | PQ |
| **Business Rule(s)** | This element captures single quantitative values or an average value for a given element. |

<quantity>  
 <numerator value="***Average***" unit="*Unit*"/>

### Low Limit

|  |  |
| --- | --- |
| **User Guidance** | Lower limit for a given quantitative value. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | PQ |
| **Business Rule(s)** | To be used to express the lower limit of a range of values. |

<quantity>  
 <numerator xsi:type="URG\_PQ">  
 <low value="***Low Limit***" unit="*Unit*"/>

The low limit can be alone to express one-sided ranges greater or equal than this limit, or together with the low limit to express a range bounded on both sides.

<quantity>  
 <numerator xsi:type="URG\_PQ">  
 <low value="*Low Limit*" unit="*Unit*"/>  
 <high value="*High Limit*" unit="*Unit*"/>

### High Limit

|  |  |
| --- | --- |
| **User Guidance** | Higher limit for a given quantitative value. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | PQ |
| **Business Rule(s)** | To be used to express the higher limit of a range of values. |

<quantity>  
 <numerator xsi:type="URG\_PQ">  
 <high value="***High Limit***" unit="*Unit*"/>

The high limit can be alone to express one-sided ranges less or equal than this limit, or together with the low limit to express a range bounded on both sides.

<quantity>  
 <numerator xsi:type="URG\_PQ">  
 <low value="*Low Limit*" unit="*Unit*"/>  
 <high value="*High Limit*" unit="*Unit*"/>

### Unit

|  |  |
| --- | --- |
| **User Guidance** | The unit of measure |
| **Example(s)** | 1, g, mg, mol, mmol, L, [IU], mg/L, mol/L, g/mol, [IU]/g |
| **Conformance** | REQUIRED – default is 1 (i.e., the dimensionless unity) |
| **Data Type** | CD |
| **Value Allowed** | UCUM / ISO 11240 units |
| **Business Rule(s)** | When used with limits, both limits must have the same dimension (and usually will have the same unit.) |

<quantity>  
 <numerator value="*Average*" unit="***Unit***"/>

<quantity>  
 <numerator xsi:type="URG\_PQ">  
 <low value="*Low Limit*" unit="***Unit***"/>  
 <high value="*High Limit*" unit="***Unit***"/>

### Non Numeric Value

|  |  |
| --- | --- |
| **User Guidance** | Qualitative, coded property value. |
| **Example(s)** | Positive, negative, Yes, No, complete, partial |
| **Conformance** | CONDITIONAL |
| **Data Type** | ST |
| **Business Rule(s)** | To be used to capture qualitative values. |

In HL7, these are all exceptional cases. Some of them are not quantities. When Property values are expressed that have non-quantitative (qualitative) values one this is done as Property value of data types other than that for true “amounts”, e.g. of coded type (CV):

<subjectOf>  
 <characteristic>  
 <code code="*Characteristic Code*" codeSystem="*Code System (OID)*"/>  
 <value xsi:type="**CV**" code="*value* *code*"  
 codeSystem="*value code system OID*"  
 displayName="*value code display name*"/>

or Boolean type (BL):

<subjectOf>  
 <characteristic>  
 <code code="*Characteristic Code*" codeSystem="*Code System (OID)*"/>  
 <value xsi:type="**BL**" value="*true/false*"/>

## Property

The section serves to provide information related to biological, physical or chemical characteristics associated with a substance. This information can be essential for the definition of each type of substance when structural elements are not sufficient to distinguish similar substances.

*Example: the pH of magnesium aluminometasilicate is necessary to distinguish low and high pH substances; viscosity is used to distinguish many polymers*).

The section is intended to only capture properties that may be essential to distinguish similar substances and not general properties that may be associated with a substance. The property information shall be provided by means of the following data elements:

<identifiedSubstance>  
 <identifiedSubstance>...</identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Property (Type/Name) Code*" codeSystem="*Code System (OID)*"  
 displayName="*Property Name*">  
 <value *According to Appropriate Data Type* />  
 <analyte>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="*Substance Id*" codeSystem="*Substance Id System (OID)*"/>  
 <name>*Substance Name*</name>  
 </identifiedSubstance>  
 </identifiedSubstance>  
 </analyte>  
 </characteristic>  
 </subjectOf>

In HL7 every property must be specified with a fully defined concept from an adequate terminology. Constructing meanings on the fly with undefined strings and names is not permitted.

### Property Type

|  |  |
| --- | --- |
| **User Guidance** | Type of the property for which the information is provided. |
| **Example(s)** | physical, chemical, enzymatic, immunological |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Each property must be provided with a fully terminologically defined concept for measurements. Such definition would include any necessary of “types” such as this.. |

<identifiedSubstance>  
 <identifiedSubstance>...</identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="***Property*** *(****Type****/Name) Code*" codeSystem="*Code System (OID)*"  
 displayName="*Property Name*">

### Property Name

|  |  |
| --- | --- |
| **User Guidance** | The name of the property shall be specified basedon a controlled vocabulary. |
| **Example(s)** | VISCOSITY, PH, CELL SURFACE ANTIGEN |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | A word or short term that describes a property. A fully defined coded concept must always be provided. |

<identifiedSubstance>  
 <identifiedSubstance>...</identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="***Property*** *(Type/****Name****) Code*" codeSystem="*Code System (OID)*"  
 displayName="***Property Name***">

### Substance ID

|  |  |
| --- | --- |
| **User Guidance** | Substance ID of a substance upon which a defining property depends; the identifier of the substance related to a defining property shall be described where applicable and based on controlled vocabulary. |
| **Example(s)** | Solubility in Water, Water ID shall be specified. |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | A substance ID should be created for every substance upon which a property is dependent. |

<identifiedSubstance>  
 <identifiedSubstance>...</identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <analyte>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Substance Id***" codeSystem="*Substance Id System (OID)*"/>

### Substance Name

|  |  |
| --- | --- |
| **User Guidance** | To be used to identify a substance related to a defining property, for example: cell surface antigens (the substance id for cd4 would be captured to defined CD4 positive cells). |
| **Example(s)** |  |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | The preferred term. |

<identifiedSubstance>  
 <identifiedSubstance>...</identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <analyte>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="*Substance Id*" codeSystem="*Substance Id System (OID)*"/>  
 <name>***Substance Name***</name>

### Amount

The amount should be captured as described in section 8.8.

<identifiedSubstance>  
 <identifiedSubstance>...</identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <value *According to Appropriate Data Type* />

## Chemical Substance

Chemical substances are a single substance type where a limited and complete molecular structure is the primary defining element. Chemical substances are defined on the basis of their covalent molecular structure, the presence of a salt (counter-ion) and/or solvates (water, alcohols). Purity, grade, physical form or particle size are not taken into account in the definition of a chemical substance or in the assignment of a Substance ID.

*Example: Purified Water, Water for Injection, Sterile Water for Injection USP, ice and steam all map to the substance Water.*

In order to assign a Substance ID for a chemical substance, a complete covalent structure with all stereochemistry (R or S and E or Z) defined is needed. The stoichiometry (mole ratio) of counterions or solvates present in the material should also be provided.

For all chemical substances a molecular representation (mol file, smiles, InChI, ChemDraw) of the substance should be provided with all stereochemistry assigned or sites of unknown stereochemistry identified.

The structural information typically provided to INN, USAN, BAN or JAN to generate assign a non-proprietary name is typically sufficient to define chemical substances. Active substances and excipients are typically defined in accordance with INN, JAN, or USAN definitions or USP/NF and European Pharmacopiea (EP) monographs. The labeling requirements provided by monographs should be taken into account when defining and distinguishing substances. Information beyond monograph requirements may occasionally be necessary to distinguish substances.

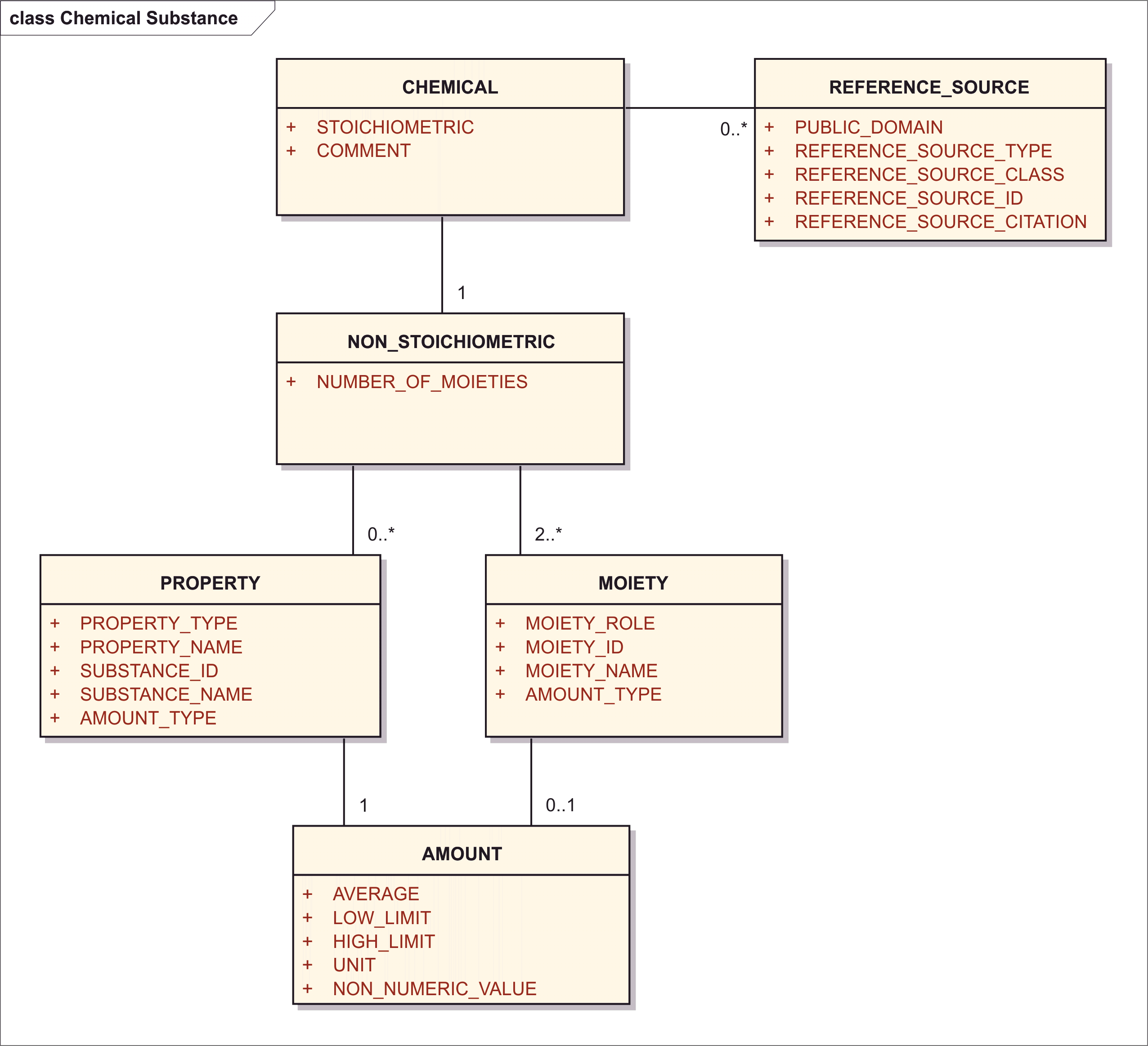


Figure 9 – Information model for the Chemical Substance

In order to assign Substance IDs for chemical substance, the following information should be provided:

### Substance Name

The information related to the name of the substance shall be provided as per specification described in Section 8.2.

### Substance Code

The information related to the code of the substance shall be provided as per specification described in Section 8.3.

### Reference Information

The information related to the reference information of the substance shall be provided as per specification described in Section 8.5.

### Version

The information related to the version shall be provided as per specification described in Section

### Structure

The information related to the structure shall be provided as per specification described in Section 8.7.

Chemical substances can either be stoichiometric.

### Stoichiometric

|  |  |
| --- | --- |
| **User Guidance** | If a substance can be represented as single covalent structural moiety or a salt or solvate with a defined and constant ratio ions or molecules and solvents. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | BL |
| **Business Rule(s)** | The Stoichiometric element is a boolean field (Y/N value applies). |

<identifiedSubstance>  
 <identifiedSubstance>...</identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Substance is Stoichiometric’*"   
 codeSystem="*Code System (OID)*">  
 <value xsi:type="BL" value="***Stoichiometric*** *(true/false)*"/>  
 </characteristic>  
 </subjectOf>

Stoichiometric Chemicals

Stoichiometric chemical substances are substances that contain complete chemical structures and definite stoichiometric ratios among moities. They can be defined by a single representation of the complete molecular structure and a stereochemistry descriptor. The representation can be of one of the four types listed above.

The following examples of defining elements for various types of chemical substances help illustrate the information needed to define substances. There are a variety of drawing tools that can generate any of the formats described.

Ethanol Example

Ethanol is an example of an achiral substance and can be completely defined by either of the three following representations:

Inchi Representation

STRUCTURAL\_REPRESENTATION\_TYPE: “INCHI”

STRUCTURAL\_REPRESENTATION: “1/C2H6O/c1-2-3/h3H,2H2,1H3”

STEREOCHEMISTRY: “ACHIRAL”

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Other Details of Structural Unit --></partMoiety>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Complete Structural Representation’*"   
 codeSystem="*Code System (OID)*">  
 <value xsi:type="ED" mediaType="application/x-**inchi**">InChI=1S/C2H6O/c1-2-3/h3H,2H2,1H3</value>  
 </characteristic>  
 </subjectOf>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘****Stereochemistry****’*"   
 codeSystem="*Code System (OID)*">  
 <value xsi:type="CV" code="*Code for ‘****achiral****’*"   
 codeSystem="*Code System (OID)*"  
 displayName="achiral"/>  
 </characteristic>  
 </subjectOf>

Smiles Representation

STRUCTURAL\_REPRESENTATION\_TYPE: “SMILES”

STRUCTURAL\_REPRESENTATION: “CCO”

STEREOCHEMISTRY: “ACHIRAL” (same as above)

<subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Complete Structural Representation’*"   
 codeSystem="*Code System (OID)*">  
 <value xsi:type="ED" mediaType="application/x-**smiles**">**CCO**</value>  
 </characteristic>  
 </subjectOf>

Mol representation

STRUCTURAL\_REPRESENTATION\_TYPE: “MOL”

STRUCTURAL\_REPRESENTATION:

3 2 0 0 0 0 0 0 0 0999 V2000  
 3.9063 -7.3125 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 4.6223 -6.8991 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 5.3383 -7.3125 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0  
 1 2 1 0 0 0 0  
 2 3 1 0 0 0 0  
M END

STEREOCHEMISTRY: “ACHIRAL” (same as above)

<subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Complete Structural Representation’*"   
 codeSystem="*Code System (OID)*">  
 <value xsi:type="ED" mediaType="application/x-mdl-**molfile**"><![CDATA[  
 -FDASRS-07251317382D  
 **3 2 0 0 0 0 0 0 0 0999 V2000  
 3.9063 -7.3125 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 4.6223 -6.8991 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 5.3383 -7.3125 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0  
 1 2 1 0 0 0 0  
 2 3 1 0 0 0 0  
M END** ]]></value>  
 </characteristic>  
 </subjectOf>

Levofloxacin Sodium

This is an example of a chiral substance and salt. An INCHI, SMILES or MOL file is sufficient to define this substance.



Inchi Representation

STRUCTURAL\_REPRESENTATION\_TYPE: “INCHI”

STRUCTURAL\_REPRESENTATION: “

InChI=1S/C18H20FN3O4.Na/c1-10-9-26-17-14-11(16(23)12(18(24)25)8-22(10)14)7-13(19)15(17)21-5-3-20(2)4-6-21;/h7-8,10H,3-6,9H2,1-2H3,(H,24,25);/q;+1/p-1/t10-;/m0./s1

”

STEREOCHEMISTRY: “CHIRAL”

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Other Details of Structural Unit --></partMoiety>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Complete Structural Representation’*"   
 codeSystem="*Code System (OID)*">  
 <value xsi:type="ED" mediaType="application/x-**inchi**">InChI=1S/C18H20FN3O4.Na/c1-10-9-26-17-14-11(16(23)12(18(24)25)8-22(10)14)7-13(19)15(17)21-5-3-20(2)4-6-21;/h7-8,10H,3-6,9H2,1-2H3,(H,24,25);/q;+1/p-1/t10-;/m0./s1</value>  
 </characteristic>  
 </subjectOf>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Stereochemistry’*"   
 codeSystem="*Code System (OID)*">  
 <value xsi:type="CV" code="*Code for ‘****chiral****’*"   
 codeSystem="*Code System (OID)*"  
 displayName="chiral"/>  
 </characteristic>  
 </subjectOf>

Smiles Representation

STRUCTURAL\_REPRESENTATION\_TYPE: “SMILES”

STRUCTURAL\_REPRESENTATION: “C[C@H]1COc2c3n1cc(c(=O)c3cc(c2N4CCN(CC4)C)F)C(=O)[O-].[Na+]”

STEREOCHEMISTRY: “CHIRAL”

<subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Complete Structural Representation’*"   
 codeSystem="*Code System (OID)*">  
 <value xsi:type="ED" mediaType="application/x-**smiles**"> **C[C@H]1COc2c3n1cc(c(=O)c3cc(c2N4CCN(CC4)C)F)C(=O)[O-].[Na+]**</value>  
 </characteristic>  
 </subjectOf>

Mol Representation

STRUCTURAL\_REPRESENTATION\_TYPE: “MOL”

STRUCTURAL\_REPRESENTATION:

27 29 0 0 1 0 0 0 0 0999 V2000  
 14.5167 -9.0507 0.0000 N 0 0 0 0 0 0 0 0 0 0 0 0  
 15.1250 -7.9548 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 13.8500 -8.7215 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 13.2625 -9.0840 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 13.8500 -8.0048 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 15.1250 -8.7048 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 14.4667 -7.6340 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 12.6334 -8.7673 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 13.2625 -9.8506 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0  
 12.6334 -8.0048 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 13.2292 -7.6548 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 12.0167 -9.1298 0.0000 N 0 0 0 0 0 0 0 0 0 0 0 0  
 15.8250 -7.6007 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 14.5167 -9.7840 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 10.7417 -9.8965 0.0000 N 0 0 0 0 0 0 0 0 0 0 0 0  
 14.4667 -6.8965 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0  
 13.9125 -10.1632 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 12.0167 -9.8798 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 11.3584 -8.7757 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 15.8250 -6.9548 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0  
 11.9542 -7.7132 0.0000 F 0 0 0 0 0 0 0 0 0 0 0 0  
 10.7417 -9.1340 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 11.4000 -10.2465 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 16.6458 -7.9631 0.0000 O 0 5 0 0 0 0 0 0 0 0 0 0  
 10.0792 -10.1715 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 15.3500 -10.3215 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 17.3000 -7.9630 0.0000 Na 0 3 0 0 0 0 0 0 0 0 0 0  
 21 10 1 0 0 0 0  
 22 19 1 0 0 0 0  
 23 18 1 0 0 0 0  
 24 13 1 0 0 0 0  
 25 15 1 0 0 0 0  
 14 26 1 1 0 0 0  
 4 9 1 0 0 0 0  
 5 7 1 0 0 0 0  
 8 10 2 0 0 0 0  
 15 23 1 0 0 0 0  
 2 6 2 0 0 0 0  
 3 1 1 0 0 0 0  
 4 3 2 0 0 0 0  
 5 3 1 0 0 0 0  
 6 1 1 0 0 0 0  
 7 2 1 0 0 0 0  
 8 4 1 0 0 0 0  
 9 17 1 0 0 0 0  
 10 11 1 0 0 0 0  
 11 5 2 0 0 0 0  
 12 8 1 0 0 0 0  
 13 2 1 0 0 0 0  
 14 1 1 0 0 0 0  
 15 22 1 0 0 0 0  
 16 7 2 0 0 0 0  
 17 14 1 0 0 0 0  
 18 12 1 0 0 0 0  
 19 12 1 0 0 0 0  
 20 13 2 0 0 0 0  
M CHG 2 24 -1 27 1  
M END

STEREOCHEMISTRY: “CHIRAL”

<subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Chemical structure*"   
 codeSystem="*Code System (OID)*">  
 <value xsi:type="ED" mediaType="application/x-mdl-molfile"><![CDATA[  
 -FDASRS-07251317382D  
27 29 0 0 1 0 0 0 0 0999 V2000  
 14.5167 -9.0507 0.0000 N 0 0 0 0 0 0 0 0 0 0 0 0  
 15.1250 -7.9548 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 13.8500 -8.7215 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 13.2625 -9.0840 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 13.8500 -8.0048 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 15.1250 -8.7048 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 14.4667 -7.6340 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 12.6334 -8.7673 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 13.2625 -9.8506 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0  
 12.6334 -8.0048 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 13.2292 -7.6548 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 12.0167 -9.1298 0.0000 N 0 0 0 0 0 0 0 0 0 0 0 0  
 15.8250 -7.6007 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 14.5167 -9.7840 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 10.7417 -9.8965 0.0000 N 0 0 0 0 0 0 0 0 0 0 0 0  
 14.4667 -6.8965 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0  
 13.9125 -10.1632 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 12.0167 -9.8798 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 11.3584 -8.7757 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 15.8250 -6.9548 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0  
 11.9542 -7.7132 0.0000 F 0 0 0 0 0 0 0 0 0 0 0 0  
 10.7417 -9.1340 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 11.4000 -10.2465 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 16.6458 -7.9631 0.0000 O 0 5 0 0 0 0 0 0 0 0 0 0  
 10.0792 -10.1715 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 15.3500 -10.3215 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0  
 17.3000 -7.9630 0.0000 Na 0 3 0 0 0 0 0 0 0 0 0 0  
 21 10 1 0 0 0 0  
 22 19 1 0 0 0 0  
 23 18 1 0 0 0 0  
 24 13 1 0 0 0 0  
 25 15 1 0 0 0 0  
 14 26 1 1 0 0 0  
 4 9 1 0 0 0 0  
 5 7 1 0 0 0 0  
 8 10 2 0 0 0 0  
 15 23 1 0 0 0 0  
 2 6 2 0 0 0 0  
 3 1 1 0 0 0 0  
 4 3 2 0 0 0 0  
 5 3 1 0 0 0 0  
 6 1 1 0 0 0 0  
 7 2 1 0 0 0 0  
 8 4 1 0 0 0 0  
 9 17 1 0 0 0 0  
 10 11 1 0 0 0 0  
 11 5 2 0 0 0 0  
 12 8 1 0 0 0 0  
 13 2 1 0 0 0 0  
 14 1 1 0 0 0 0  
 15 22 1 0 0 0 0  
 16 7 2 0 0 0 0  
 17 14 1 0 0 0 0  
 18 12 1 0 0 0 0  
 19 12 1 0 0 0 0  
 20 13 2 0 0 0 0  
M CHG 2 24 -1 27 1  
M END  
 ]]></value>  
 </characteristic>  
 </subjectOf>

### Comment

|  |  |
| --- | --- |
| **User Guidance** | Any comment can be provided in this field, if necessary. |
| **Conformance** | OPTIONAL – DISCOURAGED |
| **Data Type** | ST |

Note: Comment can always be packaged into a characteristic:

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Other Details of Structural Unit --></partMoiety>  
 <subjectOf>  
 <characteristic>  
 <text>***Comment*** *Text*</text>  
 </characteristic>  
 </subjectOf>

### Reference Source (repeat as necessary)

See Section 8.2.5.

### Non- Stoichiometric

The section applies to chemical substances that do not have defined stoichiometry. This section is used only when the Stoichiometric (Section 8.10.6) is equal to NO.

Every non-stoichiometric substance will have more than one moiety. Each moiety will be enclosed in brackets indicating stoichiometry where known and transmitted in a single structural representation. In addition each moiety will also have a substance ID in its own right. Every non-stoichiometric will have a complete structural representation in its own right associated with the substance and each of the moieties will also be described as a substance. If ranges of amount for the non-stoichiometric moieties are known they will be preferably captured as mole ratios or weight percent if mole ratios cannot be determined.

#### Number of Moieties

|  |  |
| --- | --- |
| **User Guidance** | The number of moieties specified shall be provided. Non-stoichiometric chemical substances must have at least two moieties. Each moiety shall be represented by a chemical structure. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | - The numeric value shall be always ≥ 2  - The value is automatically calculated – and therefore not explicitly represented in the HL7 SPL data. |

#### Moiety (repeat as necessary)

The moiety section serves for the description of both the moiety and the other fragment constituting the substance.

*EXAMPLE: Abacavir Sulphate*

*Abacavir and sulfate shall be described separately based on this section characteristic.*

For non-stoichiometric substances each individual ion or molecular specie is considered to be a moiety. For example the substance Aluminum Sesquichlorohydrex Propylene Glycol is defined by relationships between five moieties, the aluminum cation, chloride and hydroxide ions, propylene glycol and water.

Each moiety within the chemical substance is to be identified and the composition range of the moieties when known is to be provided by means of the following data elements:

##### Moiety Role

|  |  |
| --- | --- |
| **User Guidance** | The role of the moiety shall be described based on a controlled vocabulary. |
| **Example(s)** | COMPONENT; COUNTER ION; HYDRATE |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Such classification of role are subjective by the observer and do not change what the substance is, therefore it is not required. |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="***Moiety Role*** *Code*" codeSystem="*Code System (OID)*"  
 displayName="*Moiety Role Name*"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <code code="*Moiety Id*"   
 codeSystem="*Code System (OID)*"/>  
 <name>*Moiety Name*</name>  
 </partMoiety>  
 </moiety>  
 <moiety><!-- Next moiety, same content as above --></moiety>

##### Moiety ID

|  |  |
| --- | --- |
| **User Guidance** | The unique identifier assigned to the substance representing the moiety based on controlled vocabulary. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Should be specified if the moiety has been assigned a Substance Id. |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Moiety Role Code*" codeSystem="*Code System (OID)*"  
 displayName="*Moiety Role Name*"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <code code="***Moiety Id***"   
 codeSystem="*Code System (OID)*"/>  
 <name>*Moiety Name*</name>  
 </partMoiety>  
 </moiety>  
 <moiety><!-- Next moiety, same content as above --></moiety>

##### Moiety Name

|  |  |
| --- | --- |
| **User Guidance** | The name of the moiety shall be provided.  NOTE: If the substance name is not present in the Controlled Vocabulary, the moiety name shall be described as a free text. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Must not be used without either a Moiety Id or at least a structural representation of that moiety. |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Moiety Role Code*" codeSystem="*Code System (OID)*"  
 displayName="*Moiety Role Name*"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <code code="*Moiety Id*"   
 codeSystem="*Code System (OID)*"/>  
 <name>***Moiety Name***</name>  
 </partMoiety>  
 </moiety>  
 <moiety><!-- Next moiety, same content as above --></moiety>

##### Amount

The amount should be captured as described in Section 8.8.

