Rhea RDF documentation

version: November 2020

Preliminary remarks

SPARQL endpoints:

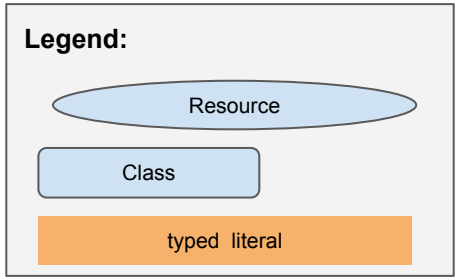
- The Rhea SPARQL endpoint (https://sparql.rhea-db.org/sparql) is built from both **rhea.rdf** and **chebi.owl** (ftp://ftp.expasy.org/databases/rhea/rdf/). The datasets are synchronized with UniProt RDF data releases.
- **rhea.rdf** does not contain the cross-references to UniProt protein entries. They are available through the UniProt SPARQL endpoint (https://sparql.uniprot.org/sparql).

In Rhea RDF (rhea.rdf), all data are represented as subclasses of rdfs:Class, *i.e* there are no instances.

Table of content

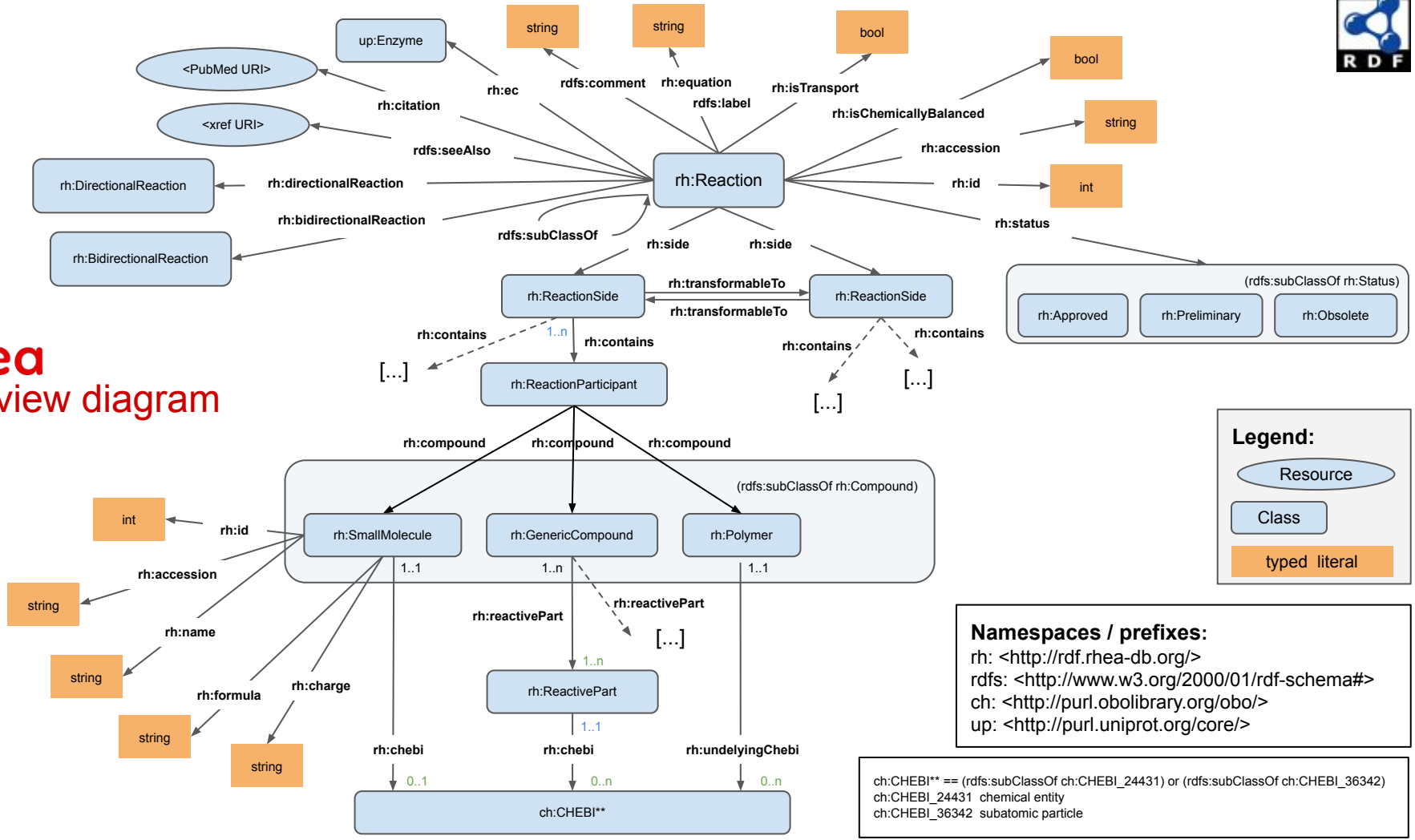
Conventions

Namespaces / prefixes:
rh: <http://rdf.rhea-db.org/>
rdfs: <http://www.w3.org/2000/01/rdf-schema#>
ch: <http://purl.obolibrary.org/obo/>
ch2: <http://purl.obolibrary.org/obo/chebi#>
ch3: <http://purl.obolibrary.org/obo/chebi/>
up: <http://purl.uniprot.org/core/>
ec: <http://purl.uniprot.org/enzyme/>
pubmed: <http://rdf.ncbi.nlm.nih.gov/pubmed/>



ch:CHEBI** == (rdfs:subClassOf ch:CHEBI_24431) or (rdfs:subClassOf ch:CHEBI_36342)
ch:CHEBI_24431 chemical entity
ch:CHEBI_36342 subatomic particle

Rhea
Overview diagram



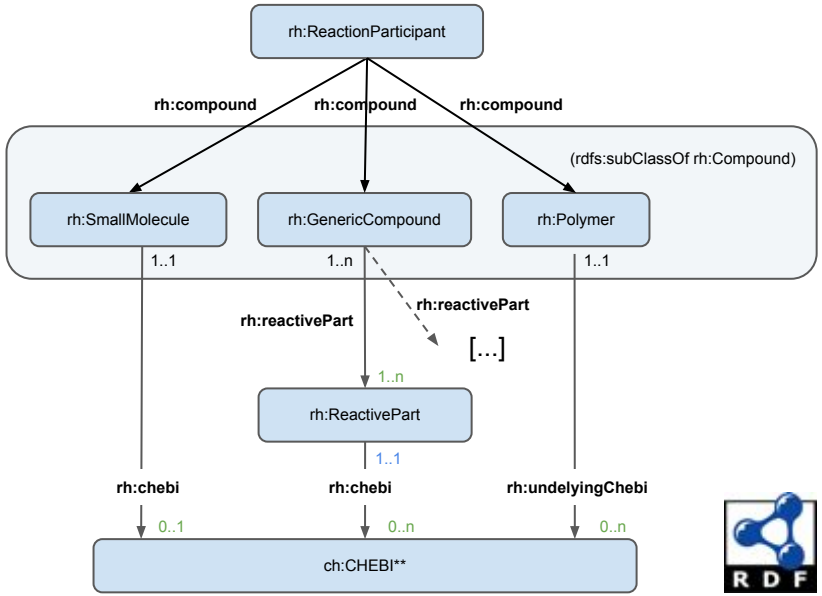
Legend:

- Resource
- Class
- typed literal

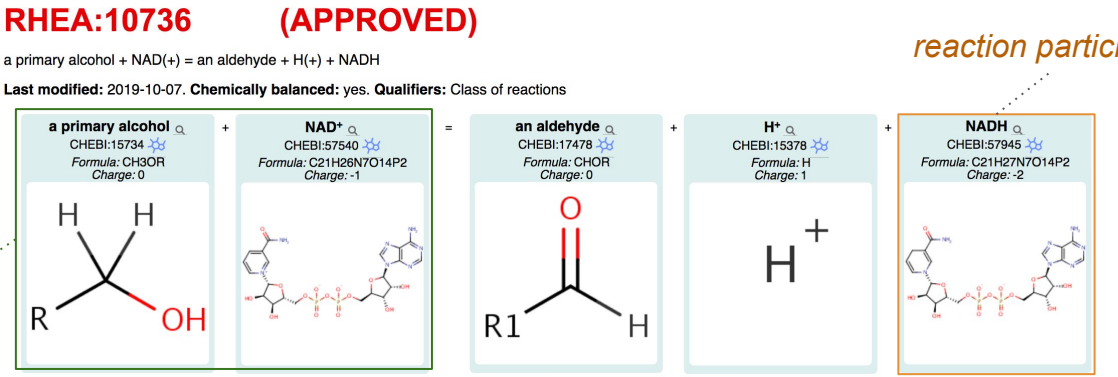
Namespaces / prefixes:

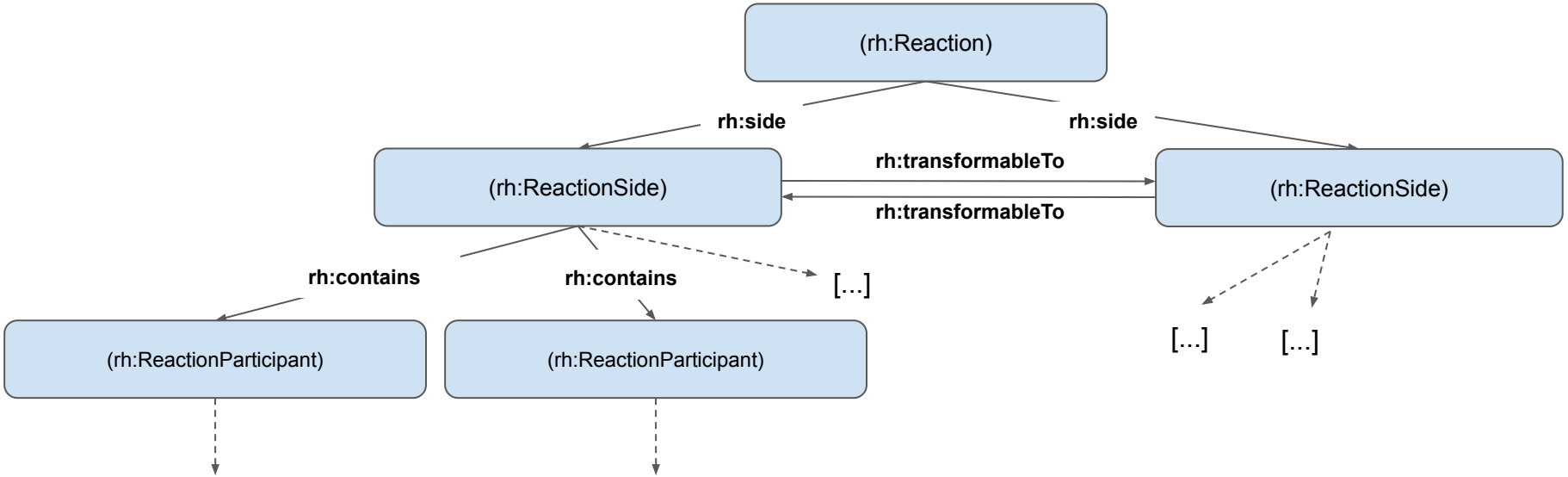
rh: <http://rdf.rhea-db.org/>
rdfs: <http://www.w3.org/2000/01/rdf-schema#>
ch: <http://purl.obolibrary.org/obo/>
up: <http://purl.uniprot.org/core/>

ch:CHEBI** == (rdfs:subClassOf ch:CHEBI_24431) or (rdfs:subClassOf ch:CHEBI_36342)
ch:CHEBI_24431 chemical entity
ch:CHEBI_36342 subatomic particle