#### Property (repeat as necessary)

See Section 8.10

##### Amount

The amount should be captured as described in Section 8.8.

## Proteins/peptides

A protein is defined as a single unit of a linear amino acid sequence, or a combination of sub-units that are either covalently linked or have a defined invariant stoichiometric relationship. This includes all synthetic, recombinant and purified proteins of defined sequence whether the use is therapeutic or prophylactic. This set of elements needs to be used to describe albumins, coagulation factors, cytokines, growth factors, peptide/protein hormones, enzymes, toxins, toxoids, recombinant vaccines, and immunomodulators.

Proteins and peptides are defined by their molecular structure based on the amino acid sequence, disulfide linkages, sites and general type of glycosylation (yeast, plant, mammal, or human host). The method of production is generally not a defining element for proteins and peptides. For a given non-glycosylated peptide, whether naturally isolated, produced by recombinant technology, or chemically synthesized, there is no resultant difference in the amino acid sequence and disulfide linkages that constitute its structure. For example, recombinant and chemically synthesized salmon calcitonin are chemically identical and therefore considered to be the same substance with the same Substance ID. Conversely, proteins that have significantly different types of glycosylation are assigned separate Substance IDs. Human glycosylation differs from other mammalian glycosylation in that terminal sialic acid residues are only acetylated and not glycolated. Therefore proteins such as Coagulation Factor VIII isolated from human blood versus that produced by recombinant technology in Chinese hamster ovary (CHO) cells are considered to be different substances even if they have the same amino acid sequence and disulfide linkages. Yeast, plant, and insect glycosylation are also significantly different from mammalian and human glycosylation and glycoproteins produced in such systems would each have a separate Substance ID even if the amino acid sequences and disulfide linkages are the same. For proteins and glycoproteins microheterogeneity is not taken into account in the assignment of a Substance ID. Therefore a given glycoprotein produced in different mammalian cell lines would have the same Substance ID.

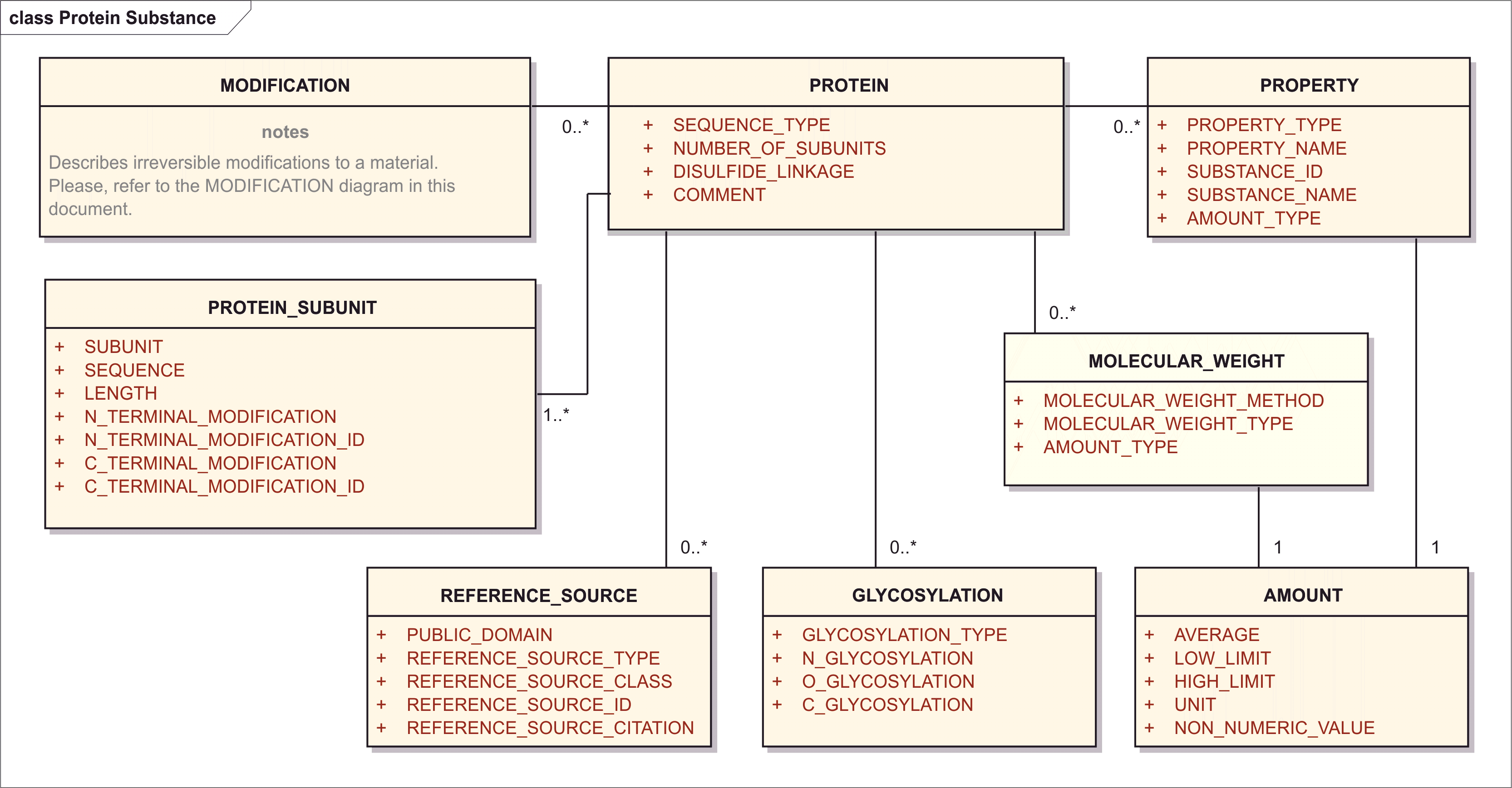


Figure 10 – Information model for the Protein Substance

In order to assign Substance IDs for proteins and peptides, the following information should be provided:

### Substance Name

The information related to the name of the substance shall be provided as per specification described in Section 8.2.

### Substance Code

The information related to the code of the substance shall be provided as per specification described in Section 8.3.

### Reference Information

The information related to the reference information of the substance shall be provided as per specification described in Section 8.5.

### Version

The information related to the version of the substance shall be provided as per specification described in Section 8.4.

### Reference Source (repeat as necessary)

See section 8.2.5.

### Molecular Weight (repeat as necessary)

NOTE: The section is **mandatory** if the substance is **protein**, **polymer** or **nucleic acid**.

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Molecular Weight’ with ‘type’ and ‘method’*"   
 codeSystem="*Code System (OID)*"  
 displayName="molecular mass"/>  
 <value xsi:type="PQ" value="*Molecular Weight*" unit="g/mol"/>

#### Molecular Weight Method

|  |  |
| --- | --- |
| **User Guidance** | The method by which the molecular weight was determined. |
| **Example(s)** | SDS-PAGE, calculated, light scattering viscosity, gel permeation, (..) |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘****Molecular Weight****’ with ‘type’ and ‘****method****’*"   
 codeSystem="*Code System (OID)*"  
 displayName="molecular mass"/>

#### Molecular Weight Type

|  |  |
| --- | --- |
| **User Guidance** | Type of molecular weight |
| **Example(s)** | Exact, average (aka. number average), weight average, (..) |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘****Molecular Weight****’ with ‘****type****’ and ‘method’*"   
 codeSystem="*Code System (OID)*"  
 displayName="molecular mass"/>

#### Amount

The information related to the amount shall be provided as per specification described in Section 8.8.

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Molecular Weight’ with ‘type’ and ‘method’*"   
 codeSystem="*Code System (OID)*"  
 displayName="molecular mass"/>  
 <value xsi:type="PQ" value="***Molecular Weight***" unit="g/mol"/>

### Structure

The information related to the structure shall be provided as per specification described in Section 8.7.

As above described, each structural unit is placed in a moiety element in HL7. For proteins and peptides, the first and foremost structural units are the peptide chain or chains, also called a “sub-unit”. Each chain is represented as one sub-unit moiety:

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Protein Sub-Unit’*"   
 codeSystem="*Code System (OID)*"  
 displayName="protein sub-unit"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <id extension="*Subunit Label (Local Id)*"   
 root="*UUID of Document Id or Set Id*"/>  
 <code code=" *Code for ‘Amino-Acid Sequence’ with Sequence Type* " displayName="amino acid sequence, complete" codeSystem=" *Code System (OID)*"/>  
 <name>”Sub-unit name”</name>  
 </partMoiety>  
 </moiety>  
 <moiety><!-- Next chain, same content as above --></moiety>

### Sequence Type

|  |  |
| --- | --- |
| **User Guidance** | The protein descriptive elements will only be used when a complete or partial amino acid sequence is available or derivable from a nucleic acid sequence. |
| **Example(s)** | COMPLETE; PARTIAL |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Amino-Acid Sequence’ with* ***Sequence Type***"   
 codeSystem="*Code System (OID)*"  
 displayName="*Name for Amino-Acid Sequence Complete/Partial*"/>

### Number of subunits

|  |  |
| --- | --- |
| **User Guidance** | Number of linear sequences of amino acids linked through peptide bonds. The number of subunit constituting the protein shall be described.  NOTE: If not specified in the reference source, the assumption is that there is 1 subunit. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | This number can be counted and is never explicitly represented. |

### Comment

|  |  |
| --- | --- |
| **User Guidance** | Any comment can be provided in this field, if necessary. |
| **Conformance** | OPTIONAL – DISCOURAGED |
| **Data Type** | ST |

Note: Comment can always be packaged into a characteristic:

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Other Details of Structural Unit --></partMoiety>  
 <subjectOf>  
 <characteristic>  
 <text>***Comment*** *Text*</text>  
 </characteristic>  
 </subjectOf>

### Protein Subunit (repeat as necessary)

This section refers to the description of each constituting the protein. A subunit is a linear sequence of amino acids linked through peptide bonds. The Subunit information shall be provided when the finished protein is a complex of multiple sequences; subunits are not used to delineate domains within a single sequence. Subunits are listed in order of decreasing length; sequences of the same length will be ordered by decreasing molecular weight; subunits that have identical sequences will be repeated multiple times.

#### Subunit (Index)

|  |  |
| --- | --- |
| **User Guidance** | Index of linear sequences of amino acids linked through peptide bonds in order of decreasing length. Sequences of the same length will be ordered by molecular weight. Subunits that have identical sequences will be repeated and have sequential subscripts. |
| **Example(s)** | 1, 2, 3 (..) |
| **Conformance** | REQUIRED |
| **Data Type** | INT |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Protein Sub-Unit’*"   
 codeSystem="*Code System (OID)*"  
 displayName="protein sub-unit"/>  
 <partMoiety>  
 <id extension="***Subunit Index***"   
 root="*UUID of Document Id or Set Id*"/>

#### Sequence

|  |  |
| --- | --- |
| **User Guidance** | The sequence information shall be provided enumerating the amino-acids from N- to C-terminal end using standard single-letter amino acid codes. Uppercase shall be used for L-amino acids and lowercase for D-amino acids. Transcribed proteins will always be described using the translated sequence; for synthetic peptide containing amino acids that are not represented with a single letter code an X will be used within the sequence. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | ST |

<identifiedSubstance>  
 <id extension="*Subunit Label (Local Id*)" root=" *UUID of Document Id or Set Id* "/>  
 <identifiedSubstance>  
 <code code=" *Code for ‘Protein Sub-Unit’*" codeSystem=" *Code System (OID)*"/>  
 <name>”Sub-unit name"</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code=" *Code for ‘Amino-Acid Sequence’ with Sequence Type* " displayName="amino acid sequence, complete" codeSystem=" *Code System (OID)*"/>  
 <value xsi:type="ED" mediaType="application/x-aa-seq"><!-- Amino-Acid Sequence in Single Letter Notation --></value>  
 </characteristic>  
 </subjectOf>

</identifiedSubstance>

#### Length

|  |  |
| --- | --- |
| **User Guidance** | Length of linear sequences of amino acids contained in the subunit |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | The sequence length is implicitly in the length property of the HL7 ST data type. |

### Post-Translational Modifications

#### Structurally specific modifications

In HL7, all structurally specific post-translational modifications are expressed as amino acid substitutions. Each substituent is described as a moiety with one or more bond elements:

<moiety>  
 <code code="*code for ‘Modification’*" codeSystem=" *Code System (OID)*"  
 displayName="peptide chain snap-in modification"/>  
 <partMoiety>  
 <id extension="*Local moiety Id*" root=" *UUID of Document Id or Set Id* "/>  
 <code code="Moiety code" codeSystem=" *Code System (OID)*"/>  
 <name>”Moiety name”</name>  
 <bond>  
 <code code="*Code for “peptide chain snap-in connection”*"  
 codeSystem="*Code System (OID)*"   
 displayName="peptide chain snap-in connection"/>  
 <positionNumber value=" *Position in the sequence*"/>  
 <distalMoiety>  
 <id extension="*Subunit Label (Local Id)*"   
 root=" *UUID of Document Id or Set Id* "/>  
 </distalMoiety>  
 </bond>

Each bond element describes one position in the sequence that is replaced by the substituent. Moiety code references the description of the substituent:

<subject>  
 <identifiedSubstance>  
 <id extension="*Local substituent Id* " root=" *UUID of Document Id or Set Id* "/>  
 <identifiedSubstance>  
 <code code="*Substituent code*" codeSystem="*Code System (OID)*"/>  
 <name>*Substituent name*</name>  
 <moiety>  
 <code code="*Code for “peptide chain snap-in bond-site”*"   
 codeSystem=" *Code System (OID)*"   
 displayName="peptide chain snap-in bond-site"/>  
 <positionNumber value="*Canonical number of N atom* "/>  
 <positionNumber value="*Canonical number of C atom* "/>  
 <partMoiety/>  
 </moiety>  
 <subjectOf>  
 <!—Elements describing the chemical structure of the substituent -->  
 </subjectOf>  
 </identifiedSubstance>  
 </identifiedSubstance>  
</subject>

Each moiety element describes a part of the molecule that substitutes one amino acid in a sequence. Canonical InChI numbering shall be used to indicate positions of N and C atoms that form peptide bonds with amino acids in the sequence.



**Example 1: Substitution of N terminal amino acid by pyroglutamic acid**

<moiety>  
 <code code="" codeSystem="2.16.840.1.113883.3.26.1.1" displayName="modification"/>  
 <partMoiety>  
 <id extension="M2" root="bf170df7-8153-45c1-9a21-6e1aa1b4c97e"/>  
 <code code="SZB83O1W42" codeSystem="2.16.840.1.113883.4.9"/>  
 <name>pyroglutamic acid</name>  
 <bond>  
 <code code=""   
 codeSystem="2.16.840.1.113883.3.26.1.1"   
 displayName="amino acid snap-n-go"/>  
 <positionNumber value="1"/>  
 <distalMoiety>  
 <id extension="SU4"   
 root="bf170df7-8153-45c1-9a21-6e1aa1b4c97e"/>  
 </distalMoiety>  
 </bond>  
 </partMoiety>  
 </moiety>

<subject>  
 <identifiedSubstance>  
 <id extension="SZB83O1W42" root="2.16.840.1.113883.4.9"/>  
 <identifiedSubstance>  
 <code code="SZB83O1W42" codeSystem="2.16.840.1.113883.4.9"/>  
 <name>pyroglutamic acid</name>  
 <moiety>  
 <code code="" codeSystem="2.16.840.1.113883.3.26.1.1"   
 displayName="N-C snap-in bond-site"/>  
 <positionNumber value="6"/>  
 <positionNumber value="5"/>  
 <partMoiety/>  
 </moiety>  
 </identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code displayName="Chemical Structure"   
 codeSystem="2.16.840.1.113883.3.26.1.1" code="C103240"/>  
 <value xsi:type="ED"   
 mediaType="application/x-inchi">  
 InChI=1S/C5H7NO3/c7-4-2-1-3(6-4)5(8)9/h3H,1-2H2,(H,6,7)(H,8,9)/t3-/m0/s1</value>  
 </characteristic>  
 </subjectOf>  
 </identifiedSubstance>  
</subject>

**Example 2: Disulfide bridge**

<moiety>  
 <code code="" codeSystem="2.16.840.1.113883.3.26.1.1" displayName="modification"/>  
 <partMoiety>  
 <id extension="BR1" root="bf170df7-8153-45c1-9a21-6e1aa1b4c97e"/>  
 <code code="CSSC" codeSystem="2.16.840.1.113883.4.9"/>  
 <name>cysteine disulfide</name>  
 <bond>  
 <code code="" codeSystem="2.16.840.1.113883.3.26.1.1"   
 displayName="amino acid snap-n-go"/>  
 <positionNumber value="220"/>  
 <distalMoiety>  
 <id extension="SU1" root="bf170df7-8153-45c1-9a21-6e1aa1b4c97e"/>  
 </distalMoiety>  
 </bond>  
 <bond>  
 <code code="" codeSystem="2.16.840.1.113883.3.26.1.1"   
 displayName="amino acid snap-n-go"/>  
 <positionNumber value="220"/>  
 <distalMoiety>  
 <id extension="SU2" root="bf170df7-8153-45c1-9a21-6e1aa1b4c97e"/>  
 </distalMoiety>  
 </bond>  
 </partMoiety>  
</moiety>

<subject>  
 <identifiedSubstance>  
 <id extension="CSSC" root="b2315501-dc0c-56d0-e044-001185133a64"/>  
 <identifiedSubstance>  
 <code code="CSSC"   
 codeSystem="b2315501-dc0c-56d0-e044-001185133a64"/>  
 <name>cysteine disulfide</name>  
 <moiety>  
 <code code=""   
 codeSystem="2.16.840.1.113883.3.26.1.1"  
 displayName="N-C snap-in bond-site"/>  
 <positionNumber value="7"/>  
 <positionNumber value="5"/>  
 <partMoiety/>  
 </moiety>  
 <moiety>  
 <code code=""   
 codeSystem="2.16.840.1.113883.3.26.1.1"   
 displayName="N-C snap-in bond-site"/>  
 <positionNumber value="8"/>  
 <positionNumber value="6"/>  
 <partMoiety/>  
 </moiety>  
 </identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="C103240"   
 codeSystem="2.16.840.1.113883.3.26.1.1"  
 displayName="Chemical Structure"/>  
 <value xsi:type="ED"   
mediaType="application/x-inchi">InChI=1S/C6H12N2O4S2/c7-3(5(9)10)1-13-14-2-4(8)6(11)12/h3-4H,1-2,7-8H2,(H,9,10)(H,11,12)/t3-,4-/m0/s1</value>  
 </characteristic>  
 </subjectOf>  
 </identifiedSubstance>  
</subject>

**Example 3: Chemical linker**



<moiety>  
 <code code="" codeSystem="2.16.840.1.113883.3.26.1.1" displayName="modification"/>  
 <partMoiety>  
 <id extension="BR2" root="af170df7-8153-45c1-9a21-6e1aa1b4c97e"/>  
 <code code="K-fumaric acid-K" codeSystem="2.16.840.1.113883.4.9"/>  
 <name>lysine diamide with fumaric acid</name>  
 <bond>  
 <code code="" codeSystem="2.16.840.1.113883.3.26.1.1"   
 displayName="amino acid snap-n-go"/>  
 <positionNumber value="99"/>  
 <distalMoiety>  
 <id extension="SU1" root="af170df7-8153-45c1-9a21-6e1aa1b4c97e"/>  
 </distalMoiety>  
 </bond>  
 <bond>  
 <code code="" codeSystem="2.16.840.1.113883.3.26.1.1"   
 displayName="amino acid snap-n-go"/>  
 <positionNumber value="99"/>  
 <distalMoiety>  
 <id extension="SU2" root="af170df7-8153-45c1-9a21-6e1aa1b4c97e"/>  
 </distalMoiety>  
 </bond>  
 </partMoiety>  
</moiety>

<subject>  
 <identifiedSubstance>  
 <id extension=" K-fumaric acid-K " root="a2315501-dc0c-56d0-e044-001185133a64"/>  
 <identifiedSubstance>  
 <code code="K-fumaric acid-K "   
 codeSystem="a2315501-dc0c-56d0-e044-001185133a64"/>  
 <name> lysine diamide with fumaric acid </name>  
 <moiety>  
 <code code="" codeSystem="2.16.840.1.113883.3.26.1.1"  
 displayName="N-C snap-in bond-site"/>  
 <positionNumber value="17"/>  
 <positionNumber value="15"/>  
 <partMoiety/>  
 </moiety>  
 <moiety>  
 <code code="" codeSystem="2.16.840.1.113883.3.26.1.1"   
 displayName="N-C snap-in bond-site"/>  
 <positionNumber value="18”/>  
 <positionNumber value="16"/>  
 <partMoiety/>  
 </moiety>  
 </identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="C103240" codeSystem="2.16.840.1.113883.3.26.1.1"  
 displayName="Chemical Structure"/>  
 <value xsi:type="ED" mediaType="application/x-inchi"> InChI=1S/C16H28N4O6/c17-11(15(23)24)5-1-3-9-19-13(21)7-8-14(22)20-10-4-2-6-12(18)16(25)26/h7-8,11-12H,1-6,9-10,17-18H2,(H,19,21)(H,20,22)(H,23,24)(H,25,26)/b8-7+/t11-,12-/m0/s1</value>  
 </characteristic>  
 </subjectOf>  
 </identifiedSubstance>  
</subject>

#### Structurally non-specific modifications

Structurally non-specific modifications shall be represented by a code of a class of substances, e.g., glycanes.

##### Glycosylation (repeat as necessary)

The exact structure of the glycane moiety is usually not fully described but only represented as a general class, using concepts such as “human glycane”, “yeast glycane” (Glycosylation Type). The N-Linked, O-Linked, and C-Linked glycosylation is specified by the single positionNumber element.

|  |  |
| --- | --- |
| **User Guidance** | The type of the glycane shall be specified basedon a controlled vocabulary. |
| **Example(s)** | HUMAN GLYCANE, MAMMALIAN GLYCANE, AVIAN GLYCANE, FUNGAL GLYCANE, BACTERIAL GLYCANE, VIRAL GLYCANE, PLANT GLYCANE, INSECT GLYCANE |
| **Conformance** | CONDITIONAL – when glycosylation is known |
| **Data Type** | CD |
| **Business Rule(s)** |  |

#### Site non-specific modifications

In case the position of the modification is not known, “Amount” descriptor shall be used instead of bond elements.







































### Modification (repeat as necessary)

The Modification is to be used to describe irreversible modifications to a protein or peptide (e.g., PEGylation, phosphorylation, etc). The modifications may be physical, chemical, enzymatic, etc. Modifications may be described by their structural result (substitution of moieties to residues, etc.) or by the process, reagents, processing time if a specific structural modification cannot be determined (aggregated albumin).

A minimal description of the modification process shall be generated when a definitive structural modification can not be determined.

This section applies to:

* Nucleic Acid
* Protein
* Polymer
* Structurally Diverse Substances
* Mixture

This section shall be repeatable to describe each modification occurred on each residue modified.

The information model for the class modification group is shown in the following figure:

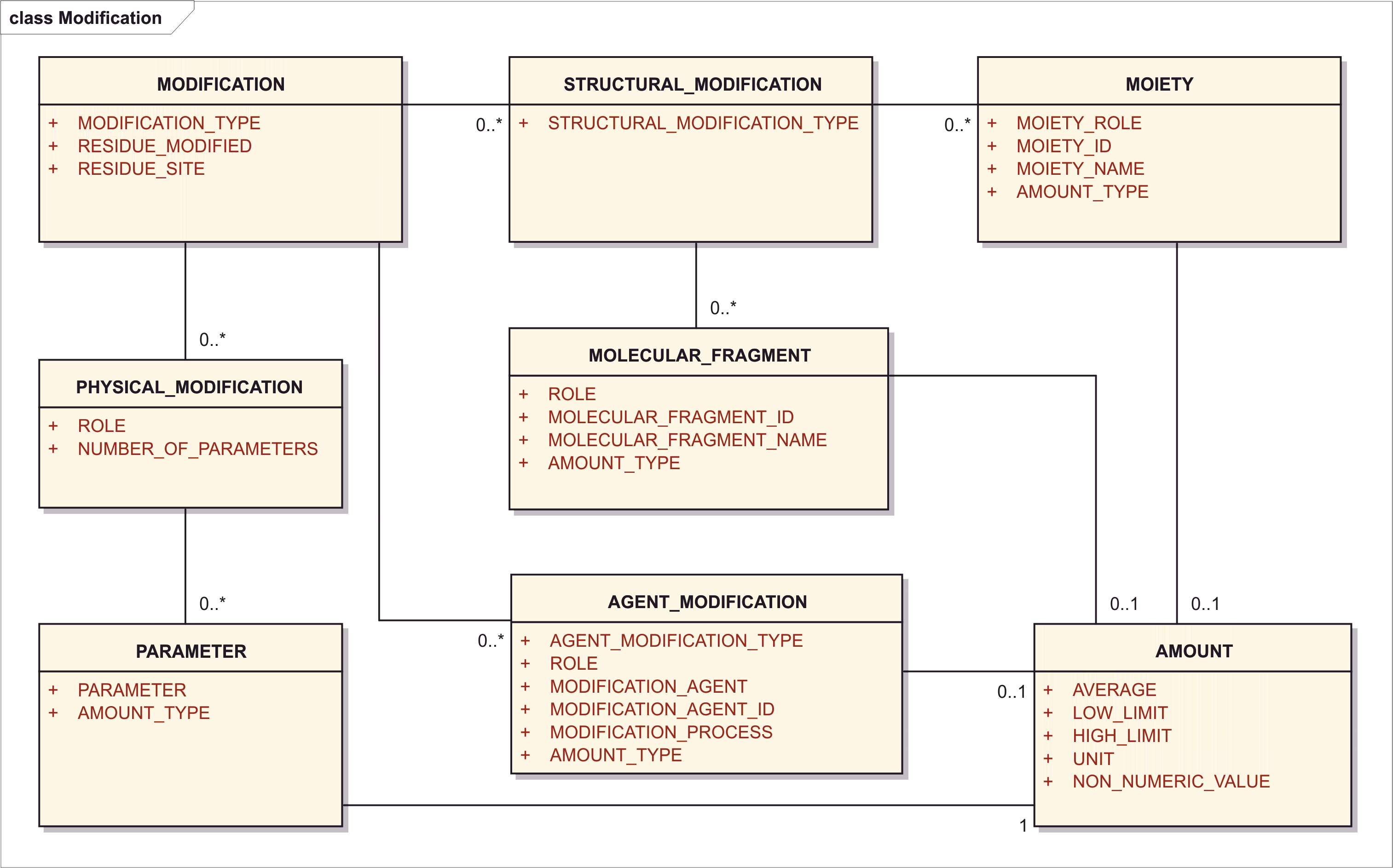


Figure 11– Information Model for Modifications

Specifically, information on the modification shall be provided by means of the following descriptors:

#### Modification Type

|  |  |
| --- | --- |
| **User Guidance** | To provide information on the general classification of modification. |
| **Example(s)** | physical, chemical, enzymatic |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

The modification type becomes apparent by its description. Structural modification is described using the structural elements <moiety> and <bond>. Agent and physical modifications are described by the <derivationProcess> element.

#### Residue Modified

|  |  |
| --- | --- |
| **User Guidance** | Every different amino acid modified is a separate modified residue.  NOTE: For virus or bacteria the value residue modified is not applicable |
| **Example(s)** | 20 amino acids plus 5 nucleotides |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | For proteins, the controlled vocabulary for amino acids applies; for nucleic acid, the Nucleotide Bases controlled vocabulary applies; for polymers, the controlled vocabulary for substance names applies; for structurally diverse substance or Mixture, the value residue modified equals to: not applicable |

The modified residue should be specified by bond elements, because if it is known then the moiety / molecular fragment will also be known at least in general terms (as in glycosylation). In that case bonds and distal moieties can be specified. If the position is not known, but we want to specify that certain types of amino acids are being modified this can be accomplished in the bond code. Further, the quantity of the bond element can be specified.

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <bond>  
 <code code="*Bond Type Code specifying* ***Residue Modified***"  
 codeSystem="*Code System (OID)*"  
 displayName="*Bond Type Name*"/>  
 <quantity><!-- amount of structural modification --></quantity>  
 <distalMoiety>  
 <code code="*Molecular Fragment (Substance) Id*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>*Molecular Fragment Name*</name>  
 </distalMoiety>

#### Residue Site

|  |  |
| --- | --- |
| **User Guidance** | The position of specific residue undergone to modifications shall be described. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <bond>  
 <code code="*Bond Type Code*" codeSystem="*Code System (OID)*"  
 displayName="*Bond Type Name*"/>  
 <position value="***Residue Site***"/>

#### Structural Modification

Modifications which can be structurally defined shall be described by means of the following descriptors:

##### Structural Modification Type

|  |  |
| --- | --- |
| **User Guidance** | Refers to the type of the structural modification. |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | CD |
| **Business Rule(s)** |  |

##### Moiety

The moiety is a defined chemical structure covalently linked to the protein. The information related to the moiety shall be provided as per specification described in the Section 8.10.9.2.

##### Molecular Fragment

Molecular Fragment and Moiety are the same thing and represented in the same way as Moiety.

###### Role

|  |  |
| --- | --- |
| **User Guidance** | The function of the molecular fragment. |
| **Example(s)** | antigen, linker, conjugate, etc. |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="***Molecular Fragment Role*** *Code*" codeSystem="*Code System (OID)*"  
 displayName="*Moiety Role Name*"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <code code="*Molecular Fragment (Substance) Id*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>*Molecular Fragment Name*</name>

###### Molecular Fragment ID

|  |  |
| --- | --- |
| **User Guidance** | The Substance Id of the molecular fragment. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Molecular Fragment Role Code*" codeSystem="*Code System (OID)*"  
 displayName="*Moiety Role Name*"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <code code="***Molecular Fragment*** *(Substance)* ***Id***"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>*Molecular Fragment Name*</name>

###### Molecular Fragment Name

|  |  |
| --- | --- |
| **User Guidance** | The name for molecular fragment shall be specified. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Molecular Fragment Role Code*" codeSystem="*Code System (OID)*"  
 displayName="*Moiety Role Name*"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <code code="*Molecular Fragment (Substance) Id*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>***Molecular Fragment Name***</name>

###### Amount

The information related to the amount shall be provided as per specification described in Section 8.8.

##### Agent Modification

This type describes modifications that do not result in the addition of a single well defined chemical moiety (i.e. formaldehyde, glutaraldehyde treatment, peroxide treatment). The following descriptors would be used to deine such a modification.

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <productOf>  
 <quantity><!-- Amount of Resulting Substance --></quantity>  
 <derivationProcess>  
 <code code="*Agent Modification Type and Process Code*"   
 codeSystem="*Code System (OID)*"  
 displayName="*Modification Type Display Name*"/>  
 <interactor>  
 <quantity><!-- Amount of Agent --></quantity>  
 <identifiedSubstance>  
 <id extension="*Agent Substance Id*"   
 root="*Substance Id System (OID)*">  
 <identifiedSubstance>  
 <code code="*Agent Substance Id*"   
 codeSystem="*Substance Id System (OID)*">  
 <name>*Agent Name (one primary)*</name>  
 </identifiedSubstance>  
 </identifiedSubject>  
 </interactor>  
 <controlVariable>  
 <characteristic>  
 <code code="*Parameter Code*"   
 codeSystem="*Code System (OID)*"  
 displayName="*Parameter Display Name*"/>  
 <value *Parameter Value* />  
 </characteristic>  
 </controlVariable>  
 <controlVariable>  
 <!-- Parameter as in Physical Modification -->  
 </controlVariable>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Molecular Weight’ with ‘method’*"   
 codeSystem="*Code System (OID)*"  
 displayName="molecular mass"/>  
 <value xsi:type="PQ" value="*Molecular Weight*" unit="g/mol"/>

###### Agent Modification Type

|  |  |
| --- | --- |
| **User Guidance** | Refers to the type of the agent caused the modification. |
| **Example(s)** | chemical, enzymatic, immunological, organism |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <productOf>  
 <quantity><!-- Amount of Resulting Substance --></quantity>  
 <derivationProcess>  
 <code code="***Agent Modification Type*** *and Process Code*"   
 codeSystem="*Code System (OID)*"  
 displayName="*Modification Type Display Name*"/>

###### Role

|  |  |
| --- | --- |
| **User Guidance** | For proteins, agent (a chemical that results in non-specific modifications of a protein) or moiety (a specific moiety added to a protein molecule, e.g., bis-mono-methoxy branched polethylene glycol 40000). |
| **Example(s)** | antigen, linker, conjugate, etc. |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

###### Modification Agent

|  |  |
| --- | --- |
| **User Guidance** | Established or primary name of the modifying agent |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <productOf>  
 <quantity><!-- Amount of Resulting Substance --></quantity>  
 <derivationProcess>  
 <interactor>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="*Modification Agent Substance Id*"   
 codeSystem="*Substance Id System (OID)*">  
 <name>***Modification Agent******Name*** *(one primary)*</name>  
 </identifiedSubstance>

###### Modification Agent ID

|  |  |
| --- | --- |
| **User Guidance** | Unique identifier for modifying agent. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | II |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <productOf>  
 <quantity><!-- Amount of Resulting Substance --></quantity>  
 <derivationProcess>  
 <interactor>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Modification Agent*** *Substance Id*"   
 codeSystem="*Substance Id System (OID)*">  
 <name>*Modification Agent Name (one primary)*</name>  
 </identifiedSubstance>

###### Modification Process

|  |  |
| --- | --- |
| **User Guidance** | Refers to the description of the modification process. |
| **Example(s)** | cell culture, UV irradiation, etc. |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <productOf>  
 <quantity><!-- Amount of Resulting Substance --></quantity>  
 <derivationProcess>  
 <code code="*Agent* ***Modification*** *Type and* ***Process*** *Code*"   
 codeSystem="*Code System (OID)*"  
 displayName="*Modification Type Display Name*"/>  
 <interactor>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="*Modification Agent Substance Id*"   
 codeSystem="*Substance Id System (OID)*">  
 <name>*Modification Agent Name (one primary)*</name>  
 </identifiedSubstance>

###### Amount

The information related to the amount shall be provided as per specification described in Section 8.8.

##### Physical Modification

The physical modification information shall be provided by means of the following descriptors:

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Physical Modification Type Code*"   
 codeSystem="*Code System (OID)*"  
 displayName="*Modification Type Display Name*"/>  
 <controlVariable>  
 <characteristic>  
 <code code="*Parameter Code*"   
 codeSystem="*Code System (OID)*"  
 displayName="*Parameter Display Name*"/>  
 <value *Parameter Value* />  
 </characteristic>  
 </controlVariable>  
 <controlVariable><!-- Another Parameter --></controlVariable>  
 ...

###### Role

|  |  |
| --- | --- |
| **User Guidance** | To be used to describe the function of the modification. |
| **Example(s)** | inactivation, activation |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Physical Modification Type Code with* ***Role***"   
 codeSystem="*Code System (OID)*"  
 displayName="*Modification Type Display Name*"/>  
 <controlVariable>  
 <characteristic>  
 <code code="*Parameter Code*"   
 codeSystem="*Code System (OID)*"  
 displayName="*Parameter Display Name*"/>  
 <value *Parameter Value* />  
 </characteristic>  
 </controlVariable>  
 <controlVariable><!-- Another Parameter --></controlVariable>  
 ...

###### Number of Parameters

|  |  |
| --- | --- |
| **User Guidance** | The number of parameter that is going to be described as result of the modification shall be provided. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | This number can be counted and is never explicitly represented. |

###### Parameters

Primarily used for nonspecific modification, specifies how a modification is quantified or extent of physical treatment.

Parameter

|  |  |
| --- | --- |
| **User Guidance** | Refers to the conditions under with the modification has been produced. |
| **Example(s)** | time, temperature, etc. |
| **Conformance** | CONDITIONAL – required parameters depend on process. |
| **Data Type** |  |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <productOf>  
 <derivationProcess>  
 <controlVariable>  
 <characteristic>  
 <code code="***Parameter*** *Code*"   
 codeSystem="*Code System (OID)*"  
 displayName="*Parameter Display Name*"/>  
 <value *Parameter Value* />  
 </characteristic>  
 </controlVariable>  
 <controlVariable><!-- Another Parameter --></controlVariable>  
 ...

###### Amount

The information related to the amount shall be provided as per specification described in Section 8.9.

<identifiedSubstance>  
 <identifiedSubstance><!-- Structural Details --></identifiedSubstance>  
 <productOf>  
 <derivationProcess>  
 <controlVariable>  
 <characteristic>  
 <value ***Parameter Value*** />

## Nucleic Acids

Nucleic acids are defined by nucleotide sequence, type of nucleic acid ( i.e DNA, RNA, DNA plasmid, modified sugars, and linkages, etc.) and any covalent modifications that exist in the substance (i.e. modified bases, sugars, or phosphate linkages). For transcribed genes, individual functional elements will also be captured (i.e. promoter, enhancer, coded gene, etc.) and assigned a substance ID. The location of each functional element within the gene will also be captured when defining the entire gene. The coding strand sequence of a functional gene is sufficient to define a substance.

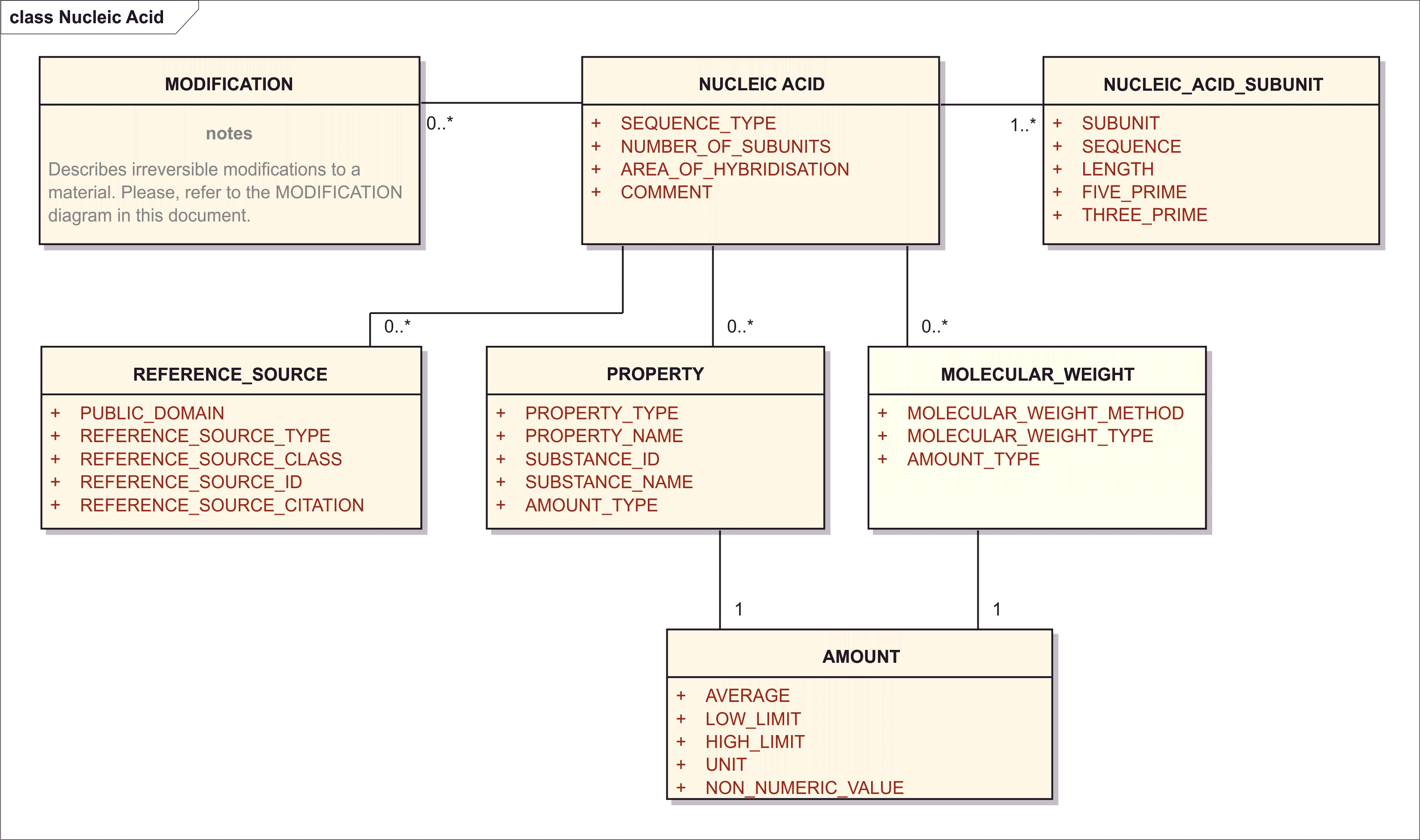


Figure 12– Information Model for Nucleic Acid

The Nucleic Acid set of elements are only to be used to describe nucleic acids that have:

- A length greater than three bases or base pairs;

- Elements to be used for oligonucleotides

- Genes used in gene therapy

- Any nucleic acid aptomers

For the description of nucleic acids the following information are mandatory to assign substance IDs:

- Nucleotide Sequence

- Nucleic Acid Type (RNA, DNA, plasmid, single or double stranded)

- Any Modifications to the substance either sequence specific (i.e. modified or unnatural base, sugar, phosphate linkage) or non-sequence specific modifications.

For the description of genes used in gene therapy the following information are mandatory to assign substance IDs:

- Gene ID (source NCBI, EMBL etc., gene id and name)

- Gene Elements

- Gene Element IDs

- Gene Loci

- Loci of each Gene Element within the gene used for Gene Therapy

- Species, Strain from which the Gene or Gene Element was derived

- Complete nucleotide sequence of the Gene and/or Gene Elements

- Any modifications made to the sequence and the purpose of the modification (i.e. codon optimization, enhancing or decreasing transcriptional activity).

The information shall be provided by means of the followings data elements:

### Substance Name

The information related to the name of the substance shall be provided as per specification described in Section 8.2.

### Substance Code

The information related to the code of the substance shall be provided as per specification described in Section 8.3.

### Reference Information

The information related to the code of the substance shall be provided as per specification described in Section 8.5.

### Version

The information related to the version shall be provided as per specification described in Section 8.4.

### Structure

The structure is only to be used for the description of short peptides that frequently have non-natural amino-acids. When proteins are fully specified by the Amino Acid letter sequence with disulfide-bonds and modifications, the structure file is not to be used.

The information related to the structure of the substance shall be provided as per specification described in Section 8.7.

In the HL7 SPL model, the Nucleotide sequences are specified analogously to Proteins:

* each chain (sequence) is one moiety
* modifications are specified by substitution elements (the bond element)
  + where the bond code defines what substitution occurs, including which fragments from the bonded moieties are being deleted (e.g, in an amid bond, X–OH + a H-N-Y --> X-N-Y deletes the hydroxyl group from the proximal moiety and a hydrogen from the distal moiety.)
  + distal moiety identifies the substituted fragment by code or by reference to a moiety id, whereby the referenced moiety may be another moiety in the substance, e.g., another sequence sub-unit, or the same moiety as the proximal moiety for cycle-forming bonds.
  + a pair of position number specifies the position in the sequence of the proximal chain and the distal chain that are so liked.

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Nucleotide Sequence Sub-Unit’*"   
 codeSystem="*Code System (OID)*"  
 displayName="Nucleotide sequence sub-unit"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <id extension="*Subunit Label (Local Id)*"   
 root="*UUID of Document Id or Set Id*"/>  
 </partMoiety>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Nucleotide Sequence’*"   
 codeSystem="*Code System (OID)*"  
 displayName="nucleotide sequence"/>  
 <value xsi:type="ED" mediaType="application/x-dna-seq">  
 <!-- Nucleotide Sequence in Single Letter Notation -->  
 </value>  
 </characteristic>  
 </subjectOf>  
 </moiety>  
 <moiety><!-- Next chain, same content as above --></moiety>

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Nucleotide Sub-Unit’*"   
 codeSystem="*Code System (OID)*"  
 displayName="protein sub-unit"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <id extension="*Subunit Label (Local Id)*"   
 root="*UUID of Document Id or Set Id*"/>  
 <bond>  
 <code code="*Bond Type Code*"   
 codeSystem="*Code System (OID)*"  
 displayName="*Bond Type Display Name*"/>  
 <positionNumber value="*Position on the Proximal Moiety*"/>  
 <positionNumber value="*Position on the Distal Moiety*"/>  
 <distalMoiety><!-- Other Moiety Id or Details --></distalMoiety>  
 </bond>

The isotope information shall be provided proteins that contain a radionuclide or a non-natural isotopic ratio. The information related to the isotope shall be provided as per specification described in Section 8.7.7.

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Nucleotide Sub-Unit’*"   
 codeSystem="*Code System (OID)*"  
 displayName="protein sub-unit"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <id extension="*Subunit Label (Local Id)*"   
 root="*UUID of Document Id or Set Id*"/>  
 <bond>  
 <code code="*Code meaning ‘Isotope Substitution’*"  
 codeSystem="*Code System (OID)*"/>  
 <position value="*Position in Sequence*"/> <distalMoiety>  
 <code code="*Nuclide Id*" codeSystem="*Nuclide Id System (OID)*"/>

### Sequence Type

|  |  |
| --- | --- |
| **User Guidance** | The type of the sequence shall be specified based on a controlled vocabulary. |
| **Example(s)** | complete / partial |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Nucleotide Sequence’ with* ***Sequence Type***"   
 codeSystem="*Code System (OID)*"  
 displayName="*Name for Nucleotide Sequence Complete/Partial*"/>

### Number of Subunits

|  |  |
| --- | --- |
| **User Guidance** | The number of linear sequences of nucleotides linked through phosphodiesteric bonds shall be described.  NOTE: If not specified in the reference source, the assumption is that there is 1 subunit. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | This number can be counted and is never explicitly represented. |

### Area of hybridisation

|  |  |
| --- | --- |
| **User Guidance** | The area of hydbridization shall be described if applicable. The number associated with the subunit followed by the number associated to the residue shall be specified in increasing order.  The underscore “\_” shall be used as separator as follow: “Subunit\_number Residue”. |
| **Example(s)** | 1\_10 refers to the 10th residue on the 1st subunit  or,  A1\_10, A1\_11  Each residue shall be specified followed by a comma and space. |
| **Conformance** | CONDITIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

### Comment

|  |  |
| --- | --- |
| **User Guidance** | Any comment can be provided in this field, if necessary. |
| **Example(s)** |  |
| **Conformance** | OPTIONAL – DISCOURAGED |
| **Data Type** | ST |
| **Business Rule(s)** |  |

Note: Comment can always be packaged into a characteristic:

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Other Details of Structural Unit --></partMoiety>  
 <subjectOf>  
 <characteristic>  
 <text>***Comment*** *Text*</text>  
 </characteristic>  
 </subjectOf>

### Nucleic Acid Subunit (repeat as necessary)

Subunits are listed in order of decreasing length; sequences of the same length will be ordered by molecular weight; subunits that have identical sequences will be repeated multiple times.

#### Subunit (Index)

|  |  |
| --- | --- |
| **User Guidance** | Index of linear sequences of nucleic acids linked through phosphate bonds in order of decreasing length. Sequences of the same length will be ordered by molecular weight. Subunits that have identical sequences will be repeated and have sequential subscripts. |
| **Example(s)** | 1, 2, 3, ... |
| **Conformance** | REQUIRED |
| **Data Type** | INT |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Protein Sub-Unit’*"   
 codeSystem="*Code System (OID)*"  
 displayName="protein sub-unit"/>  
 <partMoiety>  
 <id extension="***Subunit Index***"   
 root="*UUID of Document Id or Set Id*"/>

#### Sequence

|  |  |
| --- | --- |
| **User Guidance** | Actual nucleotide sequence notation from 5' to 3' end using standard single letter codes. In addition to the base sequence, sugar and type of phosphate or non-phosphate linkage should also be captured. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | ST |
| **Business Rule(s)** | Sequence of the letters G, C, T, A, and U |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Nucleic Acid Sub-Unit’*"   
 codeSystem="*Code System (OID)*"  
 displayName="nucleic acid sub-unit"/>  
 <partMoiety ... />  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Nucleotide Sequence’ with Sequence Type*"   
 codeSystem="*Code System (OID)*"  
 displayName="nucleotide sequence"/>  
 <value xsi:type="ED" mediaType="application/x-dna-seq">  
 <!-- Nucleotide Sequence in Single Letter Notation -->  
 </value>

#### Length

|  |  |
| --- | --- |
| **User Guidance** | The number of nucleotide residues linked through phosphate esters shall be specified. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | The sequence length is implicitly in the length property of the HL7 ST data type. |

#### Five Prime

|  |  |
| --- | --- |
| **User Guidance** | The nucleotide present at the 5’ terminal shall be specified based on a controlled vocabulary.  NOTE: If not specified in the available reference source this value shall be derived from the nucleic acid sequence. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | CD |
| **Business Rule(s)** | Since the sequence is represented from the 5' to the 3' end, the 5’ prime nucleotide is the letter at the first position in the sequence. A separate representation would be redundant. |

#### Three Prime

|  |  |
| --- | --- |
| **User Guidance** | The nucleotide present at the 3’ terminal shall be specified based on a controlled vocabulary.  NOTE: If not specified in the available reference source this value shall be derived from the nucleic acid sequence. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | Since the sequence is represented from the 5' to the 3' end, the 5’ prime nucleotide is the letter at the last position in the sequence. A separate representation would be redundant. |

### Modification (repeat as necessary)

The information related to the Modification shall be provided as per specification described in Section 8.11.14.

### Molecular Weight (repeat as necessary)

The information related to the molecular weight shall be provided as per specification described in Section 8.11.13.

### Property (repeat as necessary)

The information related to the property shall be provided as per specification described in Section 8.9.

### Reference Source (repeat as necessary)

The information related to the reference source shall be provided as per specification described in Section 8.2.5.

## Polymers

Polydisperse polymers are molecular ensembles defined using a combination structural and descriptive elements. The structural elements are the polymerized repeating units along with end groups and salt forms. The number and/or weight average molecular weight, degree of polymerization, degree or extent of substitution and/or physical properties related to molecular weight are descriptive elements that may be required to assign a substance ID. For synthetic polymers, the monomers used to prepare the polymer should also be provided. For copolymers, the type of copolymer (random, block, branched, etc) along with the block/branch size and number of blocks or branches should also be provided. A specification sheet and/or information sheet from a manufacturer is often sufficient to generate a substance ID. For polymers derived from biological matrixes, the source of the polymer should also be provided. Starch, for example, is normally isolated from wheat, corn, rice, potato, or tapioca.