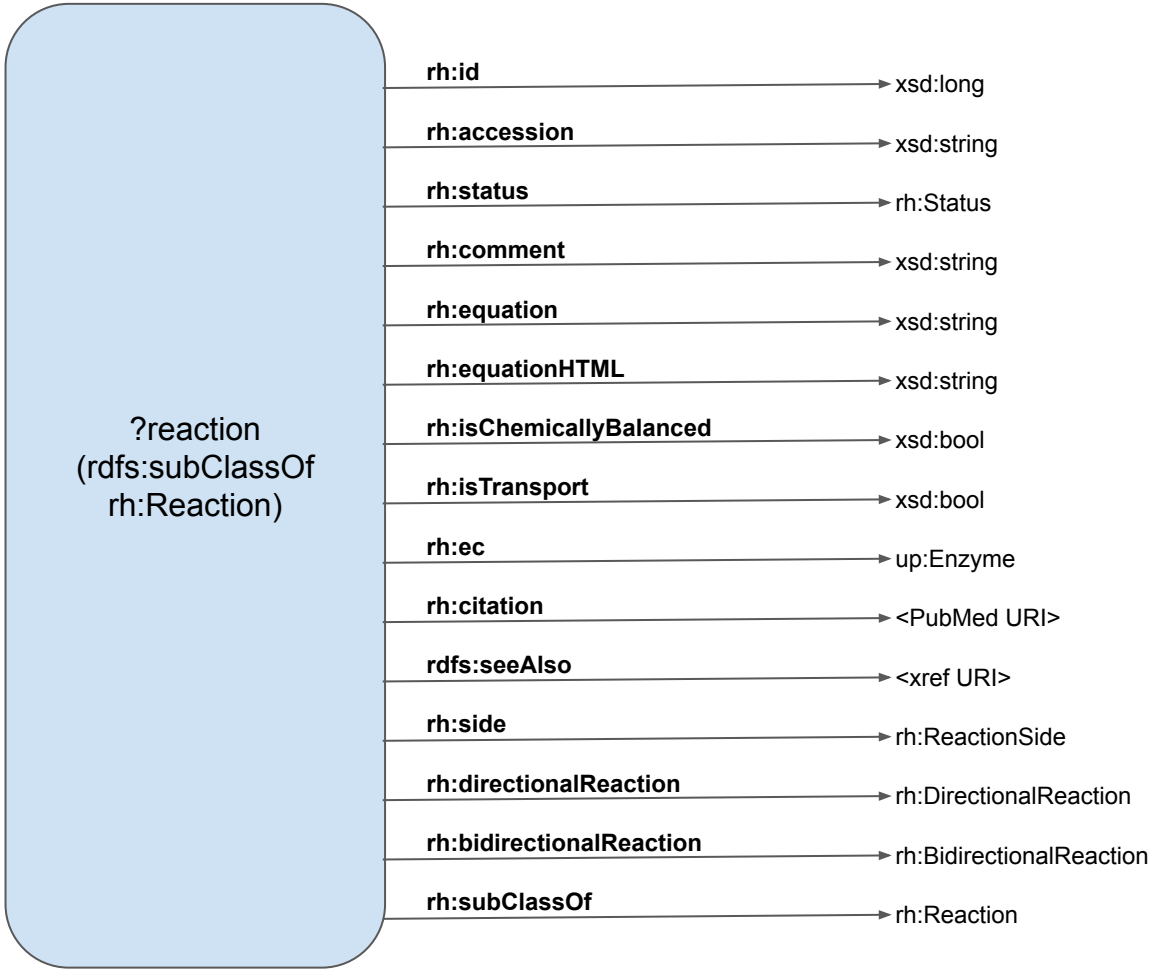

rh:ReactionParticipant
rh:compound
rh:compound
rh:compound
(rdfs:subClassOf rh:Compound)
rh:SmallMolecule
rh:GenericCompound
rh:Polymer
1..1
1..n
1..1
rh:reactivePart
rh:reactivePart
[…]
1..n
rh:ReactivePart
1..1
rh:chebi
rh:chebi
rh:undelyingChebi
0..1
0..n
0..n
ch:CHEBI**
R D F