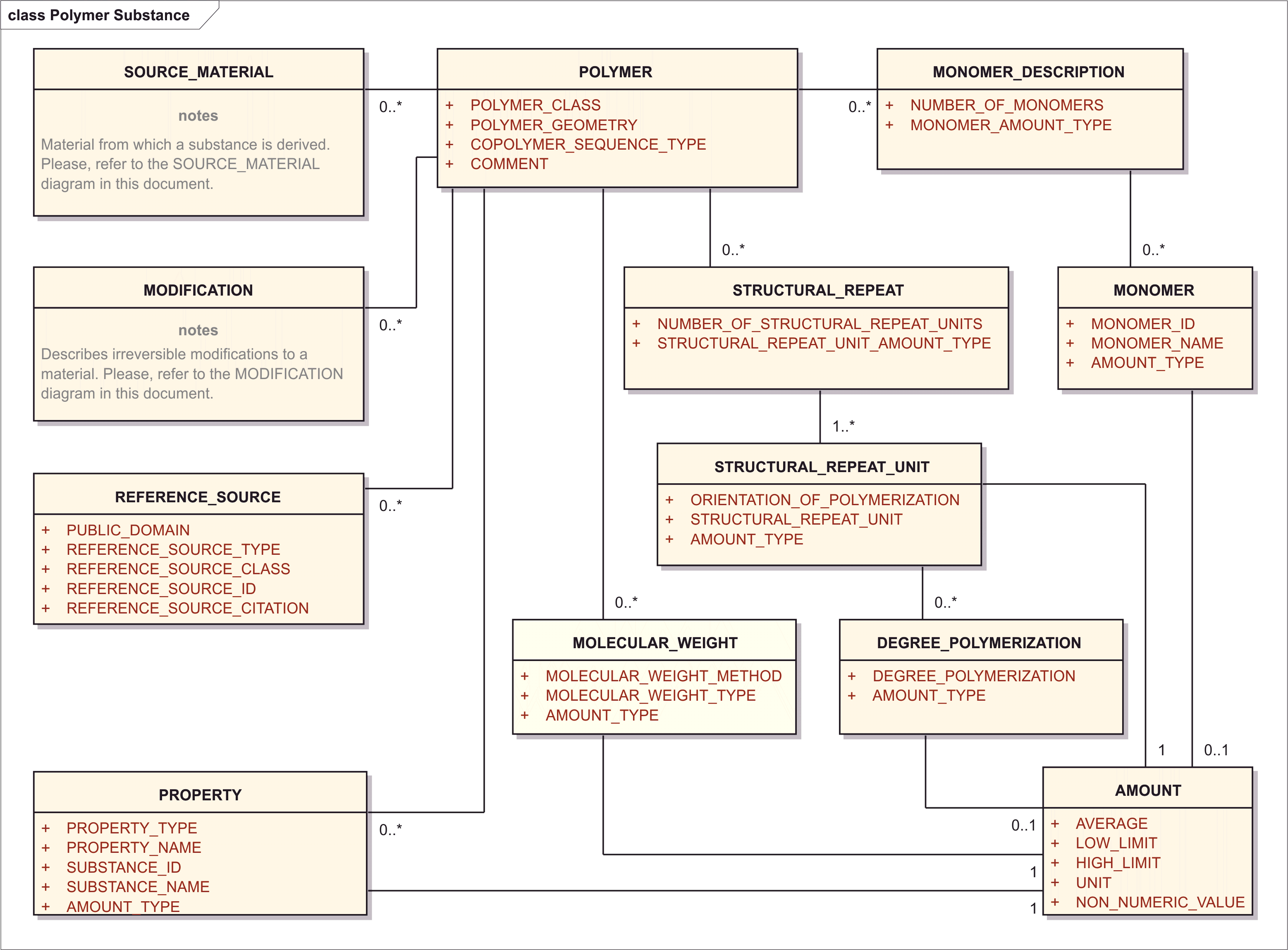


Figure 13– Information Model for Polymer Substance

For the description of polymers the following information are mandatory to assign substance IDs:

* Graphical Representation of the Repeating Units and/or Monomers
* Number Average Molecular Weight and/or Limits
* Weight Average Molecular Weight and/or Limits
* Degree of Polymerization
* Physical Properties Related to Molecular Structure (i.e. Viscosity, Density, etc.)
* For Copolymers: Relative Amounts of Monomers
* For Block Copolymers: Block Size and Block Number
* For Substituted Polymers: Substituents and Degree of Substitution
* Any Process that Modifies the Chemical Structure

The general set of information shall be provided by means of the followings data elements:

### Substance Name

The information related to the name of the substance shall be provided as per specification described in Section 8.2.

### Substance Code

The information related to the code of the substance shall be provided as per specification described in Section 8.3.

### Reference Information

The information related to the reference information of the substance shall be provided as per specification described in Section 8.5.

### Version

The information related to the version of the substance shall be provided as per specification described in Section 8.4.

### Structure

The information related to the structure shall be provided as per specification described in Section 8.7.

### Polymer Class

|  |  |
| --- | --- |
| **User Guidance** | Type of polymer. |
| **Example(s)** | homopolymer, copolymer |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <asSpecializedKind>  
 <definingMaterialKind>  
 <code code="***Polymer Class*** *Code*" codeSystem="*Code System (OID)*"/>  
 </definingMaterialKind>

### Polymer Geometry

|  |  |
| --- | --- |
| **User Guidance** | The geometry of the polymer shall be described. |
| **Example(s)** | linear, branched, cross linked, network, dendritic, star, comb, brush |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Parts --></identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Polymer Geometry’*"   
 codeSystem="*Code System (OID)*"  
 displayName="polymer geometry"/>  
 <value code="***Polymer Geometry*** *Code*" xsi:type="CV"  
 codeSystem="*Polymer Geometry Code System (OID)*"  
 displayName="*Polymer Geometry Name*"/>

### Copolymer Sequence type

|  |  |
| --- | --- |
| **User Guidance** | The type of copolymer of the polymer sequence shall be described |
| **Example(s)** | random, alternating, block, graft, statistical |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Parts --></identifiedSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Copolymer Sequence Type’*"   
 codeSystem="*Code System (OID)*"  
 displayName="Copolymer Sequence Type"/>  
 <value code="***Copolymer Sequence Type*** *Code*" xsi:type="CV"  
 codeSystem="*Copolymer Sequence Type Code System (OID)*"  
 displayName="*Copolymer Sequence Type Name*"/>

### Comment

|  |  |
| --- | --- |
| **User Guidance** | Any comment can be provided in this field, if necessary. |
| **Example(s)** |  |
| **Conformance** | OPTIONAL – DISCOURAGED |
| **Data Type** | ST |
| **Business Rule(s)** |  |

Note: Comment can always be packaged into a characteristic:

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Other Details of Structural Unit --></partMoiety>  
 <subjectOf>  
 <characteristic>  
 <text>***Comment*** *Text*</text>  
 </characteristic>  
 </subjectOf>

### Monomer Description (repeat as necessary)

For synthetic polymers this set of descriptors intends to specify and quantify the monomers used for the synthesis of the polymer or copolymer. The information on the monomer shall be provided by means of the following data elements:

#### Number of Monomers

|  |  |
| --- | --- |
| **User Guidance** | The number of diverse monomers used to synthesize the polymer shall be specified. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | This number can be counted and is never explicitly represented. |

#### Monomer (repeat as necessary)

This section identifies and quantifies the monomer(s) used in the synthesis of the polymer.

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <quantity value="*Amount (Average) of Polymer*" unit="*Amount Unit*"/>  
 <derivationProcess>  
 <code code="*Code meaning ‘Polymerization of Monomers’*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <quantity value="*Amount (Average) of Monomer*" unit="*Amount Unit*"/>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="*Monomer Substance Id*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>*Monomer Name*</name>  
 </identifiedSubstance>  
 </identifiedSubstance>  
 </interactor typeCode>  
 <interactor typeCode="CSM" ... *Additional Monomer* .../>

##### Monomer ID

|  |  |
| --- | --- |
| **User Guidance** | The unique identifier assigned to the monomer. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Monomer*** *Substance* ***Id***"   
 codeSystem="*Substance Id System (OID)*"/>

##### Monomer Name

|  |  |
| --- | --- |
| **User Guidance** | The name of the monomer shall be provided; established or primary name of monomer. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <name>***Monomer Name***</name>

##### Amount

The information related to the amount shall be provided as per specification described in Section 8.8.

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <quantity value="***Amount*** *(Average) of Polymer*" unit="*Amount Unit*"/>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <quantity value="***Amount*** *(Average) of Monomer*" unit="*Amount Unit*"/>

### Structural Repeat (repeat as necessary)

This section specifies and quantifies the repeated units and their configuration. A graphical representation is also captured. Information on the structural repeat unit shall be provided by means of the following data elements:

#### Number of Structural Repeat Units

|  |  |
| --- | --- |
| **User Guidance** | The number of diverse repeated units represented in the structure of the polymer shall be specified; the number of time that the monomer is repeated in the polymer. |
| **Example(s)** | PEG 40 is 40 |
| **Conformance** | IMPLIED |
| **Data Type** | INT |
| **Business Rule(s)** | This number can be counted and is never explicitly represented. |

#### Structural Repeat Unit (repeat as necessary)

Structural repeat units are represented as moieties:

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Structural Repeat Unit’ with orientation*"   
 codeSystem="*Code System (OID)*"  
 displayName="polymer repeat-unit"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <id extension="*Repeat Unit Label (Local Id, e.g. A, B, C, etc.)*"   
 root="*UUID of Document Id or Set Id*"/>  
 </partMoiety>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Molecular Structure’*"   
 codeSystem="*Code System (OID)*"  
 displayName="structure"/>  
 <value xsi:type="ED" mediaType="application/x-smiles">  
 <!-- Structure Representation of Repeat Unit -->  
 </value>  
 </characteristic>  
 </subjectOf>  
 </moiety>  
 <moiety><!-- Next repeat unit, same content as above --></moiety>

##### Orientation of Polymerization

|  |  |
| --- | --- |
| **User Guidance** | The orientation of the polymerization shall be described. |
| **Example(s)** | head-tail, head-head, random |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Structural Repeat Unit’ with* ***Orientation***"   
 codeSystem="*Code System (OID)*"  
 displayName="polymer repeat-unit, *orientation*"/>

##### Structural Repeat Unit

|  |  |
| --- | --- |
| **User Guidance** | (A, B, C,…) relates back to the repeating units described in the structure in order of decreasing molecular weight; |
| **Example(s)** | a,b,c,... |
| **Conformance** | OPTIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Structural Repeat Unit’ with orientation*"   
 codeSystem="*Code System (OID)*"  
 displayName="polymer repeat-unit"/>  
 <partMoiety>  
 <id extension="***Structural Repeat Unit*** *Label (Local Id, e.g. A, B, C, etc.)*"   
 root="*UUID of Document Id or Set Id*"/>  
 </partMoiety>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Molecular Structure’*"   
 codeSystem="*Code System (OID)*"  
 displayName="structure"/>  
 <value xsi:type="ED" mediaType="application/x-smiles">  
 <!-- Structure Representation of Repeat Unit -->  
 </value>

##### Amount

The information related to the amount shall be provided as per specification described in Section 8.8.

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <quantity><!-- Moiety **Amount** --></quantity>

##### Degree of Polymerization (repeat as necessary)

This section applies to homopolymer and block co-polymers where the degree of polymerization within a block can be described. The following information shall be provided where available:

###### Degree Polymerization

|  |  |
| --- | --- |
| **User Guidance** | The type of the degree of polymerization shall be described. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Repeat Unit Label etc. --><partMoiety>  
 <subjectOf><!-- Repeat Unit Structural Representation --></subjectOf>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Type of Degree of Polymerization’*"   
 codeSystem="*Code System (OID)*"  
 displayName="Copolymer Sequence Type"/>  
 <value code="***Degree of Polymerization*** *(Code)*" xsi:type="CV"  
 codeSystem="*Code System (OID)*"  
 displayName="*Display Name*"/>

###### Amount

The information related to the amount shall be provided as per specification described in Section 8.8.

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Repeat Unit Label etc. --><partMoiety>  
 <subjectOf><!-- Repeat Unit Structural Representation --></subjectOf>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Degree of Polymerization (Quantitative)’*"   
 codeSystem="*Code System (OID)*"  
 displayName="Copolymer Sequence Type"/>  
 <value ***Amount*** />

### Molecular Weight (repeat as necessary)

The information related to the molecular weight shall be provided as per specification described in Section 8.11.13.

### Property (repeat as necessary)

The information related to the property of the polymer shall be provided as per specification described in Section 8.9.

### Reference Source (repeat as necessary)

The information related to the property of the polymer shall be provided as per specification described in Section 8.2.5.

### Source Material (repeat as necessary)

Source material is shall capture information on the taxonomic and anatomical origins as well as the fraction of a material that can result in or can be modified to form a substance. This set of data elements shall be used to define polymer substances isolated from biological matrices. Taxonomic and anatomical origins shall be described using controlled vocabulary as required. This information is captured for naturally derived polymers (i.e. starch) and structurally diverse substances.

The information model for the source material is shown in the following figure:

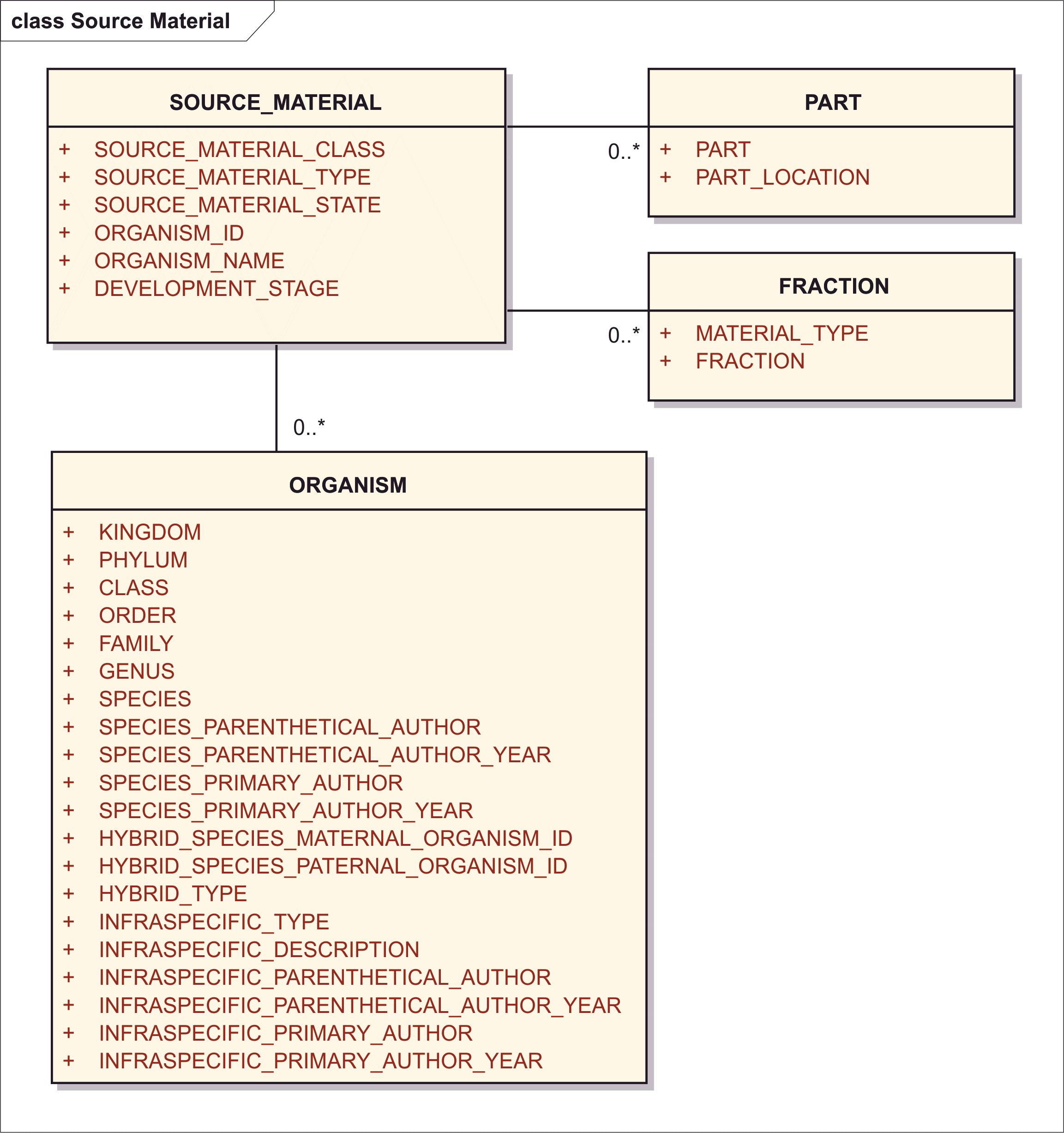


Figure 14– Information Model for Source Material

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Isolation from Matrix’ with method if needed*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <presentSubstance>  
 <code code="*Substance Id of Source Material (if available)*"   
 codeSystem="*Code System (OID)*"/>  
 <name><!-- Name of Source Material (if available) --></name>  
 <asSpecializedKind>  
 <definingMaterialKind>  
 <code code="*Source Material Class&Type Code*"   
 codeSystem="*Code System (OID)*" displayName="*Display Name*"/>  
 </definingMaterialKind>  
 </asSpecializedKind>  
 </presentSubstance>  
 <environmentOrganismStrain>  
 <code code="*Organism Code*"   
 codeSystem="*Organism Code System (OID)*"/>  
 <name><!-- Organism Name --></name>  
 <strainText>*Organism Strain*</strainText>  
 </environmentOrganismStrain>  
 </presentSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Developmental Stage’*"   
 codeSystem="*Code System (OID)*"/>  
 <value code="*Developmental Stage*" xsi:type="CV"  
 codeSystem="*Code System (OID)*"/>  
 </characteristic>  
 </subjectOf>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Source Material State’*"   
 codeSystem="*Code System (OID)*"/>  
 <value code="*Source Material State*" xsi:type="CV"  
 codeSystem="*Code System (OID)*"/>  
 </characteristic>  
 </subjectOf>  
 <productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Preparing the Source Material’* "   
 codeSystem="*Code System (OID)*"/>  
 <!-- Preparation Steps follow -->  
 <component>  
 <derivationProcess>  
 <code code="*Code meaning ‘Taking a Part’ with Location*"/>  
 codeSystem="*Code System (OID)*"/>  
 </derivationProcess>  
 </component>  
 <component>  
 <derivationProcess>  
 <code code="Code meaning ‘*Fractionation’ with Material Type*"/>  
 codeSystem="*Code System (OID)*"/>  
 </derivationProcess>  
 </component>

Specifically, the information on the source material shall be provided by means of the following set of data elements:

#### Source Material Class

|  |  |
| --- | --- |
| **User Guidance** | General classification of the source material; for vaccines this is the class of infectious agents. |
| **Example(s)** | organic, inorganic |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | The “source material class” values make different “Source\_material\_type” values applicable. Because organic / inorganic is just a class best captured in terminology of the source material types, only one data element is needed for Type which implies Class. |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Isolation from Matrix’ with method if needed*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <presentSubstance>  
 <asSpecializedKind>  
 <definingMaterialKind>  
 <code code="***Source Material Class****&Type Code*"   
 codeSystem="*Code System (OID)*" displayName="*Display Name*"/>  
 </definingMaterialKind>

#### Source Material Type

|  |  |
| --- | --- |
| **User Guidance** | The type of the source material shall be specified based on controlled vocabulary. For vaccines this section refers to the class of infectious agent. |
| **Example(s)** | bacterium, human, mammal, fungus, virus, plant |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Allowed Values** | * inorganic * organic   + plant   + virus   + bacterium   + fungus   + mammal     - human |
| **Business Rule(s)** | The “Source\_material\_type” values are presented based on the “source material class” values above. Because organic / inorganic is just a class best captured in terminology of the source material types, only one data element is needed for Type which implies Class. |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Isolation from Matrix’ with method if needed*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <presentSubstance>  
 <asSpecializedKind>  
 <definingMaterialKind>  
 <code code="***Source Material*** *Class&****Type*** *Code*"   
 codeSystem="*Code System (OID)*" displayName="*Display Name*"/>  
 </definingMaterialKind>

#### Source Material state

|  |  |
| --- | --- |
| **User Guidance** | The state of the source material when extracted. |
| **Example(s)** | live, activated, inactivated, attenuated, conjugated, live (attenuated). |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Isolation from Matrix’ with method if needed*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <presentSubstance>  
 <code code="*Substance Id of Source Material (if available)*"   
 codeSystem="*Code System (OID)*"/>  
 <name><!-- Name of Source Material (if available) --></name>  
 </presentSubstance>  
 </presentSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Source Material State’*"   
 codeSystem="*Code System (OID)*"/>  
 <value code="***Source Material State***" xsi:type="CV"  
 codeSystem="*Code System (OID)*"/>  
 </characteristic>

#### Organism ID

|  |  |
| --- | --- |
| **User Guidance** | The unique identifier associated with the source material parent organism shall be specified. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <environmentOrganismStrain>  
 <code code="***Organism Id***"   
 codeSystem="*Organism Code System (OID)*"/>  
 <name><!-- Organism Name --></name>  
 </environmentOrganismStrain>  
 </presentSubstance>

#### Organism Name

|  |  |
| --- | --- |
| **User Guidance** | The organism accepted Latin name shall be provided based on the binomial Latin nomenclature. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <environmentOrganismStrain>  
 <code code="*Organism Id*"   
 codeSystem="*Organism Code System (OID)*"/>  
 <name><!-- **Organism Name** --></name>  
 </environmentOrganismStrain>  
 </presentSubstance>

#### Development Stage

|  |  |
| --- | --- |
| **User Guidance** | Stage of life for animals, plants, insects and microorganisms. This information shall be provided only when the substance is significantly different in these stages (e.g., fetal bovine serum). |
| **Example(s)** | fetus, infant, juvenile, adult, senescent, leafing, pre-flowering, flowering, fruiting, etc. |
| **Conformance** | CONDITIONAL – if it is a distinguishing factor, even if it is implicitly understood to experts. |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance> ... </presentSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Developmental Stage’*"   
 codeSystem="*Code System (OID)*"/>  
 <value code="***Developmental Stage***" xsi:type="CV"  
 codeSystem="*Code System (OID)*"/>  
 </characteristic>  
 </subjectOf>

#### Part Description (repeat as necessary)

Information of the part of the material used to produce the substance shall be provided by means of the following data elements:

##### Part

|  |  |
| --- | --- |
| **User Guidance** | The portion of an organism with a definable anatomical location. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | Depending on the Source\_Material\_Type values (e.g. human/mammalian or plant) the relevant CV applies (e.g. organism or plant). |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Isolation from Matrix’ with method if needed*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <presentSubstance .../>  
 <environmentOrganismStrain .../>  
 </presentSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Preparing the Source Material’* "   
 codeSystem="*Code System (OID)*"/>  
 <!-- Preparation Steps follow -->  
 <component>  
 <derivationProcess>  
 <code code="*Code meaning ‘Taking a* ***Part****’ with Location*"/>  
 codeSystem="*Code System (OID)*"/>  
 </derivationProcess>  
 </component>

##### Part Location

|  |  |
| --- | --- |
| **User Guidance** | The detail anatomic location when the part can be extracted from different anatomical location of the organism. Multiple alternative locations may apply. |
| **Example(s)** | For cartilage: knee, elbow , stomach, shoulder |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | Applicable only if Source\_Material\_Type is human or mammalian |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Isolation from Matrix’ with method if needed*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <presentSubstance .../>  
 <environmentOrganismStrain .../>  
 </presentSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Preparing the Source Material’* "   
 codeSystem="*Code System (OID)*"/>  
 <!-- Preparation Steps follow -->  
 <component>  
 <derivationProcess>  
 <code code="*Code meaning ‘Taking a* ***Part’ with Location***"/>  
 codeSystem="*Code System (OID)*"/>  
 </derivationProcess>  
 </component>

#### Fraction (repeat as necessary)

##### Material type

|  |  |
| --- | --- |
| **User Guidance** | The specific type of the material constituting the component. |
| **Example(s)** | plasmid, extra-chromosomal, chondrocyte, lipase, triglycerides (..) |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Isolation from Matrix’ with method if needed*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <presentSubstance .../>  
 <environmentOrganismStrain .../>  
 </presentSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Preparing the Source Material’* "   
 codeSystem="*Code System (OID)*"/>  
 <!-- Preparation Steps follow -->  
 <component>  
 <derivationProcess>  
 <code code="Code meaning ‘*Fractionation’ with* ***Material Type***"/>  
 codeSystem="*Code System (OID)*"/>  
 </derivationProcess>  
 </component>

##### Fraction

|  |  |
| --- | --- |
| **User Guidance** | Displayed name; primarily used for gene therapy where the component is described within the nucleic acid description (e.g. P-LLO-E7; HPV-16).  For cell therapy this expresses the protein which is expressed within the relevant cell (e.g. IL12).  For conjugated vaccines this describe the organism strain of the carrier.  For vaccines, this may identify the antigen used in the vaccine.  Fractions of fractions may be specified. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

As per the description, this is a very diverse “field” with many different meanings. Most of the time they all fall under the Fractionation terminology:

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Isolation from Matrix’ with method if needed*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <presentSubstance .../>  
 <environmentOrganismStrain .../>  
 </presentSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Preparing the Source Material’* "   
 codeSystem="*Code System (OID)*"/>  
 <!-- Preparation Steps follow -->  
 <component>  
 <derivationProcess>  
 <code code="Code meaning ‘***Fractionation****’ with Material Type*"/>  
 codeSystem="*Code System (OID)*"/>  
 </derivationProcess>  
 </component>

Organism strain of the source organism is not about fractionation but about cultivation, however, the strain is specified in the organism:

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <environmentOrganismStrain>  
 <strainText>***Organism Strain***</strainText>

#### Organism (repeat as necessary)

This section describes the organism which the substance derived from. For vaccines, the parent organism shall be specified based on these section elements.

##### Kingdom

|  |  |
| --- | --- |
| **User Guidance** | The kingdom of an organism shall be specified. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | CD |
| **Business Rule(s)** | This is implied by the species is referenced from an appropriate terminology / taxonomy (e.g., the NCBI Entrez Taxonomy Id implies all higher levels.) |

##### Phylum

|  |  |
| --- | --- |
| **User Guidance** | The phylum of an organism shall be specified. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | CD |
| **Business Rule(s)** | This is implied by the species is referenced from an appropriate terminology / taxonomy (e.g., the NCBI Entrez Taxonomy Id implies all higher levels.) |

##### Class

|  |  |
| --- | --- |
| **User Guidance** | The class of an organism shall be specified. |
| **Example(s)** | Mammalia, maxillopoda, Sauropsida. |
| **Conformance** | IMPLICIT |
| **Data Type** | CD |
| **Business Rule(s)** | This is implied by the species is referenced from an appropriate terminology / taxonomy (e.g., the NCBI Entrez Taxonomy Id implies all higher levels.) |

##### Order

|  |  |
| --- | --- |
| **User Guidance** | The order of an organism shall be specified, |
| **Example(s)** | Primate |
| **Conformance** | IMPLICIT |
| **Data Type** | CD |
| **Business Rule(s)** | This is implied by the species is referenced from an appropriate terminology / taxonomy (e.g., the NCBI Entrez Taxonomy Id implies all higher levels.) |

##### Family

|  |  |
| --- | --- |
| **User Guidance** | The family of an organism shall be specified. |
| **Example(s)** | Homidae |
| **Conformance** | IMPLICIT |
| **Data Type** | CD |
| **Business Rule(s)** | This is implied by the species is referenced from an appropriate terminology / taxonomy (e.g., the NCBI Entrez Taxonomy Id implies all higher levels.) |

##### Genus

|  |  |
| --- | --- |
| **User Guidance** | The genus of an organism shall be specified; refers to the Latin epithet of the genus of the plant/animal; it is present in names for genera, species and infraspecies. |
| **Example(s)** | Tanacetum |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

##### Species

|  |  |
| --- | --- |
| **User Guidance** | The species of an organism shall be specified; refers to the Latin epithet of the species of the plant/animal; it is present in names for species and infraspecies. |
| **Example(s)** | Parthenium |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | Must be specified as a code from an appropriate taxonomy terminology (e.g., NCBI Taxonomy Id) |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <environmentOrganismStrain>  
 <asSpecializedKind>  
 <definingSpecies>  
 <code code="***Species*** *Code*" codeSystem="*Taxonomy Code System (OID)*"   
 displayName="*Species Name*"/>

##### Species Literature Reference

Species are described by original articles, and the taxonomy may vary over time, so the reference to the original article is critical. A reference to an article should be complete and not just using author name abbreviations. The ITIS database provides a good example for species defined by way of complete literature refrerneces.

The structure is the same as reference source, with the addition of adding structured authors and years:

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <environmentOrganismStrain>  
 <asSpecializedKind>  
 <definingSpecies .../>  
 <subjectOf>  
 <document>  
 <id extension="*Reference Source (Document) Id*"   
 root="*Reference Source Type Document Id System (OID)*"/>  
 <code code="*Reference Source Type (Code)*"   
 codeSystem="*Reference Source Type Code Sytem (OID)*"  
 displayName="*Reference Source Type Display Name*"/>  
 <text>  
 <reference value="*Reference Source Document Id (URL)*"/>  
 </text>  
 <effectiveTime value="*Publication Year*"/>  
 <bibliographicDesignationText>  
 *Reference Source Citation*</bibliographicDesignationText>  
 <author>  
 <assignedEntity>  
 <assignedPerson>  
 <name>*Author Name*</name>  
 </assignedPerson>  
 <representedOrganization>  
 <name>*Author Organization Name*</name>  
 </representedOrganization>  
 </assignedEntity>  
 </author>  
 </document>  
 </subjectOf>

###### Species Parenthetical Author

|  |  |
| --- | --- |
| **User Guidance** | The parenthetical author of an organism species shall be specified; refers to the first author, who published the plant/animal name (of any rank). |
| **Example(s)** | Linnaeus |
| **Conformance** | OPTIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <subjectOf>  
 <document>  
 <author>  
 <assignedEntity>  
 <assignedPerson>  
 <name>***Author Name***</name>

###### Species Parenthetical Author Year

|  |  |
| --- | --- |
| **User Guidance** | The parenthetical author year of an organism species an organism shall be specified; refers to the year in which the first author published the plant/animal name (of any rank).  For vaccines names, it refers to the year of isolation. |
| **Example(s)** | 2010/2011 |
| **Conformance** | OPTIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <subjectOf>  
 <document>  
 <effectiveTime value="***Publication Year***"/>

###### Species Primary Author

|  |  |
| --- | --- |
| **User Guidance** | The primary author of an organism species shall be specified; refers to the first author, who published the plant/animal name (of any rank). |
| **Example(s)** | Sch.Bip |
| **Conformance** | OPTIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <subjectOf>  
 <document>  
 <author>  
 <assignedEntity>  
 <assignedPerson>  
 <name>***Author Name***</name>

###### Species Primary Author Year

|  |  |
| --- | --- |
| **User Guidance** | The primary author year of an organism species shall be specified; refers to the year in which the first author published the plant/animal name (of any rank). |
| **Example(s)** |  |
| **Conformance** | OPTIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <subjectOf>  
 <document>  
 <effectiveTime value="***Publication Year***"/>

##### Hybrid Species Maternal Organism ID

|  |  |
| --- | --- |
| **User Guidance** | The identifier of the maternal species constituting the hybrid organism shall be specified based on a controlled vocabulary. |
| **Example(s)** |  |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Isolation from Matrix’ with method if needed*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <presentSubstance .../>  
 <environmentOrganismStrain .../>  
 </presentSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Preparing the Source Material’* "   
 codeSystem="*Code System (OID)*"/>  
 <!-- Preparation Steps follow -->  
 <component>  
 <derivationProcess>  
 <code code="*Code meaning ‘Hybridization’*"/>  
 codeSystem="*Code System (OID)*"/>  
 <directTarget>  
 <functionCode code="*Code for ‘Maternal Organism’*"   
 codeSystem="*Code System (OID)*"  
 displayName="maternal organism"/>  
 <presentSubstance>  
 <identifiedSubstance>  
 <code code="***Hybrid Species Maternal Organism Id***"   
 codeSystem="*Organism Id System (OID)*"/>  
 </identifiedSubstance>  
 </presentSubstance>  
 </directTarget>  
 <directTarget>  
 <functionCode code="*Code for ‘Paternal Organism’*"   
 codeSystem="*Code System (OID)*"  
 displayName="paternal organism"/>  
 <presentSubstance>  
 <identifiedSubstance>  
 <code code="*Hybrid Species Paternal Organism Id*"   
 codeSystem="*Organism Id System (OID)*"/>  
 </identifiedSubstance>  
 </presentSubstance>  
 </directTarget>  
 </derivationProcess>  
 </component>

##### Hybrid Species Paternal Organism ID

|  |  |
| --- | --- |
| **User Guidance** | The identifier of the paternal species constituting the hybrid organism shall be specified based on a controlled vocabulary. |
| **Example(s)** |  |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Isolation from Matrix’ with method if needed*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <presentSubstance .../>  
 <environmentOrganismStrain .../>  
 </presentSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Preparing the Source Material’* "   
 codeSystem="*Code System (OID)*"/>  
 <!-- Preparation Steps follow -->  
 <component>  
 <derivationProcess>  
 <code code="*Code meaning ‘Hybridization’*"/>  
 codeSystem="*Code System (OID)*"/>  
 <directTarget>  
 <functionCode code="*Code for ‘Maternal Organism’*"   
 codeSystem="*Code System (OID)*"  
 displayName="maternal organism"/>  
 <presentSubstance>  
 <identifiedSubstance>  
 <code code="*Hybrid Species Maternal Organism Id*"   
 codeSystem="*Organism Id System (OID)*"/>  
 </identifiedSubstance>  
 </presentSubstance>  
 </directTarget>  
 <directTarget>  
 <functionCode code="*Code for ‘Paternal Organism’*"   
 codeSystem="*Code System (OID)*"  
 displayName="paternal organism"/>  
 <presentSubstance>  
 <identifiedSubstance>  
 <code code="***Hybrid Species Paternal Organism Id***"   
 codeSystem="*Organism Id System (OID)*"/>  
 </identifiedSubstance>  
 </presentSubstance>  
 </directTarget>  
 </derivationProcess>  
 </component>

##### Hybrid Type

|  |  |
| --- | --- |
| **User Guidance** | The hybrid type of an organism shall be defined. |
| **Example(s)** | Yes, No |
| **Conformance** | IMPLICIT |
| **Data Type** | BL |
| **Business Rule(s)** | If the hybridization step is present, the organism is a hybrid. |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Isolation from Matrix’ with method if needed*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor typeCode="CSM">  
 <presentSubstance>  
 <presentSubstance .../>  
 <environmentOrganismStrain .../>  
 </presentSubstance>  
 <productOf>  
 <derivationProcess>  
 <code code="*Code meaning ‘Preparing the Source Material’* "   
 codeSystem="*Code System (OID)*"/>  
 <!-- Preparation Steps follow -->  
 **<component>  
 <derivationProcess>  
 <code code="*Code meaning ‘Hybridization’*"/>  
 codeSystem="*Code System (OID)*"/>**

##### Infraspecific Type

|  |  |
| --- | --- |
| **User Guidance** | The infraspecific type of an organism shall be specified. |
| **Example(s)** |  |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance> ... </presentSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Infraspecific Type’*"   
 codeSystem="*Code System (OID)*"/>  
 <value code="***Infraspecific Type***" xsi:type="CV"  
 codeSystem="*Code System (OID)*"/>  
 </characteristic>  
 </subjectOf>

##### Infraspecific Description

|  |  |
| --- | --- |
| **User Guidance** | The infraspecific description of an organism shall be specified based on a controlled vocabulary. For Influenza Vaccine, the infraspecifc description shall contain the syntax of the antigen in line with the WHO convention. |
| **Example(s)** | A/BRISBANE/59/2007(H1N1) |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance> ... </presentSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Infraspecific Description’*"   
 codeSystem="*Code System (OID)*"/>  
 <value code="***Infraspecific Description***" xsi:type="CV"  
 codeSystem="*Code System (OID)*"/>  
 </characteristic>  
 </subjectOf>

##### Infraspecific Literature Reference

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance> ... </presentSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Infraspecific Description’*"   
 codeSystem="*Code System (OID)*"/>  
 <value code="*Infraspecific Description*" xsi:type="CV"  
 codeSystem="*Code System (OID)*"/>  
 <subjectOf>  
 <document>  
 <id extension="*Reference Source (Document) Id*"   
 root="*Reference Source Type Document Id System (OID)*"/>  
 <code code="*Reference Source Type (Code)*"   
 codeSystem="*Reference Source Type Code Sytem (OID)*"  
 displayName="*Reference Source Type Display Name*"/>  
 <text>  
 <reference value="*Reference Source Document Id (URL)*"/>  
 </text>  
 <effectiveTime value="*Publication Year*"/>  
 <bibliographicDesignationText>  
 *Reference Source Citation*</bibliographicDesignationText>  
 <author>  
 <assignedEntity>  
 <assignedPerson>  
 <name>*Author Name*</name>  
 </assignedPerson>  
 <representedOrganization>  
 <name>*Author Organization Name*</name>  
 </representedOrganization>  
 </assignedEntity>  
 </author>  
 </document>  
 </subjectOf>  
 </characteristic>  
 </subjectOf>

###### Infraspecific Parenthetical Author

|  |  |
| --- | --- |
| **User Guidance** | The infraspecific parenthetical author of an organism shall be specified; refers to the first author, who published the infraspecific plant/animal name (of any rank). |
| **Example(s)** |  |
| **Conformance** | OPTIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance> ... </presentSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Infraspecific Description’*"   
 codeSystem="*Code System (OID)*"/>  
 <subjectOf>  
 <document>  
 <effectiveTime value="*Publication Year*"/>  
 <author>  
 <assignedEntity>  
 <assignedPerson>  
 <name>***Author Name***</name>

###### Infraspecific Parenthetical Author Year

|  |  |
| --- | --- |
| **User Guidance** | The infraspecific parenthetical author year of an organism shall be specified; refers to the year in which the first author published the infraspecific plant/animal name (of any rank). |
| **Example(s)** |  |
| **Conformance** | OPTIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance> ... </presentSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Infraspecific Description’*"   
 codeSystem="*Code System (OID)*"/>  
 <subjectOf>  
 <document>  
 <effectiveTime value="***Publication Year***"/>

###### Infraspecific Primary Author

|  |  |
| --- | --- |
| **User Guidance** | The infraspecific primary author of an organism shall be specified; refers to the first author, who published the infraspecific plant/animal name (of any infraspecific rank). |
| **Example(s)** |  |
| **Conformance** | OPTIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance> ... </presentSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Infraspecific Description’*"   
 codeSystem="*Code System (OID)*"/>  
 <subjectOf>  
 <document>  
 <effectiveTime value="*Publication Year*"/>  
 <author>  
 <assignedEntity>  
 <assignedPerson>  
 <name>***Author Name***</name>

###### Infraspecific Primary Author Year

|  |  |
| --- | --- |
| **User Guidance** | The infraspecific primary author year of an organism shall be specified; refers to the year in which the first author published the infraspecific plant/animal name (of any infraspecific rank). |
| **Example(s)** |  |
| **Conformance** | OPTIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!—Other Substance Details --></identifiedSubstance>

<productOf>  
 <derivationProcess>  
 <interactor typeCode="CSM">  
 <presentSubstance> ... </presentSubstance>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Infraspecific Description’*"   
 codeSystem="*Code System (OID)*"/>  
 <subjectOf>  
 <document>  
 <effectiveTime value="***Publication Year***"/>

### Modification (repeat as necessary)

Information on irreversible modifications of a polymer when derived from natural polymers shall be specified. Synthetic variations such as r-group modifications will be described structurally as specific fragments. Information related to the modification shall be provided as per specification described in Section 8.11.14.

## Structurally-Diverse Substance

This category is for the description of all substances that cannot be fully described by the other classes such as:

- Herbals and mineral

- Vaccines

- Purified blood products

- Immunoglobulins

- Substance derived from cells and tissues

For organism-based substances, the parent organism is essential defining information.

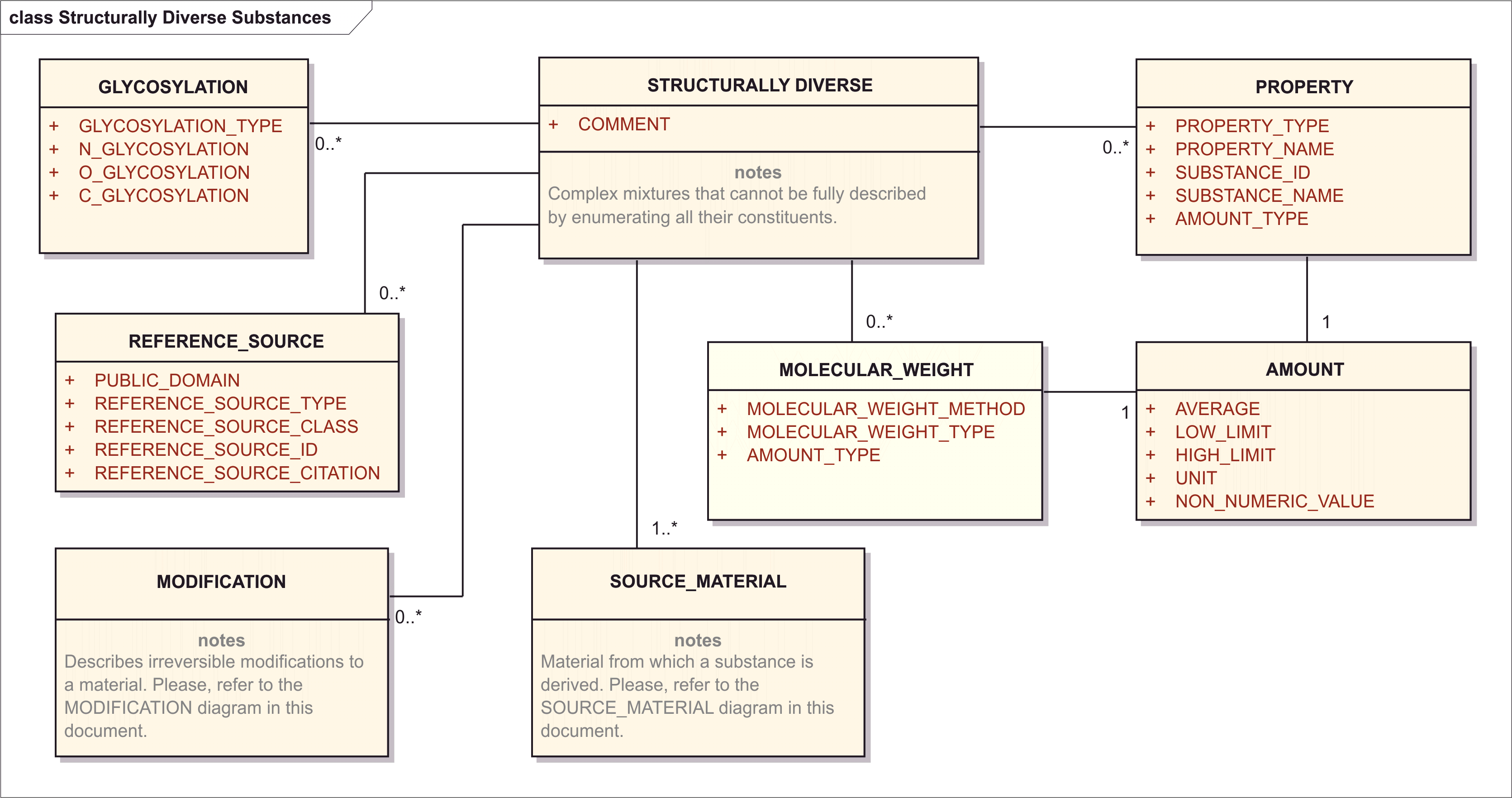


Figure 15– Information Model for Structurally-Diverse Substance

The majority of these substances will be derived from a biological organism but they could be complex natural materials such as coal tar or mineral oil. Herbals are typically described by parent organism genus, species, and part or parts. If specific parts of a plant are used, identification requires lists of individual parts such as flower, leaf, and stem or an indication of the plant life cycle segment such as flowering top. Because of variability in extraction processes (solvent, temperature, time), these processes are not captured at the substance level (broccoli and broccoli extract would be assigned the same SUBSTANCE ID). Time and place of harvest, type of soil and fertilizer, amount of daylight and water, and degree of plant maturity are also not captured at the substance level. A cultivar or variety of a plant may be defined as a different substance if known differences in constituents exist (e.g., broccoli and cauliflower are defined as different substances even though they share the same genus and species). Commodity oils, juices, and exudates such as olive oil, orange juice, or acacia are separate substances, but extracts such as coffee or tea map to the parent organism and part or parts due to their ambiguous nature. Oils and juices are described as organism part components rather than as parts so the part is necessary to accurately describe the oil or juice unless the entire organism is used. Olive oil is Olea Europaea fruit oil. Orange juice is Citrus sinensis fruit juice.

Some organisms such as influenza A virus require identification of subspecies, variety, strain, serovar, type, form, or cultivar group to accurately describe them. Relevant taxonomic identification numbers are helpful in parent organism identification.

Purified blood products (distinct clotting factors, human serum albumin) and monoclonal immunoglobulins are described as proteins. Polyclonal immunoglobulins are described as structurally diverse materials and require identification of the immunoglobulin type and targeted antigen.

Cells and tissues are also described as structurally diverse substances. Information on individual donors or extent of pooling is not captured at the substance level.

Many natural substances are modified chemically, physically, or biologically. A new substance will be generated if such a modification changes the chemical structure of one or more chemical entities in the substance. A polysaccharide conjugate vaccine is described as a component of an organism part chemically conjugated to another substance or substances. The type of conjugation chemistry and the identity of the conjugated chemical entity or entities are needed to describe the resulting structurally diverse conjugate. When a genetically-modified organism or cell is used, the inserted gene biological genetic modifications must be described.

Structurally-diverse substances not derived from organisms often require property descriptions as defining characteristics. In this way, light mineral oil is distinguished from mineral based on density or viscosity while petroleum distillates require boiling range. These natural mixtures may be chemically, physically, or biologically modified prior to use so information about these modifications is necessary to define them.

In order to assign a substance ID for all organism-based structurally-diverse substances, the following information shall be provided:

* Genus
* Species
* Subspecies, variety or cultivar
* Taxon Author
* Part
* Any chemical process that would modify the covalent structure of any of the constituents in the material (i.e. formaldehyde treatment, proteolytic digestion, etc).

In order to assign a substance ID for other structurally-diverse substances (non organism derived), the following information shall be provided:

* Source Material (e.g, Coal, Petroleum)
* Physical Properties Related to Molecular Structure (i.e. Viscosity, Density, etc.)
* Any Process that Fractionates or Modifies the Source Chemically

The general set of information shall be provided by means of the followings data elements:

### Substance Name

The information related to the name of the substance shall be provided as per specification described in Section 8.2.

### Substance Code

The information related to the code of the substance shall be provided as per specification described in Section 8.3.

### Reference Information

The information related to the reference information of the substance shall be provided as per specification described in Section 8.5.

### Version

The information related to the version of the substance shall be provided as per specification described in Section 8.4.

### Source Material (repeat as necessary)

The Source Material contains all the organism, part and fraction information and is specifed as detailed in Section 8.14.15.

### Comment

|  |  |
| --- | --- |
| **User Guidance** | Any comment can be provided in this field, if necessary. |
| **Example(s)** |  |
| **Conformance** | OPTIONAL – DISCOURAGED |
| **Data Type** | ST |
| **Business Rule(s)** |  |

Note: Comment can always be packaged into a characteristic:

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <partMoiety><!-- Other Details of Structural Unit --></partMoiety>  
 <subjectOf>  
 <characteristic>  
 <text>***Comment*** *Text*</text>  
 </characteristic>  
 </subjectOf>

### Modification (repeat as necessary)

Information related to the modification shall be provided as per specification described in Section 8.11.14.

### Source material (repeat as necessary)

The information related to the source material shall be provided as per specification described in Section 8.13.15.

### Glycosylation (repeat as necessary)

The information related to the glycosylation shall be provided as per specification described in Section 8.11.12.

### Molecular Weight (repeat as necessary)

The information related to the molecular weight shall be provided as per specification described in Section 8.11.13.

#### Amount

The information related to the amount shall be provided as per specification described in Section 8.8.

### Property (repeat as necessary)

The information related to the property shall be provided as per specification described in Section 8.9. For cells or other complex material an additional field is necessary that will contain a substance id related to the qualitative property that describes the material. This field will be used predominantly to describe specific antigens present or absent on cell surfaces. The presence of such antigens is a defining property of the cells. (Example: CD54+ cells would have Substance ID to identify the CD54 antigen as part of the property description.)

#### Amount

The information related to the amount shall be provided as per specification described in Section 8.8.

### Reference Source (repeat as necessary)

The information related to the reference source shall be provided as per specification described in Section 8.2.5.

## Mixture Substance (repeat as necessary)

Material that contains multiple substances can be mixture substance if the substances are either isolated or synthesized together. Racemic mixtures or substances containing unknown or mixed stereochemistry will not be defined as mixtures but will be represented as substances that contain impurities or degradents will not be described as mixtures. Mixtures of mixtures will not be allowed. Each component of a mixture should be listed. Substances present in trace amounts will not be listed in a mixture unless they are known to have a specific effect. Mixtures are also used when substance ambiguity exists in authoritative sources (aloe).

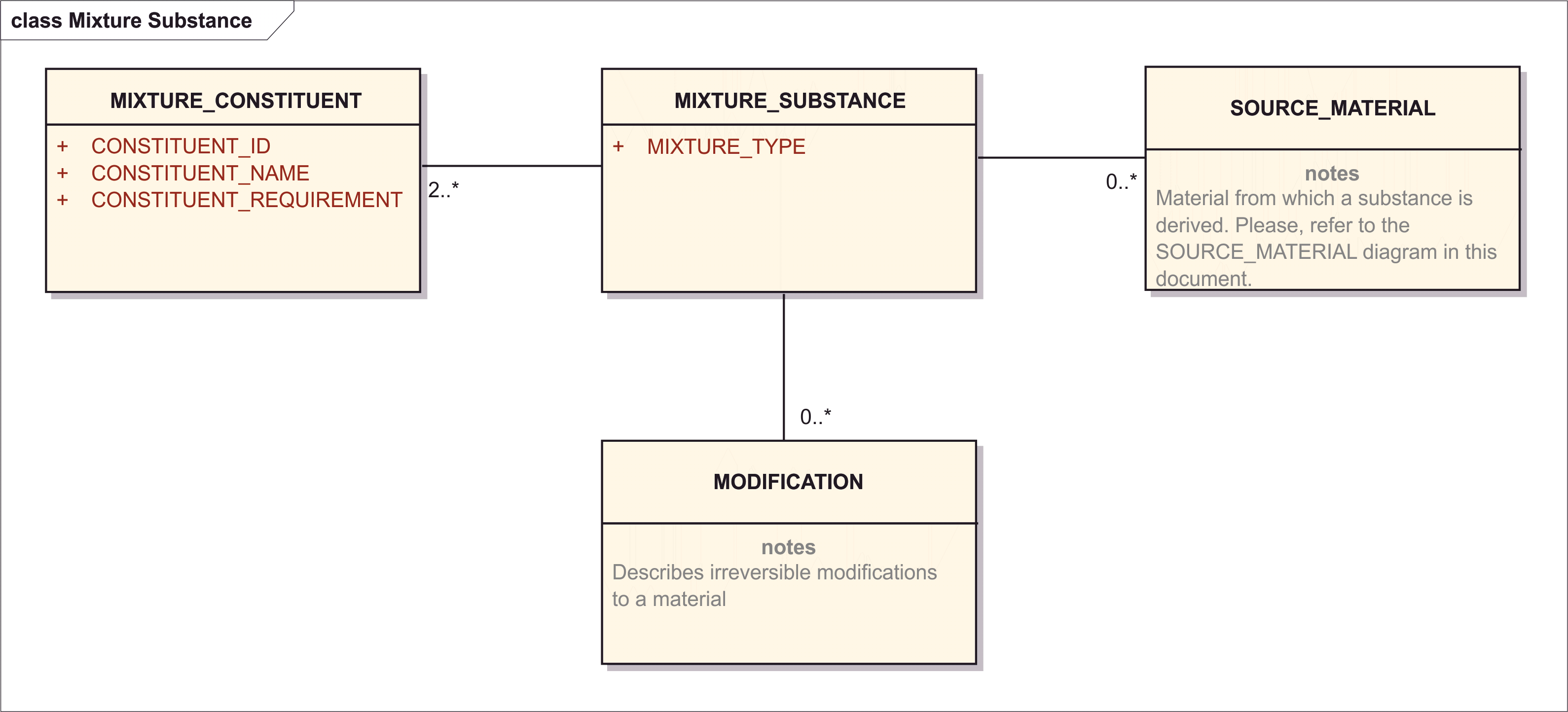


Figure 16– Information Model for Mixture Substance

The following information on mixture shall be provided:

### Mixture Type

|  |  |
| --- | --- |
| **User Guidance** |  |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <asSpecializedKind>  
 <definingMaterialKind>  
 <code code="***Mixture Type*** *Code*" codeSystem="*Code System (OID)*"/>  
 </definingMaterialKind>  
 </asSpecializedKind>

### Mixture Constituent (repeat as necessary)

The mixture constituent set of elements aimed to describe the constituents of the mixture substance. Such information shall be provided by means of the following data elements:

#### Constituent ID

|  |  |
| --- | --- |
| **User Guidance** | The unique identifier assigned to the constituent substance.  NOTE: If a unique “Substance ID” has been assigned, this “Substance ID” shall be specified based on the Substance Name controlled vocabulary. In the absence of a unique “Substance ID” e. g. for the initial submission of the substance this data element is not required. |
| **Example(s)** |  |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Mixture Constituent’*"   
 codeSystem="*Code System (OID)*"  
 displayName="compound"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <code code="***Constituent Substance Id***"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name><!-- Constituent Name --></name>  
 </partMoiety>  
 <!-- any characteristics for structures -->

#### Constituent Name

|  |  |
| --- | --- |
| **User Guidance** | The name of the constituent of the mixture shall be described; established or primary substance name. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Mixture Constituent’*"   
 codeSystem="*Code System (OID)*"  
 displayName="compound"/>  
 <quantity><!-- Moiety Amount --></quantity>  
 <partMoiety>  
 <code code="*Constituent Substance Id*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name><!-- **Constituent Name** --></name>  
 </partMoiety>  
 <!-- any characteristics for structures -->

#### Constituent Requirement

|  |  |
| --- | --- |
| **User Guidance** | A flag indicating if this component is required. |
| **Example(s)** | Always present, may be present. |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <moiety>  
 <code code="*Code for ‘Mixture Constituent’*"   
 codeSystem="*Code System (OID)*"  
 displayName="compound"/>  
 <quantity>  
 <numerator xsi:type="URG\_PQ">  
 <low value="***0 = Not Required, > 0 = Required***" unit="*Amount Unit*"/>  
 <high value="*Amount High Boundary*" unit="*Amount Unit*"/>  
 <numerator>  
 <denominator value="*Reference Amount*" unit="*Reference Unit*"/>  
 </quantity>  
 <partMoiety>  
 <code code="*Constituent Substance Id*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name><!-- Constituent Name --></name>  
 </partMoiety>  
 <!-- any characteristics for structures -->

Constituent Requirement is represented in the Amount numerator low boundary: if the component is not required, the low boundary is 0. If it is required, the low boundary is a value greater than 0.