Reactions, sides and participants



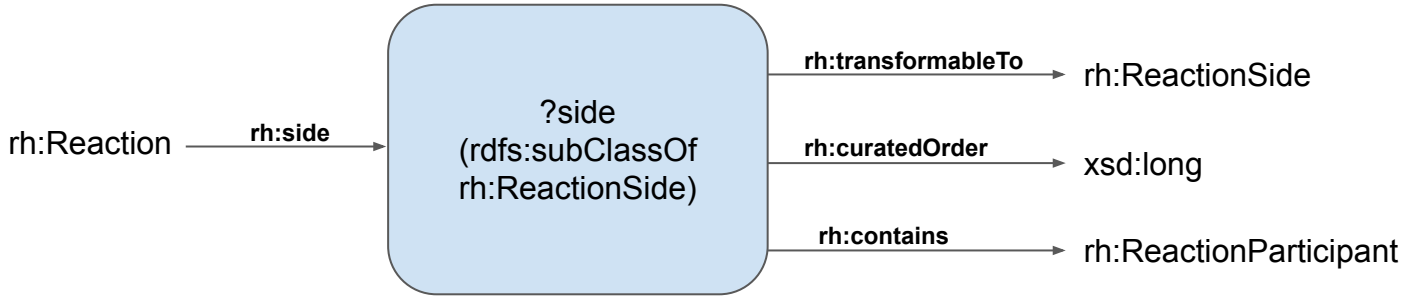


rh:Reaction



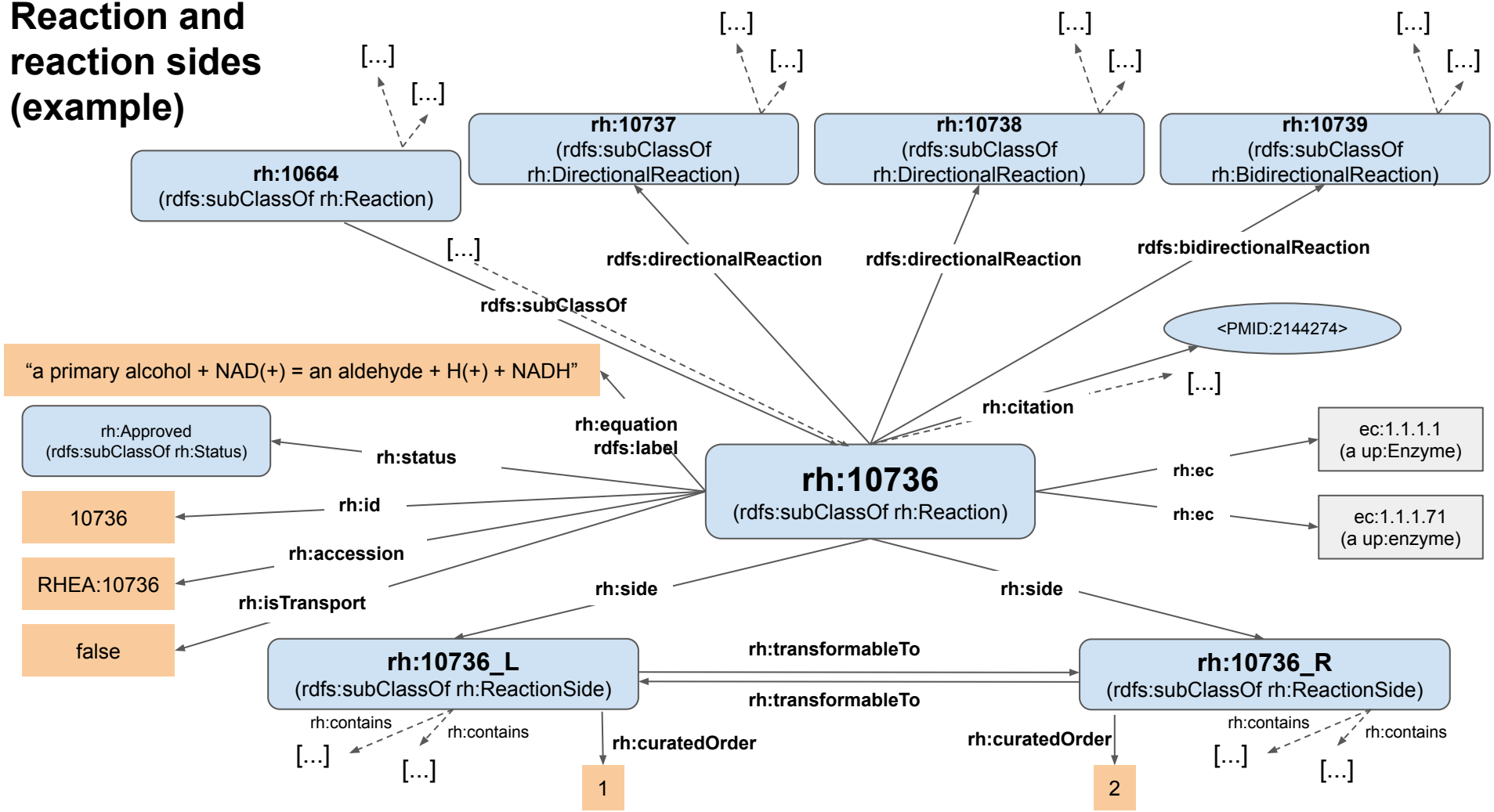
Definition:
A chemical or transport reaction of biological interest with unspecified