### Source Material (repeat as necessary)

The information related to the source material shall be provided as per specification described in Section 8.13.15.

### Modification (repeat as necessary)

The information related to the modification shall be provided as per specification described in Section 8.11.14.

### Substance Name

The information related to the name shall be provided as per specification described in Section 8.2.

### Substance Code

The information related to the code shall be provided as per specification described in Section 8.3.

### Reference Information

The information related to the reference information shall be provided as per specification described in Section 8.5.

### Version

The information related to the version shall be provided as per specification described in Section 8.4.

# Specified Substance

There are a number of regulatory needs when defining materials in medicinal products. The globalization of supply chains, widespread contract manufacturing and the high value of many medicinal products places a much greater burden on both regulatory agencies and companies to ensure that the material used in medicinal products is correctly identified with known pedigree.

Although the substance model captures information essential to the description of materials in medicinal products there is often a strong regulatory need for additional information that is not captured at the substance level. The four groups of Specified Substance Elements allow for the explicit capture of information essential for the evaluation and tracking of material used in medicinal products. Each of the four groups of elements provides information essential for these regulatory needs in a manner that should facilitate complaince.

## Group 1 Specified Substance (repeat as necessary)

Multiple substance materials (e.g. simethicone, aluminum lakes), extracts and tinctures (herbal, and polymorphic forms of materials will be described and differentiated using Group 1 specified substance elements.

Group 1 specified substances will always have at least one constituent; all constituents will either be substances or group 1 specified substances.

Single component substances that differ in physical form will be defined as different group 1 specified substances. For example crystalline insulin and amorphous insulin are two different group1 specified substances.

The information model for Group 1 specified substances is listed below.

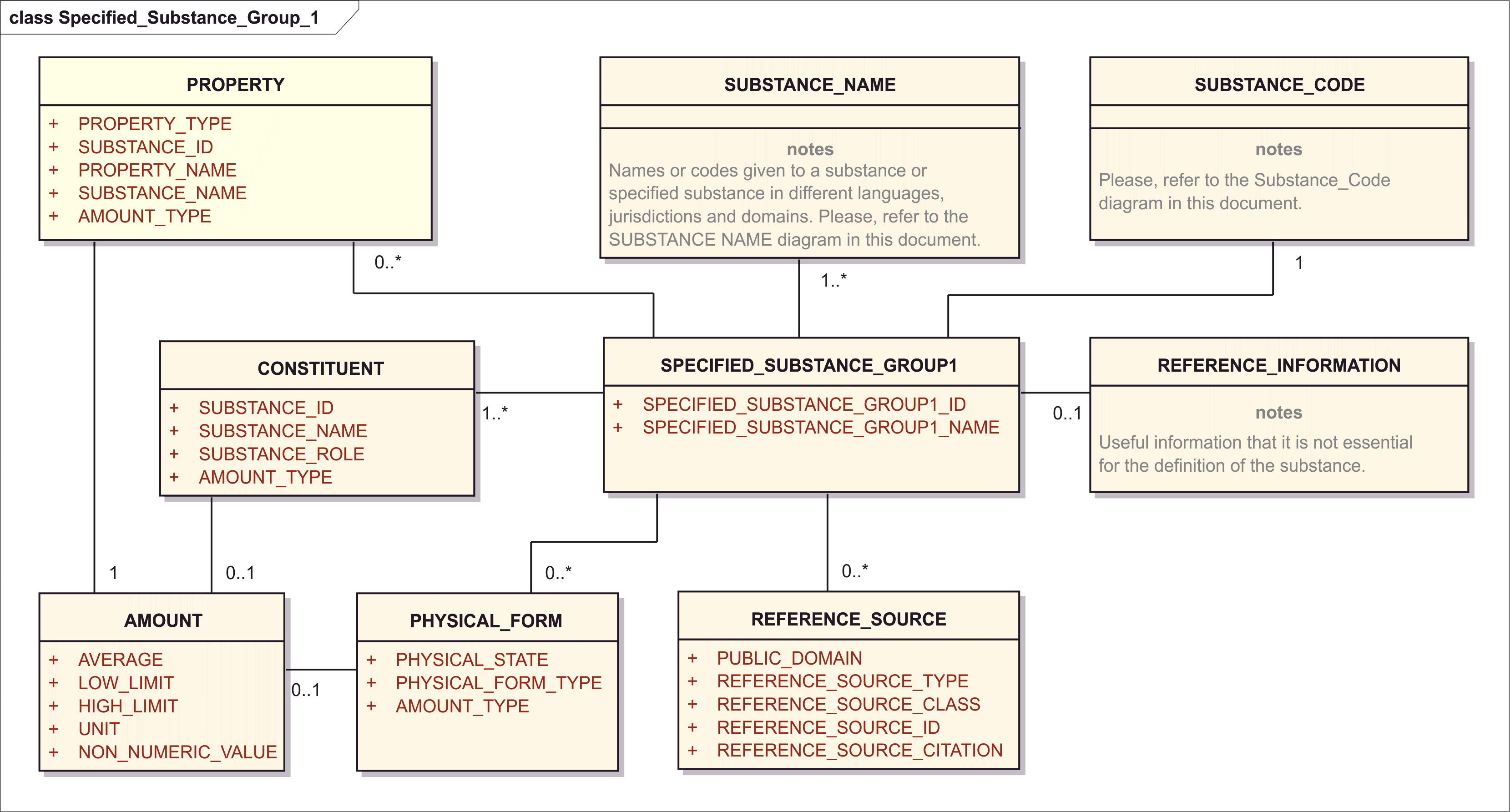


Figure 17– Information model for the Group 1 specified substance

### Specified\_Substance\_Group1\_ID

|  |  |
| --- | --- |
| **User Guidance** | The unique identifier assigned to the specified substance group1 shall be specified.  NOTE: If a unique “Specified Substance ID” has been assigned, this “Specified Substance ID” shall be specified based on the Substance Name controlled vocabulary.  In the absence of a unique “Specified Substance ID” e. g. for the initial submission of the substance this data element is not required. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | II |
| **Business Rule(s)** | The ID of the substance will be automatically assigned by the system once the message will be processed. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="*Substance Specification Group 1 Id*"   
 codeSystem="*Substance Specification Group 1 Code System (OID)*"  
 displayName="***Substance Specification Group 1 Name***"/>

### Specified\_Substance\_Group1\_Name

|  |  |
| --- | --- |
| **User Guidance** | The name of the specified substance group 1 shall be provided in this field. |
| **Example(s)** | The latest version of the Naming Conventions document (Ref. Doc.: EMA/720247/2011). |
| **Conformance** | REQUIRED |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 1 Id***"   
 codeSystem="*Substance Specification Group 1 Code System (OID)*"  
 displayName="*Substance Specification Group 1 Name*"/>

### Substance Name (repeat as necessary)

Names associated with a Group 1 specified substance will be captured in manner similar to Substance Name (see Section 8.2)

### Substance Code

Codes associated with a Group 1 specified substance will be captured in manner similar to Substance Codes (see Section 8.3)

### Version (repeat as necessary)

Version associated with a Group 1 specified substance will be captured in manner similar to Substance Version (see Section 8.4)

### Reference Information

Reference Information will also be captured in a manner similar to that describe in a manner similar to the substance reference information (see Section 8.5).

### Property (repeat as necessary)

Property values will be captured in a manner similar to those described for substances (see Section 8.9). It is possible to capture both definitional properties and properties that provide additional information about a substance. For example the melting point of a polymorphic substance or solubility or rate of dissolution of a crystalline material could be a defining property of a given specified substance group 1 substance.

### Reference Source (repeat as necessary)

Reference source will be captured in a manner similar to the one described in Section 8.2.5).

### Constituent (repeat as necessary)

The constituents group of elements serves several roles, each substance in multiple substance materials will be captured as a component. Signature substances and extraction solvents will also be captured for herbal extracts. The amount of components, constituents and extraction solvents will also be captured. A different Group 1 Specified Substance will be created if these amounts consistently vary.

#### Substance Role

|  |  |
| --- | --- |
| **User Guidance** | The role of the substance in the specified substance shall be described based on a controlled vocabulary. An impurity component is not described in the Specified Substance Group 1. The parent ID of a specified substance group 1 shall always be specified. |
| **Example(s)** | active marker, component, degradent, metabolite, active moiety |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Value Allowed** | COMPONENT, EXTRACTION SOLVENT, SIGNATURE SUBSTANCE; ACTIVE MARKER |
| **Business Rule(s)** |  |

At the level of specified substance, such constituents are all considered “observations” because they would be measured in quality control processes.

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 1 Name****/Code*"   
 codeSystem="***Substance Specification Group 1*** *Code System (OID)*"  
 displayName="***Substance Specification Group 1 Name***"/>  
 <component>  
 <observation>  
 <code code="*Observation Code*"   
 codeSystem="*Code System (OID)*"/>  
 <value ***Amount*** */*>  
 <analyte>  
 <identifiedSubstance>  
 <code code="*Constituent* ***Substance Id*** *(Substance Id)*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>*Constituent* ***Substance Name*** *(Substance Name)*</name>  
 </identifiedSubstance>

Since the terminology for the Constituent Role is not fully defined, and HL7 representation depends on defined business concepts, the possible representations of the example concepts are being discussed in the following:

COMPONENT – When a substance is considered a “component” of a mixture, it should be represented as a mixture component moiety in the substance definition itself. A component at the level of specified substance would be a component recognized in quality assurance measurements, either to ensure it is below a certain upper limit or above a certain lower limit.

SIGNATURE SUBSTANCE – this is a component substance used as a marker for quality or potency, and is controlled at a certain threshold.

ACTIVE MARKER – this too is a component substance used as quality or potency measure, and is controlled at a certain threshold.

EXTRACTION SOLVENT – Specifying a solvent used in a manufacturing process of extraction is done by representing that process step and associating the extraction solvent as an interactor:

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 1 Name****/Code*"   
 codeSystem="***Substance Specification Group 1*** *Code System (OID)*"  
 displayName="***Substance Specification Group 1 Name***"/>  
 <component>  
 <processStep>  
 <code code="*Code meaning ‘Extraction by Solvent’*"   
 codeSystem="*Code System (OID)*"/>  
 <interactor>  
 <identifiedSubstance>  
 <code code="*Constituent* ***Substance Id*** *(Substance Id)*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>*Constituent* ***Substance Name*** *(Substance Name)*</name>  
 </identifiedSubstance>

#### Substance ID

|  |  |
| --- | --- |
| **User Guidance** | The unique identifier assigned to the substance that constitutes the specified substance group 1 shall be defined. For vaccines, the constituents refer to adjuvant (s). For herbals, the constituents refer to herbal markers or the herbal substance (s) which is present in the herbal preparation (e.g. Sennosides A; Senna Acutifolia et Senna Angustifolia leaf tincture). For multiple substances, the constituent refers to the substances complexed together to form the final multiple substance (e.g. Insulin human and protamine are constituents of the Insulin Isophane). |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | Mandatory, at least one constituent is necessary for every group 1 specified substance. A constituent can be a substance or group 1 specified substance. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <analyte>  
 <identifiedSubstance>  
 <code code="*Constituent* ***Substance Id*** *(Substance Id)*"   
 codeSystem="*Substance Id System (OID)*"/>

#### Substance Name

|  |  |
| --- | --- |
| **User Guidance** | The name of the substance which is the constituent of the specified substance group 1 shall be described; Preferred or Official Name of the constituent. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Value Allowed** | NSC, NCI, CAS, ATC, EINICS, INN |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <analyte>  
 <identifiedSubstance>  
 <name>*Constituent* ***Substance Name*** *(Substance Name)*</name>  
 <asNamedEntity .../>

See Section 8.2 for all details of Substance Name.

#### Amount

The information related to the amount shall be provided as per specification described in Section 8.8.

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 1 Name****/Code*"   
 codeSystem="***Substance Specification Group 1*** *Code System (OID)*"  
 displayName="***Substance Specification Group 1 Name***"/>  
 <component>  
 <observation>  
 <code code="*Observation Code*"   
 codeSystem="*Code System (OID)*"/>  
 <value ***Amount*** */*>  
 <analyte>  
 <identifiedSubstance>  
 <code code="*Constituent* ***Substance Id*** *(Substance Id)*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>*Constituent* ***Substance Name*** *(Substance Name)*</name>  
 </identifiedSubstance>

### Physical Form (repeat as necessary)

This section is to capture information on the physical form of the specified substance (e.g. crystalline amorphous, tincture, dry extract), the state of matter, slightly more detailed form of the final specified substance (e.g., solid, liquid, gas, emulsion). Biphasic insulin is one example of a crystal in which part is dissolved and part is in solution. There are also instances of substances being partially crystalline and partially amorphous which would be two different form types and a mixture of Specified Substances, different polymorphic crystalline types.If two polymorphs are distinguished by crystalline type, the symmetry group that distinguish the crystalline types will be captured as a property.

#### Physical State

|  |  |
| --- | --- |
| **User Guidance** | The actual state of the substance in packaged product or administered product |
| **Example(s)** | solid, liquid |
| **Conformance** | IMPLIED |
| **Data Type** | CD |
| **Value Allowed** | SOLID, LIQUID, GAS, EMULSION, GEL |
| **Business Rule(s)** | This is implied by the appropriate terminology for Physical Form Type |

#### Physical Form Type

|  |  |
| --- | --- |
| **User Guidance** | Type of the physical form |
| **Example(s)** | crystaline, amorphous, tincture, dry, extract |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 1 Name****/Code*"   
 codeSystem="***Substance Specification Group 1*** *Code System (OID)*"  
 displayName="***Substance Specification Group 1 Name***"/>  
 <component>  
 <observation>  
 <code code="*Code meaning ‘Physical Form’*"   
 codeSystem="*Code System (OID)*"/>  
 <value xsi:type="CV" codeSystem="Code System (OID)"   
 code="***Physical State/Form Type*** *detail Code*"   
 displayName="*Code Display Name*">  
 <originalText>***Physical Form Type*** *(text)*</originalText>

#### Amount

The information related to the amount shall be provided as per specification described in Section 8.8.

When there is an Amount, it must be specified as an observation with a measurable property code.

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <code code="*Code for Quantitative Measurement*"   
 codeSystem="*Code System (OID)*"/>  
 <value ***Amount*** */*>

## Group 2 Specified Substance (repeat as necessary)

Group 2 Specified Substances will be used to capture the manufacturer of a given substance as well as limited manufacturing information. The defining elements are the parent substance which will be either a substance or a group 1 specified substance. Many biosimilar substances will be distinguished at this level. Each manufacturer should be identifier with a code such as DUNS number and a name. The manufacturing type can be manufacturer or repackager. A repackager will be provided with an ID independent of the manufacturer but should identify the original manufacturer or provide group 2 specified substance id for the manufacturer. A repackager need only provide their own id to each user of their material.

It is also important to capture the production method type (i.e. synthetic, extractive, biosynthetic) production type and the production system that describes the cell line or animal from which a given substance is isolated from.

The manufacturer is not necessarily the entity that has market authorization but the entity that manufacturers or repackages the material. If a substantial change in a critical manufacturing process occurs a new specified substance group 2 id should be generated.

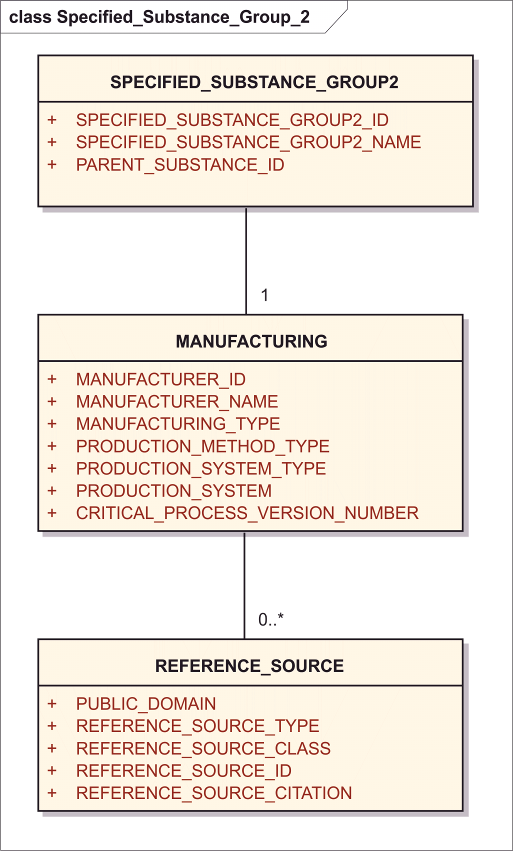


Figure 18– Information model for the Group 2 specified substance

<identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Parent Substance Id***" codeSystem="*Substance Id System (OID)*"  
 <name>***Parent Substance Name***</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 2 Name****/Code*"   
 codeSystem="***Substance Specification* Group 2** *Code System (OID)*"   
 displayName="***Substance Specification Group 2 Name***"/>  
 <author>  
 <assignedEntity>  
 <representedOrganization>  
 <id extension="***Manufacturer Id***"   
 root="*Manufacturer Id System (OID)*"/>  
 <name>***Manufacturer Name***</name>  
 </representedOrganization>  
 <performance>  
 <actDefinition>  
 <code code="***Manufacturer Type*** *Code*"  
 codeSystem="*Manufacturer Type Code System (OID)*"  
 displayName="*Manufacturer Type Name*"/>  
 </actDefinition>  
 </performance>  
 </assignedEntity>  
 </author>  
 <component>  
 <processStep>  
 <code code="***Production Method and System Type*** *Code*"   
 codeSystem="*Code System (OID)*"/>  
 <subjectOf>  
 <document>  
 <id root="***Critical Process*** *Step Document* ***Version*** *Id*"/>  
 <setId root="*Critical Process Step Document Set Id*"/>  
 <versionNumber value="***Critical Process*** *Step* ***Version*** *Number*"/>

### Specified\_Substance\_Group2\_ID

|  |  |
| --- | --- |
| **User Guidance** | The unique identifier assigned to the specified substance group2 shall be specified.  NOTE: If a unique “Specified Substance ID” has been assigned, this “Specified Substance ID” shall be specified based on the Substance Name controlled vocabulary.  In the absence of a unique “Specified Substance ID” e. g. for the initial submission of the substance this data element is not required. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | II |
| **Business Rule(s)** | The ID of the substance will be automatically assigned by the system once the message will be processed. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 2 Id***"   
 codeSystem="*Substance Specification Group 2 Code System (OID)*"  
 displayName="*Substance Specification Group 2 Name*"/>

### Specified\_Substance\_Group2\_Name

|  |  |
| --- | --- |
| **User Guidance** | The name of the specified substance group 2 shall be provided in this field. The name should be coined from the preferred term of the substance and the manufacturer. |
| **Example(s)** | The latest version of the Naming Conventions document (Ref. Doc.: EMA/720247/2011). |
| **Conformance** | REQUIRED |
| **Data Type** | ST |
| **Business Rule(s)** | The name should be coined from a concatenation of the preferred term of the parent substance and the manufacturer (i.e. Human Insulin Lilly). |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="*Substance Specification Group 2 Id*"   
 codeSystem="*Substance Specification Group 2 Code System (OID)*"  
 displayName="***Substance Specification Group 2 Name***"/>

### Parent Substance ID

|  |  |
| --- | --- |
| **User Guidance** | The substance or specified substance group 1 ID that identifies the manufactured substance. The unique identifier assigned to the substance that is the parent of the specified substance shall be provided based on controlled vocabulary. The ID of the parent for the specified substance group 1 may refer to either a substance ID or a specified substance group 1 ID.  NOTE: If a unique “ID” has been assigned, this “ID” shall be specified based on the Substance Name CV. In the absence of a unique “ID” e. g. for the initial submission of specified substance and parent substance, the name of the parent substance shall be specified. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Parent Substance Id***" codeSystem="*Substance Id System (OID)*"  
 <name>***Parent Substance Name***</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>

### Manufacturing

#### Manufacturer ID

|  |  |
| --- | --- |
| **User Guidance** | The unique code used to track manufacturers. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | An ID associated with each manufacturer may or may not be linked to a manufacturer site, |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <author>  
 <assignedEntity>  
 <representedOrganization>  
 <id extension="***Manufacturer Id***"   
 root="*Manufacturer Id System (OID)*"/>

#### Manufacturer Name

|  |  |
| --- | --- |
| **User Guidance** | The name of the manufacturer or repackager of the specified substance. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | ST |
| **Business Rule(s)** | The general name of the manufacturer not linked to a specific site |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <author>  
 <assignedEntity>  
 <representedOrganization>  
 <name>***Manufacturer Name***</name>

#### Manufacturing Type

|  |  |
| --- | --- |
| **User Guidance** | The type of operation performed by the entity associated with the substance. |
| **Example(s)** | manufacturer, repackager, distributor |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | If a repackager the original manufacturer should also be captured |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <author>  
 <assignedEntity>  
 <representedOrganization><!-- Manufacturer --></representedOrganization>  
 <performance>  
 <actDefinition>  
 <code code="***Manufacturer Type*** *Code*"  
 codeSystem="*Manufacturer Type Code System (OID)*"  
 displayName="*Manufacturer Type Name*"/>  
 </actDefinition>

#### Production Method Type

|  |  |
| --- | --- |
| **User Guidance** | The overall type of production system, synthetic, extractive, biosynthetic. |
| **Example(s)** | synthetic, extractive, biosynthetic, semi-synthetic |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | Captured for all |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <code code="***Production Method*** *and System* ***Type*** *Code*"   
 codeSystem="*Code System (OID)*"/>

#### Production System Type

|  |  |
| --- | --- |
| **User Guidance** | Should be captured for material derived from both extractive and biosynthetic production method, for synthesized peptides and nucleic acids. |
| **Example(s)** | Plant, animal, bacterial, fungal, insect cell line, yeast, mammalian cell line, human cell line, animal tissue, human tissue, solid phase chemistry, solution chemistry |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | Should be captured for all material derived from extractive or biosynthetic production methods and some synthetic peptides and nucleic acids. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <code code="***Production*** *Method and* ***System Type*** *Code*"   
 codeSystem="*Code System (OID)*"/>

#### Production System

|  |  |
| --- | --- |
| **User Guidance** | The production system description shall be provided when available. |
| **Example(s)** | cho cell, goat, bovine lungs |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Production System, System Type and Method are all to be represented in one terminology. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <code code="***Production*** *Method and* ***System*** *Type Code*"   
 codeSystem="*Code System (OID)*"/>  
 <subjectOf>  
 <document>  
 <id root="***Critical Process*** *Step Document* ***Version*** *Id*"/>  
 <setId root="*Critical Process Step Document Set Id*"/>  
 <versionNumber value="***Critical Process*** *Step* ***Version*** *Number*"/>

#### Critical Process Version Number

|  |  |
| --- | --- |
| **User Guidance** | Should be captured for material derived from both extractive and biosynthetic production method, for synthesized peptides and nucleic acids. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | INT |
| **Value Allowed** | Start at one and increases if a major critical process changes occur (ie changes in master cell bank; elimination or addition of a chromatographic purification process). |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <subjectOf>  
 <document>  
 <id root="***Critical Process*** *Step Document* ***Version*** *Id*"/>  
 <setId root="*Critical Process Step Document Set Id*"/>  
 <versionNumber value="***Critical Process*** *Step* ***Version*** *Number*"/>

#### Reference Source (repeat as necessary)

Reference source will be captured in a manner similar to the one described in Section 8.2.5).

### Substance Name

A Group 2 Specified Substance name and associated information will be described in a matter similar as that for substances (see Section 8.2).

### Substance Code

Codes associated with a Group 2 specified substance will be captured in manner similar to Substance Codes (see Section 8.3)

### Version

Version associated with a Group 2 specified substance will be captured in manner similar to Substance Version (see Section 8.4)

### Reference Information

Reference Information will also be captured in a manner similar to that describe in a manner similar to the substance reference information (see Section 8.5).

## Group 3 Specified Substance (repeat as necessary)

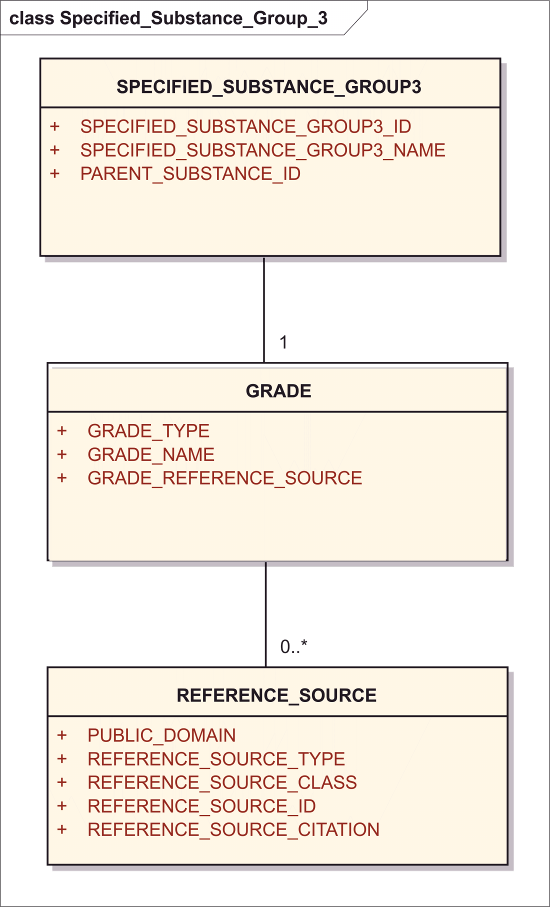


Figure 19– Information model for the Group 3 specified substance

This section captures the reference to a set of specifications typically described in a pharmacopeia indicating the quality of the specified substance.

The information on the grade shall be specified by means of the following data elements:

<identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Parent Substance Id***" codeSystem="*Substance Id System (OID)*"  
 <name>***Parent Substance Name***</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 3 Id****/Code*"   
 codeSystem="***Substance Specification Group 3*** *Code System (OID)*"  
 displayName="***Substance Specification Group 3 Name***"/>  
 <component>  
 <observation>  
 <code code="***Grade*** *Type/****Name*** *Code*" codeSystem="*Code System (OID)*"  
 displayName="***Grade Name***"/>  
 <value ***Amount*** */*>  
 <subjectOf>  
 <document><!-- **Grade Reference Source** --></document>

### Specified\_Substance\_Group3\_ID

|  |  |
| --- | --- |
| **User Guidance** | The unique identifier assigned to the specified substance group3 shall be specified.  NOTE: If a unique “Specified Substance ID” has been assigned, this “Specified Substance ID” shall be specified based on the Substance Name controlled vocabulary.  In the absence of a unique “Specified Substance ID” e. g. for the initial submission of the substance this data element is not required. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | II |
| **Business Rule(s)** | The ID of the substance will be automatically assigned by the system once the message will be processed. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 3 Id****/Code*"   
 codeSystem="*Substance Specification Group 3 Code System (OID)*"  
 displayName="*Substance Specification Group 3 Name*"/>

### Specified\_Substance\_Group3\_Name

|  |  |
| --- | --- |
| **User Guidance** | The name of the specified substance group 3 shall be provided in this field or level 1 Specified Substance Name (Prefered Term) |
| **Example(s)** | Name in monograph appended to Grade Type (i.e. sterile water for injection USP) |
| **Conformance** | REQUIRED |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="*Substance Specification Group 3 Id/Code*"   
 codeSystem="*Substance Specification Group 3 Code System (OID)*"  
 displayName="***Substance Specification Group 3 Name***"/>

### Parent Substance ID

|  |  |
| --- | --- |
| **User Guidance** | The unique identifier assigned to the substance that is the parent of the specified substance shall be provided based on controlled vocabulary. The ID of the parent for the specified substance group 3 may refer to either a substance ID or specified substance group 1 ID.  NOTE:  If a unique “ID” has been assigned, this “ID” shall be specified based on the Substance Name controlled vocabulary. In the absence of a unique “ID” e. g. for the initial submission of specified substance and parent substance, the name of the parent substance shall be specified. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Parent Substance Id***" codeSystem="*Substance Id System (OID)*"  
 <name>***Parent Substance Name***</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>

### Grade

The characteristics of the grade of the substance shall be specified.

#### Grade Type

|  |  |
| --- | --- |
| **User Guidance** | Pharmacopoeial specification type or other specification type. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | Each Pharmacopieal Specification shall be given a separate record. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <code code="***Grade*** *Type/****Name*** *Code*" codeSystem="*Code System (OID)*"  
 displayName="*Grade Name*"/>

#### Grade Name

|  |  |
| --- | --- |
| **User Guidance** | Typically the Monograph Title that refers to a given substance or specified substance; for herbal substances standardized or unstandardized will be appended to the name along with standardization. |
| **Example(s)** | e.g. for Ph. Eur,, JP, USP, for most active ingredients and some excipients. Herbals: quantified, standardized, i.e. St. John’s Wort Standardized (3 mg Hyperforin per gram); Sterile Water USP; purified water, water for injection, Insulin Human Injection |
| **Conformance** | REQUIRED |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <code code="***Grade*** *Type/****Name*** *Code*" codeSystem="*Code System (OID)*"  
 displayName="***Grade Name***"/>

#### Grade Reference Source

|  |  |
| --- | --- |
| **User Guidance** | Refers to the relevant reference source of the grade or specific source of monograph. |
| **Example(s)** | For Insulin zinc injectable suspension (crystalline) monograph. Ph. Eu. Ed 7th 2011 (7.2) |
| **Conformance** | REQUIRED |
| **Data Type** | ST |
| **Value Allowed** | Mandatory every grade must have a source. USP 33. |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <subjectOf>  
 <document><!-- **Grade Reference Source** --></document>

#### Reference Source (repeat as necessary)

Reference source will be captured in a manner similar to the one described in Section 8.2.5).

### Substance Name

A Group 3 Specified Substance name and associated information will be described in a matter similar as that for Substance Name (see Section 8.2).

### Substance Code

Codes associated with a Group 3 specified substance will be captured in manner similar to Substance Code (see Section 8.3)

### Version

Version associated with a Group 3 specified substance will be captured in manner similar to Substance Version (see Section 8.4)

### Reference Information

Reference Information will also be captured in a manner similar to that describe in a manner similar to the substance reference information (see Section 8.5).

## Group 4 Specified Substance (repeat as necessary)

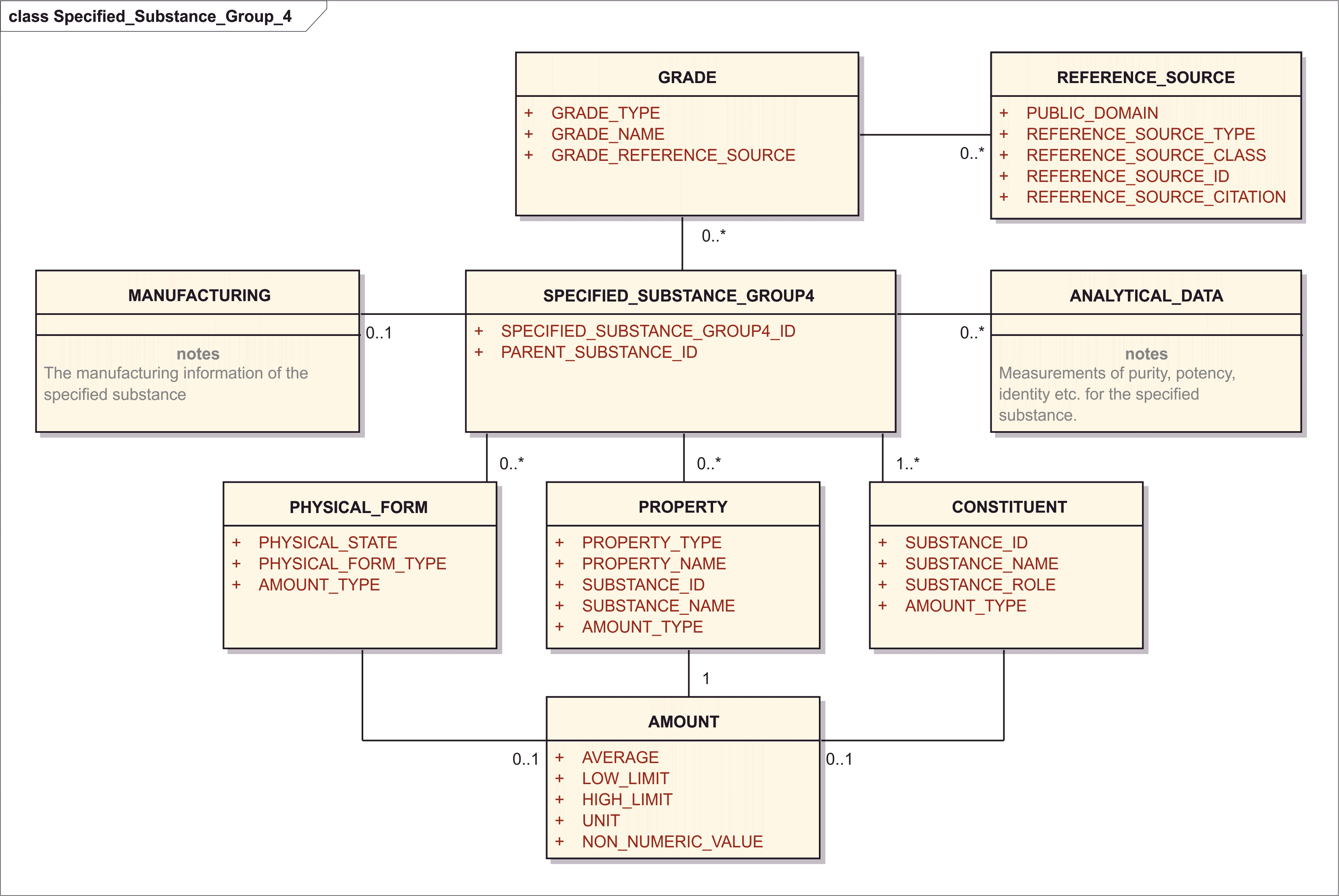


Figure 20– Information model for the Group 4 specified substance

### Specified\_Substance\_Group4\_ID

|  |  |
| --- | --- |
| **User Guidance** | The unique identifier assigned to the specified substance group4 shall be specified.  NOTE: If a unique “Specified Substance ID” has been assigned, this “Specified Substance ID” shall be specified based on the Substance Name controlled vocabulary. In the absence of a unique “Specified Substance ID” e. g. for the initial submission of the substance this data element is not required. |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | II |
| **Business Rule(s)** | The ID of the substance will be automatically assigned by the system once the message will be processed. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 4 Id****/Code*"   
 codeSystem="*Substance Specification* Group 4 *Code System (OID)*"  
 displayName="*Substance Specification Group 4 Name*"/>

### Parent Substance ID

|  |  |
| --- | --- |
| **User Guidance** | The substance or specified substance group 1 ID that identifies the manufactured substance. The unique identifier assigned to the substance that is the parent of the specified substance shall be provided based on controlled vocabulary. The ID of the parent for the specified substance group 1 may refer to either a substance ID or a specified substance group 1 ID.  NOTE: If a unique “ID” has been assigned, this “ID” shall be specified based on the Substance Name CV. In the absence of a unique “ID” e. g. for the initial submission of specified substance and parent substance, the name of the parent substance shall be specified. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Parent Substance Id***" codeSystem="*Substance Id System (OID)*"  
 <name>***Parent Substance Name***</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>

### Grade

Grade shall be captured in a manner similar to the one described in Section 9.3.4. A given manufacturing process may result in material that will meet multiple grade specifications.

#### Reference Source (repeat as necessary)

Reference source shall be captured in a manner similar to the one described in Section 8.2.5).

### Physical Form (repeat as necessary)

Physical Form shall be captured in a manner similar to the one described in Section 9.1.10).

### Constituent (repeat as necessary)

Impurities and degradents in additions to the substances listed in group 1 specified would be captured and listed as such. They may also be capatured as analytical darex. Constituent shall be captured in a manner similar to the one described in Section 9.1.9.

### Analytical Data (repeat as necessary)

The information model allows the capture of specification data in an organized fashion. A given specification can be associated with more than one analytical method. This section will allow the capture of both known and unknown impurities, related substances and limit substances. Known impurities, related and limit substances will be defined as substances and assigned substance ids. This section will allow the capture of substance specification summaries in a standardized and organized manner.

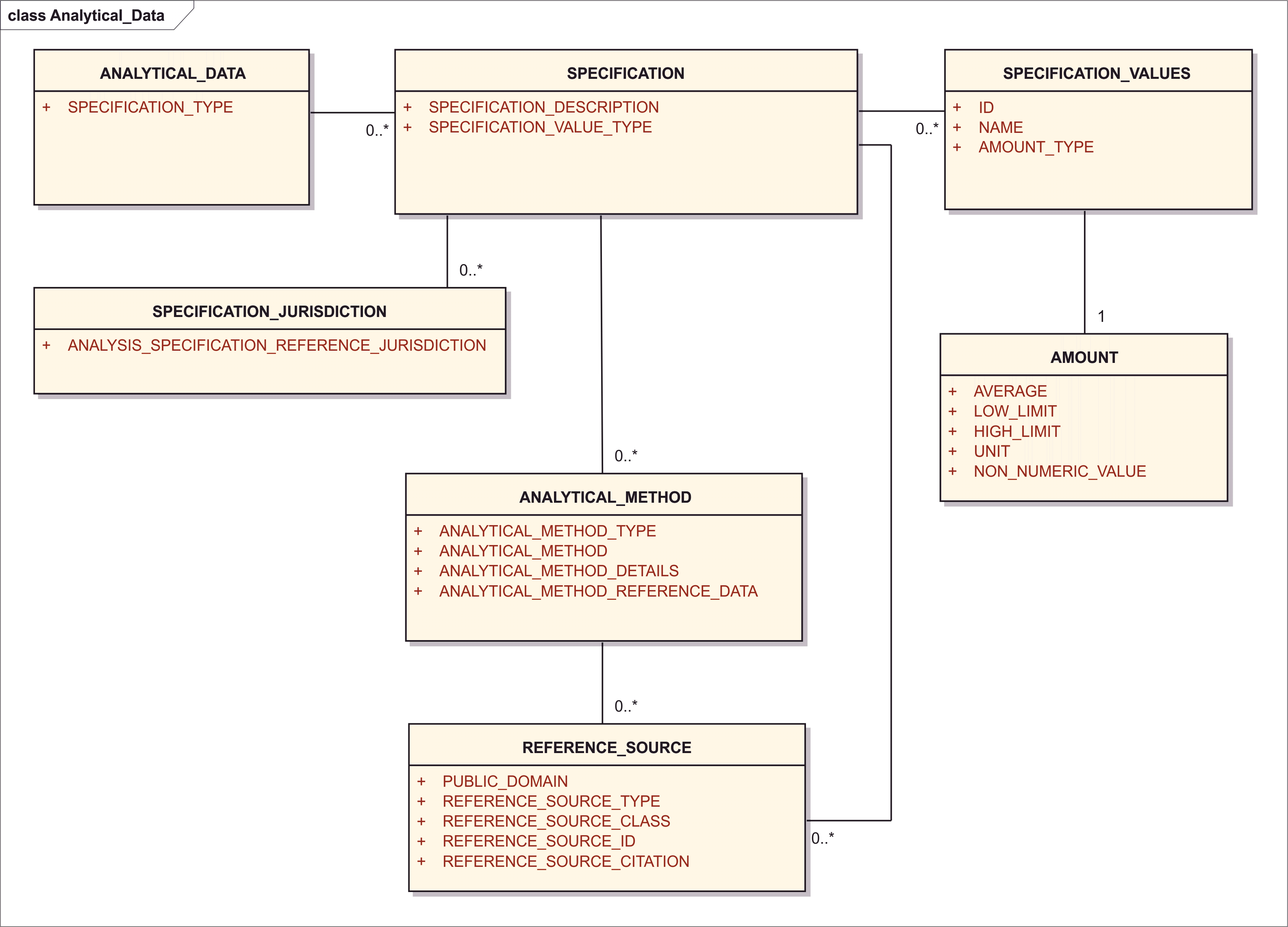


Figure 21– Information model for the Analytical Data

The analytical data type is associated with the specification and not the analytical method; the same analytical method could be used for both purity and potency.

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 4 Id****/Code*"   
 codeSystem="***Substance Specification* Group 4** *Code System (OID)*"  
 displayName="***Substance Specification Group 4 Name***"/>  
 <component>  
 <observation>  
 <code code="***Specification Type, Name, and Value Type*** *Code*"   
 codeSystem="*Code System (OID)*"/>  
 <text><!-- **Specification Description** with **Method Details** --></text>  
 <methodCode code="***Analytical Method*** *Code with* ***Type***"   
 codeSystem="*Code* System (OID)"/>  
 <value ***Amount*** */*>  
 <analyte>  
 <identifiedSubstance>  
 <code code="***Specification Value Id*** *(Substance Id)*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>***Specification Value Name***</name>  
 </identifiedSubstance>  
 </analyte>  
 <subjectOf>  
 <document>  
 <code code="*Code for ‘****Analytical Value Specification****’*"  
 codeSystem="*Code System (OID)*"/>  
 <!-- *Standard Document Attributes for actual specification* -->  
 <author>  
 <assignedEntity>  
 <representedOrganization>  
 <id extension="***Specification Source***"   
 root="*Organization Id Root (OID)*"/>  
 </representedOrganization>  
 </assignedEntity>  
 </author>  
 </document>  
 </subjectOf>  
 <subjectOf>  
 <document>  
 <code code="*Code for ‘****Analytical Method Specification****’*"  
 codeSystem="*Code System (OID)*"/>  
 <!-- *Standard Document Attributes for actual specification* -->  
 <text><!-- **Analytical Method Reference Data** --></text>  
 <author>  
 <assignedEntity>  
 <representedOrganization>  
 <id extension="***Analytical Method Reference Source***"   
 root="*Organization Id Root (OID)*"/>  
 </representedOrganization>  
 </assignedEntity>  
 </author>  
 </document>  
 </subjectOf>

Note that in international standards of metrology, including IUPAC CNPU and LOINC terminologies, any measurable quantity is defined as a dedicated kind of quantity, including the kind of quantity in a general sense (here called “amount type”) and system having the property regarding the component. In IUPAC CNPU, even the unit is included in the concept of the observation code, however, in HL7, the UNIT (if any) should always be placed in the observatiuon/value/@unit code.

#### Specification Type

|  |  |
| --- | --- |
| **User Guidance** | Type of analytical data that is going to be described. |
| **Example(s)** | Identity, purity, potency, limit |
| **Conformance** |  |
| **Data Type** | CD |
| **Business Rule(s)** | One and only one of the above values will be associated with each specification. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <code code="***Specification Type****, Name, and Value Type Code*"   
 codeSystem="*Code System (OID)*"/>

#### Specification (repeat as necessary)

The analytical method information shall be provided by means of the following data elements:

##### Specification Description

|  |  |
| --- | --- |
| **User Guidance** | A brief textual description of the specification |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | ST |
| **Business Rule(s)** | A brief. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <text><!-- **Specification Description** --></text>

##### Specification Value Type

|  |  |
| --- | --- |
| **User Guidance** | The way the value is expressed for a given specification. |
| **Example(s)** | Mole percent, mg per gram, units per gram, area percent, weight percent |
| **Conformance** |  |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <code code="***Specification*** *Type, Name, and* ***Value Type*** *Code*"   
 codeSystem="*Code System (OID)*"/>

##### Reference Source (repeat as necessary)

Reference source shall be captured in a manner similar to the one described in Section 8.2.5).

##### Specification Values (repeat as necessary)

###### ID

|  |  |
| --- | --- |
| **User Guidance** | If the specification value relates to a substance or specified substance group 1 the id should be captured; substance id or id of controlled analytical term. |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <analyte>  
 <identifiedSubstance>  
 <code code="***Specification Value Id*** *(Substance Id)*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>*Specification Value Name*</name>  
 </identifiedSubstance>  
 </analyte>

###### Name

|  |  |
| --- | --- |
| **User Guidance** | Will typically be the name of a substance on which a specification value is set. It could also be a controlled analytical term such total impurities, other impurities, etc. This will be the name of a related substance or impurity for a defined substance. The specification value name could also be free text value for certain specifications. |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <analyte>  
 <identifiedSubstance>  
 <code code="*Specification Value Id (Substance Id)*"   
 codeSystem="*Substance Id System (OID)*"/>  
 <name>***Specification Value Name***</name>  
 </identifiedSubstance>  
 </analyte>

###### Amount

Amount shall be captured in a manner similar to the one described in Section 8.8. To be used to capture the specific amount of a quantitative specification.

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <value ***Amount*** */*>

##### Reference Source (repeat as necessary)

Reference source shall be captured in a manner similar to the one described in Section 8.2.5).

##### Analytical Method (repeat as necessary)

Each specification should be linked to at least one analytical method. More than one analytical method could be used to.

###### Analytical Method Type

|  |  |
| --- | --- |
| **User Guidance** | The general type of analytical method. |
| **Example(s)** | Chromatographic, spectroscopic, bioassay, titration |
| **Conformance** | IMPLICIT |
| **Data Type** | CD |
| **Business Rule(s)** | One terminology to code both method and method type implicit. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <methodCode code="***Analytical Method*** *Code with* ***Type***"   
 codeSystem="*Code* System (OID)"/>

###### Analytical Method

|  |  |
| --- | --- |
| **User Guidance** | Source of the analytical method |
| **Example(s)** | NMR, LC-MS, ELISA |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | One terminology to code both method and method type implicit. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <methodCode code="***Analytical Method*** *Code with Type*"   
 codeSystem="*Code* System (OID)"/>

###### Analytical Method Details

|  |  |
| --- | --- |
| **User Guidance** |  |
| **Example(s)** | HPLC, L1, 4.6X250 MM, ISOCRATIC, 80:20 WATER:ACETONITRILE, 1 ML/MIN; BRIEF FREE TEXT DESCRIPTION |
| **Conformance** |  |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <text><!-- Specification Description with **Method Details** --></text>

###### Analytical Method Reference Data

|  |  |
| --- | --- |
| **User Guidance** | Reference to validation or qualification report |
| **Example(s)** | Validation , Qualification report |
| **Conformance** |  |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <subjectOf>  
 <document>  
 <code code="*Code for ‘Analytical Method Specification’*"  
 codeSystem="*Code System (OID)*"/>  
 <!-- *Standard Document Attributes for actual specification* -->  
 <text><!-- **Analytical Method Reference Data** --></text>

###### Reference Source (repeat as necessary)

Reference source shall be captured in a manner similar to the one described in Section 8.2.5).

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <observation>  
 <subjectOf>  
 <document>  
 <code code="*Code for ‘Analytical Method Specification’*"  
 codeSystem="*Code System (OID)*"/>  
 <!-- *Standard Document Attributes for actual specification* -->  
 <text><!-- **Analytical Method Reference Data** --></text>  
 <author>  
 <assignedEntity>  
 <representedOrganization>  
 <id extension="***Analytical Method Reference Source***"   
 root="*Organization Id Root (OID)*"/>

#### Specification Jurisdiction (repeat as necessary)

##### Analysis Specification Reference Jurisdiction

|  |  |
| --- | --- |
| **User Guidance** |  |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | CD |
| **Business Rule(s)** |  |

### Manufacturing

This section is to describe information on the manufacturing information of the specified substance.

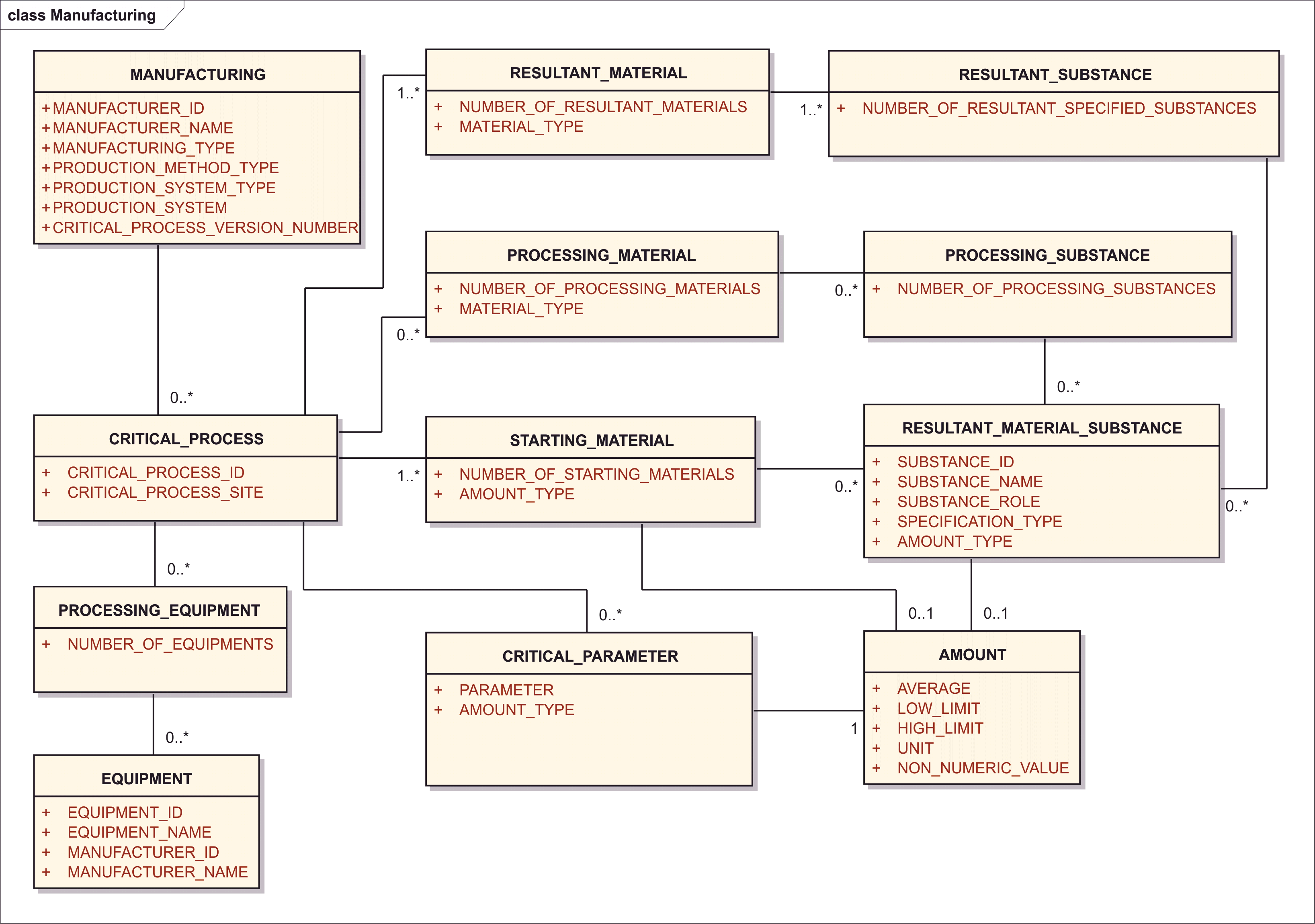


Figure 22 – Information model for the Manufacturing Data

The following information shall be provided where available

<identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Parent Substance Id***" codeSystem="*Substance Id System (OID)*"  
 <name>***Parent Substance Name***</name>  
 </identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Group 4 Name****/Code*"   
 codeSystem="***Substance Specification* Group 4** *Code System (OID)*"  
 displayName="***Substance Specification Group 4 Name***"/>  
 <author>  
 <assignedEntity>  
 <representedOrganization>  
 <id extension="***Manufacturer Id***"   
 root="*Manufacturer Id System (OID)*"/>  
 <name>***Manufacturer Name***</name>  
 </representedOrganization>  
 <performance>  
 <actDefinition>  
 <code code="***Manufacturer Type*** *Code*"  
 codeSystem="*Manufacturer Type Code System (OID)*"  
 displayName="*Manufacturer Type Name*"/>  
 </actDefinition>  
 </performance>  
 </assignedEntity>  
 </author>  
 <component>  
 <processStep>  
 <id root="***Critical Process Id*** *(UUID)*"/>  
 <code code="***Production Method and System Type*** *Code*"   
 codeSystem="*Code System (OID)*"/>  
 <device>  
 <instanceOfKind>  
 <instanceOfProductKind>  
 <code code="***Equipment ID*** *(Product Item**Code)*"   
 codeSystem="*Product Item Code System (OID)*"/>  
 <name>***Equipment Name****</*name>  
 <asManufacturedProduct>  
 <assignedEntity>  
 <representedOrganization>  
 <id extension="***Manufacturer ID***"   
 root="*Organization ID System (OID)*"/>  
 <name><!-- **Manufacturer Name** --></name>  
 </representedOrganization>  
 </assignedEntity>  
 </asManufacturedProduct>  
 </instanceOfProductKind>  
 </instanceOfKind>  
 </device>  
 <interactor typeCode="CSM">  
 <!-- repeat per each **Starting Material** or **Processing Material** -->  
 <functionCode code="***Substance Role / Material Type*** *Code*"  
 codeSystem="*Code System (OID)*" displayName="*Display**Name*"/>  
 <quantity ***Amount*** />  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Substance ID***"   
 codeSystem="*Substance ID System (OID)*"/>  
 <name>***Substance Name***</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification ID***"   
 codeSystem="*Substance Specification ID System (OID)*"/>  
 </substanceSpecification>  
 </subjectOf>  
 </identifiedSubstance>  
 </interactor>  
 <interactor typeCode="PRD">  
 <!-- repeat per each **Resulting Material** -->  
 <functionCode code="***Substance Role / Material Type*** *Code*"  
 codeSystem="*Code System (OID)*" displayName="*Display**Name*"/>  
 <quantity ***Amount*** />  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Substance ID***"   
 codeSystem="*Substance ID System (OID)*"/>  
 <name>***Substance Name***</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification ID***"   
 codeSystem="*Substance Specification ID System (OID)*"/>  
 </substanceSpecification>  
 </subjectOf>  
 </identifiedSubstance>  
 </interactor>  
 <primaryPerformer>  
 <assignedEntity>  
 <id extension="***Critical Process Site*** *ID*"   
 root="*Site ID System (OID)*"/>  
 </assignedEntity>  
 </primaryPerformer>  
 <controlVariable>  
 <characteristic>  
 <code code="***Parameter*** *Code*"   
 displayName="*Parameter Name*"   
 codeSytem="*Code System (OID)*"/>  
 <value ***Amount*** />  
 </characteristic>  
 </controlVariable>  
 <subjectOf>  
 <document>  
 <id root="***Critical Process*** *Step Document* ***Version*** *Id*"/>  
 <setId root="*Critical Process Step Document Set Id*"/>  
 <versionNumber value="***Critical Process*** *Step* ***Version*** *Number*"/>

NOTE: Resultant Material is represented by interactor element with typeCode = “PRD”, i.e., the product of the reaction. Both Starting Material and Processing Material is represented by the interactor elenment with typeCode = “CSM” for consumable, because both enter into the processing step or reaction. The material Substance Role terminology placed into interactor/functionCode distinguishes nuances such as “starting material” from “processing material”, There is one theoretical exception: to specify a processing material which acts as a catalyst the interactor/@typeCode = “CAT” for catalyst might be used; however, such physical catalysts as platinum sheets, would more likely be specified as devices (equipment) or their parts.

#### Manufacturer ID

|  |  |
| --- | --- |
| **User Guidance** | The unique code used to track manufacturers. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | An ID associated with each manufacturer may or may not be linked to a manufacturer site, |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <author>  
 <assignedEntity>  
 <representedOrganization>  
 <id extension="***Manufacturer Id***"   
 root="*Manufacturer Id System (OID)*"/>

#### Manufacturer Name

|  |  |
| --- | --- |
| **User Guidance** | The name of the manufacturer or repackager of the specified substance. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | ST |
| **Business Rule(s)** | The general name of the manufacturer not linked to a specific site |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <author>  
 <assignedEntity>  
 <representedOrganization>  
 <name>***Manufacturer Name***</name>

#### Manufacturing Type

|  |  |
| --- | --- |
| **User Guidance** | The type of operation performed by the entity associated with the substance. |
| **Example(s)** | manufacturer, repackager, distributor |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | If a repackager the original manufacturer should also be captured |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <author>  
 <assignedEntity>  
 <representedOrganization><!-- Manufacturer --></representedOrganization>  
 <performance>  
 <actDefinition>  
 <code code="***Manufacturer Type*** *Code*"  
 codeSystem="*Manufacturer Type Code System (OID)*"  
 displayName="*Manufacturer Type Name*"/>  
 </actDefinition>

#### Production Method Type

|  |  |
| --- | --- |
| **User Guidance** | The overall type of production system, synthetic, extractive, biosynthetic. |
| **Example(s)** | synthetic, extractive, biosynthetic, semi-synthetic |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | Captured for all |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <code code="***Production Method*** *and System* ***Type*** *Code*"   
 codeSystem="*Code System (OID)*"/>

#### Production System Type

|  |  |
| --- | --- |
| **User Guidance** | Should be captured for material derived from both extractive and biosynthetic production method, for synthesized peptides and nucleic acids. |
| **Example(s)** | Plant, animal, bacterial, fungal, insect cell line, yeast, mammalian cell line, human cell line, animal tissue, human tissue, solid phase chemistry, solution chemistry |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** | Should be captured for all material derived from extractive or biosynthetic production methods and some synthetic peptides and nucleic acids. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <code code="***Production*** *Method and* ***System Type*** *Code*"   
 codeSystem="*Code System (OID)*"/>

#### Production System

|  |  |
| --- | --- |
| **User Guidance** | The production system description shall be provided when available. |
| **Example(s)** | cho cell, goat, bovine lungs |
| **Conformance** | OPTIONAL |
| **Data Type** | CD |
| **Business Rule(s)** | Production System, System Type and Method are all to be represented in one terminology. |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <code code="***Production*** *Method and* ***System*** *Type Code*"   
 codeSystem="*Code System (OID)*"/>  
 <subjectOf>  
 <document>  
 <id root="***Critical Process*** *Step Document* ***Version*** *Id*"/>  
 <setId root="*Critical Process Step Document Set Id*"/>  
 <versionNumber value="***Critical Process*** *Step* ***Version*** *Number*"/>

#### Critical Process Version Number

|  |  |
| --- | --- |
| **User Guidance** | Should be captured for material derived from both extractive and biosynthetic production method, for synthesized peptides and nucleic acids. |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | INT |
| **Value Allowed** | Start at one and increases if a major critical process changes occur (ie changes in master cell bank; elimination or addition of a chromatographic purification process). |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Details --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <subjectOf>  
 <document>  
 <id root="***Critical Process*** *Step Document* ***Version*** *Id*"/>  
 <setId root="*Critical Process Step Document Set Id*"/>  
 <versionNumber value="***Critical Process*** *Step* ***Version*** *Number*"/>

#### Critical Process (repeat as necessary)

The critical processes steps shall be described by means of the following set of data elements:

##### Critical Process ID

|  |  |
| --- | --- |
| **User Guidance** | code used to track critical process steps |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | II |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <id root="***Critical Process Id*** *(UUID)*"/>

##### Critical Process Site

|  |  |
| --- | --- |
| **User Guidance** | code used by regulators to track manufacturers (e.g. duns number) |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <primaryPerformer>  
 <assignedEntity>  
 <id extension="***Critical Process Site*** *Id*"   
 root="*Site ID System (OID)*"/>

##### Starting Material (repeat as necessary)

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <interactor typeCode="CSM">  
 <functionCode code="***Substance Role*** */ Material Type**Code*"  
 codeSystem="*Code System (OID)*" displayName="*Display**Name*"/>  
 <quantity ***Amount*** />  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Substance ID***"   
 codeSystem="*Substance ID System (OID)*"/>  
 <name>***Substance Name***</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification ID***"   
 codeSystem="*Substance Specification ID System (OID)*"/>  
 </substanceSpecification>  
 </subjectOf>  
 </identifiedSubstance>  
 </interactor>

###### Number of Starting material

|  |  |
| --- | --- |
| **User Guidance** | Number of starting materials used by the manufacturer in the critical process step. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** |  |
| **Business Rule(s)** | This number can be counted and is never explicitly represented. |

###### Amount

The information related to the amount shall be provided as per specification described in section 8.8.

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <interactor typeCode="CSM">  
 <quantity ***Amount*** />

###### Starting Material Substance (repeat as necessary)

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <interactor typeCode="CSM">  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Substance ID***"   
 codeSystem="*Substance ID System (OID)*"/>  
 <name>***Substance Name***</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification ID***"   
 codeSystem="*Substance Specification ID System (OID)*"/>  
 </substanceSpecification>  
 </subjectOf>  
 </identifiedSubstance>  
 </interactor>

Substance ID

|  |  |
| --- | --- |
| **User Guidance** | Substance or specified substance unique identifier |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <interactor typeCode="CSM">  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Substance Id***"   
 codeSystem="*Substance ID System (OID)*"/>

Substance Name

|  |  |
| --- | --- |
| **User Guidance** | Established or primary name substance or specified substance |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <interactor typeCode="CSM">  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <name>***Substance Name***</name>

Substance Role

|  |  |
| --- | --- |
| **User Guidance** | Purpose of the substance or specified substance in the critical process step |
| **Example(s)** |  |
| **Conformance** | REQUIRED |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <interactor typeCode="CSM">  
 <functionCode code="***Substance Role*** */ Material Type Code*"  
 codeSystem="*Code System (OID)*" displayName="*Display**Name*"/>

Specification Type

|  |  |
| --- | --- |
| **User Guidance** | Specification of starting material used (grade) |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | CD |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <interactor typeCode="CSM">  
 <identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification Id***"   
 codeSystem="*Substance Specification Id System (OID)*"/>

##### Processing Material (repeat as necessary)

Information on the material used to produce the substance or specified substance shall be provided by means of the following data elements:

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <interactor typeCode="CSM">  
 <!-- repeat per each **Starting Material** or **Processing Material** -->  
 <functionCode code="***Substance Role / Material Type*** *Code*"  
 codeSystem="*Code System (OID)*" displayName="*Display Name*"/>  
 <quantity ***Amount*** />  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Substance ID***"   
 codeSystem="*Substance ID System (OID)*"/>  
 <name>***Substance Name***</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification ID***"   
 codeSystem="*Substance Specification ID System (OID)*"/>  
 </substanceSpecification>  
 </subjectOf>  
 </identifiedSubstance>  
 </interactor>

###### Number of Processing Materials

|  |  |
| --- | --- |
| **User Guidance** | Number of materials in direct contact with starting material but not present in the product |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | This number can be counted and is never explicitly represented. |

###### Material Type

|  |  |
| --- | --- |
| **User Guidance** | General role of the material in the process |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | CD |
| **Value Allowed** | Extraction solvent, solvent, (..) |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <interactor typeCode="CSM">  
 <!-- repeat per each **Starting Material** or **Processing Material** -->  
 <functionCode code="*Substance Role /* ***Material Type*** *Code*"  
 codeSystem="*Code System (OID)*" displayName="*Display**Name*"/>

###### Processing Substance (repeat as necessary)

Number of Processing Substances

|  |  |
| --- | --- |
| **User Guidance** | Number of substance or specified substances used as processing materials in the critical process step |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | This number can be counted and is never explicitly represented. |

Processing Material Substance (repeat as necessary)

See Section 9.4.7.8.3.3

##### Resultant Material (repeat as necessary)

Information on the material resulting from the process shall be provided by means of the following data elements:

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <interactor typeCode="PRD">  
 <!-- repeat per each **Resulting Material** -->  
 <quantity ***Amount*** />  
 <functionCode code="***Substance Role / Material Type*** *Code*"  
 codeSystem="*Code System (OID)*" displayName="*Display Name*"/>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Substance ID***"   
 codeSystem="*Substance ID System (OID)*"/>  
 <name>***Substance Name***</name>  
 </identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <code code="***Substance Specification ID***"   
 codeSystem="*Substance Specification ID System (OID)*"/>  
 </substanceSpecification>  
 </subjectOf>  
 </identifiedSubstance>  
 </interactor>

###### Number of Resultant Materials

|  |  |
| --- | --- |
| **User Guidance** | Number of materials produced in the critical process step. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | This number can be counted and is never explicitly represented. |

###### Material Type

|  |  |
| --- | --- |
| **User Guidance** | Type of the material resulting from the process |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | CD |
| **Value Allowed** | Api, intermediate product, formulation, etc. |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <interactor typeCode="PRD">  
 <!-- repeat per each **Resulting Material** -->  
 <quantity ***Amount*** />  
 <functionCode code="*Substance Role /* ***Material Type*** *Code*"  
 codeSystem="*Code System (OID)*" displayName="*Display Name*"/>

###### Resultant Substance (repeat as necessary)

Number of Resultant Specified Substance

|  |  |
| --- | --- |
| **User Guidance** | Number of substances or specified substances in the resultant material |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | This number can be counted and is never explicitly represented. |

Resultant Material Substance (repeat as necessary)

See Section 9.4.7.8.3.3

##### Processing Equipment (repeat as necessary)

Information related to the equipment used for the production of the substance or specified substance shall be provided by means of the following data elements:

###### Number of Equipments

|  |  |
| --- | --- |
| **User Guidance** | Number of critical piexes of equipment used in the Critical Process step. |
| **Example(s)** |  |
| **Conformance** | IMPLICIT |
| **Data Type** | INT |
| **Business Rule(s)** | This number can be counted and is never explicitly represented. |

###### Equipment (repeat as necessary)

Detailed description of the equipment shall be provided by means of the following data elements:

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <device>  
 <instanceOfKind>  
 <instanceOfProductKind>  
 <code code="***Equipment ID*** *(Product Item**Code)*"   
 codeSystem="*Product Item Code System (OID)*"/>  
 <name>***Equipment Name****</*name>  
 <asManufacturedProduct>  
 <assignedEntity>  
 <representedOrganization>  
 <id extension="***Manufacturer ID***"   
 root="*Organization ID System (OID)*"/>  
 <name><!-- **Manufacturer Name** --></name>  
 </representedOrganization>  
 </assignedEntity>  
 </asManufacturedProduct>  
 </instanceOfProductKind>  
 </instanceOfKind>  
 </device>

Equipment ID

|  |  |
| --- | --- |
| **User Guidance** | Model or product number from manufacturer |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | II |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <device>  
 <instanceOfKind>  
 <instanceOfProductKind>  
 <code code="***Equipment ID*** *(Product Item**Code)*"   
 codeSystem="*Product Item Code System (OID)*"/>

Equipment Name

|  |  |
| --- | --- |
| **User Guidance** | General name for equipment |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <device>  
 <instanceOfKind>  
 <instanceOfProductKind>  
 <name>***Equipment Name****</*name>

Manufacturer ID

|  |  |
| --- | --- |
| **User Guidance** | Code used by regulators to track manufacturers (e.g. duns number) |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | II |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <device>  
 <instanceOfKind>  
 <instanceOfProductKind>  
 <asManufacturedProduct>  
 <assignedEntity>  
 <representedOrganization>  
 <id extension="***Manufacturer ID***"   
 root="*Organization ID System (OID)*"/>

Manufacturer Name

|  |  |
| --- | --- |
| **User Guidance** | Process equipment manufacturer. |
| **Example(s)** |  |
| **Conformance** |  |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <device>  
 <instanceOfKind>  
 <instanceOfProductKind>  
 <asManufacturedProduct>  
 <assignedEntity>  
 <representedOrganization>  
 <name><!-- **Manufacturer Name** --></name>

##### Critical Parameter (repeat as necessary)

Information on the parameters used in the critical processes shall be provided by means of the following data elements:

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <controlVariable>  
 <characteristic>  
 <code code="***Parameter*** *Code*"   
 displayName="*Parameter Name*"   
 codeSytem="*Code System (OID)*"/>  
 <value ***Amount*** />  
 </characteristic>  
 </controlVariable>

###### Parameter

|  |  |
| --- | --- |
| **User Guidance** |  |
| **Example(s)** |  |
| **Conformance** | CONDITIONAL |
| **Data Type** | ST |
| **Business Rule(s)** |  |

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <controlVariable>  
 <characteristic>  
 <code code="***Parameter*** *Code*"   
 displayName="*Parameter Name*"   
 codeSytem="*Code System (OID)*"/>

###### Amount

The information related to the amount shall be provided as per specification described in Section 8.9.

<identifiedSubstance>  
 <identifiedSubstance><!-- Substance Detail --></identifiedSubstance>  
 <subjectOf>  
 <substanceSpecification>  
 <component>  
 <processStep>  
 <controlVariable>  
 <characteristic>  
 <value ***Amount*** />

### Constituent

This section is to be provide the description of the substance(s) or specified substance(s) that are mixed together to form another specified substance. Most of the specified substance shall be associated with a particular substance. It can also be used to capture putative active or signature substances, impurities or degradants and even known metabolites if so desired.

The mixing together of substances is a process step and so these substances are nothing else than Starting Material of process steps described in Section 9.4.7.8.3.

The use for active or signature substance and impurities or degradents, which are effectively quality control measurement specifications, are as defined in Section 9.1.9.

### Property (repeat as necessary)

The information related to the property shall be provided as per specification described in Section 8.9.

### Physical Form

The information related to the physical form shall be provided as per specification described in Section 9.1.10.

### Substance Name

A Group 4 Specified Substance name and associated information will be described in a matter similar as that for substance name (Section 8.2).

### Substance Code

Codes associated with a Group 4 specified substance will be captured in manner similar to Substance Code (see Section 8.3)

### Version

Version associated with a Group 4 specified substance will be captured in manner similar to Substance Version (see Section 8.4)

### Reference Information

Reference Information will also be captured in a manner similar to that describe in a manner similar to the substance reference information (see Section 8.5).

# Annex

## Example AMOXICILLIN

USAN defines amoxicillin as a trihydrate, INN as an anhydrous substance.

Graphical Structural Representation



<document>  
 <component>  
 <structuredBody>  
 <component>  
 <section>  
 <subject>

<identifiedSubstance>  
 <id extension="**804826J2HU**" root="**2.16.840.1.113883.4.9**"/>  
 <identifiedSubstance>  
 <code code="**804826J2HU**" codeSystem="**2.16.840.1.113883.4.9**"/>

<name>**amoxicillin**</name>  
 <asNamedEntity>  
 <code code="*Substance Name Type (Code)*"  
 codeSystem="*Substance Name Type Code System (OID)*"  
 displayName="**OFFICIAL NAME**"/>  
 <name xml:lang="**en**">**amoxicillin**</name>  
 <statusCode code="**active**"/>  
 <assigningTerritory>  
 <code code="**USA**" codeSystem="2.16.840.1.113883.5.28"/>  
 <assigningTerritory>  
 <subjectOf>  
 <document>  
 <id extension="**USAN**"   
 root="*Reference Source Type Document Id System (OID)*"/>  
 <bibliographicDesignationText>**USP DICTIONARY 2011**</bibliographicDesignationText>  
 </document>  
 </subjectOf>  
 <subjectOf>  
 <policy>  
 <code code="***Official Name Domain*** *(Code)*" codeSystem="*Code System (OID)*"   
 displayName="*Display Name*"/>  
 </policy>  
 </subjectOf>  
 </asNamedEntity>

<asNamedEntity>  
 <code code="*Substance Name Type (Code)*"  
 codeSystem="*Substance Name Type Code System (OID)*"  
 displayName="**OFFICIAL NAME**"/>  
 <name xml:lang="**en**">**amoxicillin trihydrate**</name>  
 <statusCode code="**active**"/>  
 <assigningTerritory>  
 <code code="**EUR**" codeSystem="2.16.840.1.113883.5.28"/>  
 <assigningTerritory>  
 <subjectOf>  
 <document>  
 <id extension="**EP**"   
 root="*Reference Source Type Document Id System (OID)*"/>  
 <bibliographicDesignationText>**European Pharmacopoiea   
 2011**</bibliographicDesignationText>  
 </document>  
 </subjectOf>  
 <subjectOf>  
 <policy>  
 <code code="*Official Name Domain (Code)*" codeSystem="*Code System (OID)*"   
 displayName="**DRUG**"/>  
 </policy>  
 </subjectOf>  
 </asNamedEntity>

<asNamedEntity>  
 <code code="*Substance Name Type (Code)*"  
 codeSystem="*Substance Name Type Code System (OID)*"  
 displayName="**OFFICIAL NAME**"/>  
 <name xml:lang="**en**">**amoxicillin trihydrate**</name>  
 <statusCode code="**active**"/>  
 <assigningTerritory>  
 <code code="**JPN**" codeSystem="2.16.840.1.113883.5.28"/>  
 <assigningTerritory>  
 <subjectOf>  
 <document>  
 <id extension="**JAN**"   
 root="*Reference Source Type Document Id System (OID)*"/>  
 <bibliographicDesignationText>**Japanese Approved Names   
 2011**</bibliographicDesignationText>  
 </document>  
 </subjectOf>  
 <subjectOf>  
 <policy>  
 <code code="*Official Name Domain (Code)*" codeSystem="*Code System (OID)*"   
 displayName="**DRUG**"/>  
 </policy>  
 </subjectOf>  
 </asNamedEntity>

<asEquivalentSubstance>  
 <statusCode code="**active**"/>  
 <definingMaterialKind>  
 <code code="**61336-70-7**" codeSystem="*Code System OID for CAS*"/>  
 </definingMaterialKind>  
 <subjectOf>  
 <document>  
 <bibliographicDesignationText>***USP Dictionary 2011***</bibliographicDesignationText>  
 </document>  
 </subjectOf>  
 </asEquivalentSubstance>

<asEquivalentSubstance>  
 <statusCode code="**active**"/>  
 <definingMaterialKind>  
 <code code="**J01CA04**" codeSystem="*Code System OID for WHO ATC*"/>  
 </definingMaterialKind>  
 <subjectOf>  
 <document>  
 <bibliographicDesignationText>***WHO ATC***</bibliographicDesignationText>  
 </document>  
 </subjectOf>  
 </asEquivalentSubstance>

<asEquivalentSubstance>  
 <statusCode code="**active**"/>  
 <definingMaterialKind>  
 <code code="**QG51AX01**" codeSystem="*Code System OID for WHO ATC*"/>  
 </definingMaterialKind>  
 <subjectOf>  
 <document>  
 <bibliographicDesignationText>***WHO ATC***</bibliographicDesignationText>  
 </document>  
 </subjectOf>  
 </asEquivalentSubstance>

<asSpecializedKind>  
 <definingMaterialKind>  
 <code code="**N0000175497**"   
 codeSystem="**2.16.840.1.113883.3.26.1.5**"  
 displayName="**Penicillin-class Antibacterial [EPC]**"/>  
 </definingMaterialKind>  
 </asSpecializedKind>

<moiety><!-- Structural Unit -->  
 <partMoiety>  
 <code code="**804826J2HU**" codeSystem="**2.16.840.1.113883.4.9**"/>  
 <moiety typeCode="ACTM"><!-- Substance Relationship -->  
 <partMoiety>  
 <code code="**9EM05410Q9**" codeSystem="**2.16.840.1.113883.4.9**"/>  
 <name>**amoxicillin anhydrous**</name>  
 </partMoiety>  
 <subjectOf>  
 <document>  
 <bibliographicDesignationText>***USP Dictionary   
 2011***</bibliographicDesignationText>  
 </document>  
 </subjectOf>  
 </moiety>  
 </partMoiety>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Complete Structural Representation’*"   
 codeSystem="*Code System (OID)*">  
 <value xsi:type="ED" mediaType="application/x-inchi">**1S/C16H19N3O5S.3H2O/c1-16(2)11(15(23)24)19-13(22)10(14(19)25-16)18-12(21)9(17)7-3-5-8(20)6-4-7;;;/h3-6,9-11,14,20H,17H2,1-2H3,(H,18,21)(H,23,24);3\*1H2/t9-,10-,11+,14-;;;/m1.../s1**</value>  
 </characteristic>  
 </subjectOf>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code for ‘Stereochemistry’*" codeSystem="*Code System (OID)*">  
 <value xsi:type="CV" code="*Code for ‘chiral’*" codeSystem="*Code System (OID)*"  
 displayName="chiral"/>  
 </characteristic>  
 </subjectOf>  
 </moiety>  
 </identifiedSubstance>  
 <interactsIn>  
 <interaction>  
 <code code="*Interaction Type Code* for ‘inhibition"   
 codeSystem="*Code System (OID)*"/>  
 <interactor>  
 <functionCode code="***Target Class*** *Code*"  
 codeSystem="*Target Class Code System (OID)*"   
 displayName="*Target Class Name*"/>  
 <identifiedSubstance>  
 <identifiedSubstance>  
 <code code="***Code for the class of ‘bacterial penicillin binding protein’***"   
 codeSystem="*Target Id System (OID)*"/>  
 <name>**bacterial penicillin binding protein**</name>  
 </identifiedSubstance>  
 </identifiedSubject>  
 </interactor>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Target Type’*" codeSystem="*Code Sytem (OID)*"/>  
 <value code="*Target Type Code for ‘bacterial cell wall synthesis’*" xsi:type="CE"  
 codeSystem="*Code Sytem (OID)*"  
 displayName="**bacterial cell wall synthesis**"/>  
 </value>  
 </characteristic>  
 </subjectOf>  
 <subjectOf>  
 <document>  
 <code code="34391-3" codeSystem="2.16.840.1.113883.6.1"   
 displayName="Indications Human Prescription Drug Label"/>  
 <text>  
 <reference value="**http://www.accessdata.fda.gov/spl/data/c4f92484-5755-4046-93dd-772db30acebb/c4f92484-5755-4046-93dd-772db30acebb.xml**"/>  
 </text>  
 </document>  
 </subjectOf>  
 </interaction>  
 </interactsIn>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Optical Activity*" codeSystem="*Code System (OID)*"  
 displayName="optical activity"/>  
 <value code="*Optical Activity Code for ‘+’*" xsi:type="CV"  
 codeSystem="*Optical Activity Code System (OID)*"  
 displayName="**+**"/>  
 </characteristic>  
 </subjectOf>  
 <subjectOf>  
 <characteristic>  
 <code code="*Code meaning ‘Substance is Stoichiometric’*"   
 codeSystem="*Code System (OID)*">  
 <value xsi:type="BL" value="**true**"/>  
 </characteristic>  
 </subjectOf>  
</identifiedSubstance>