direction. It is composed of two reaction sides.
rh:Reaction is a top class.
***All the reactions published by Rhea are rdfs:subClassOf of rh:Reaction**.*

rh:ReactionSide

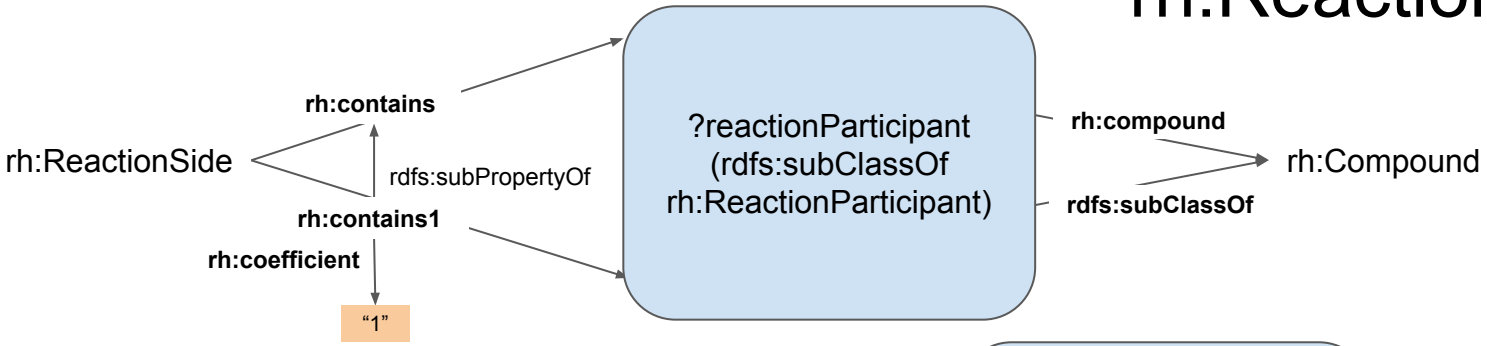


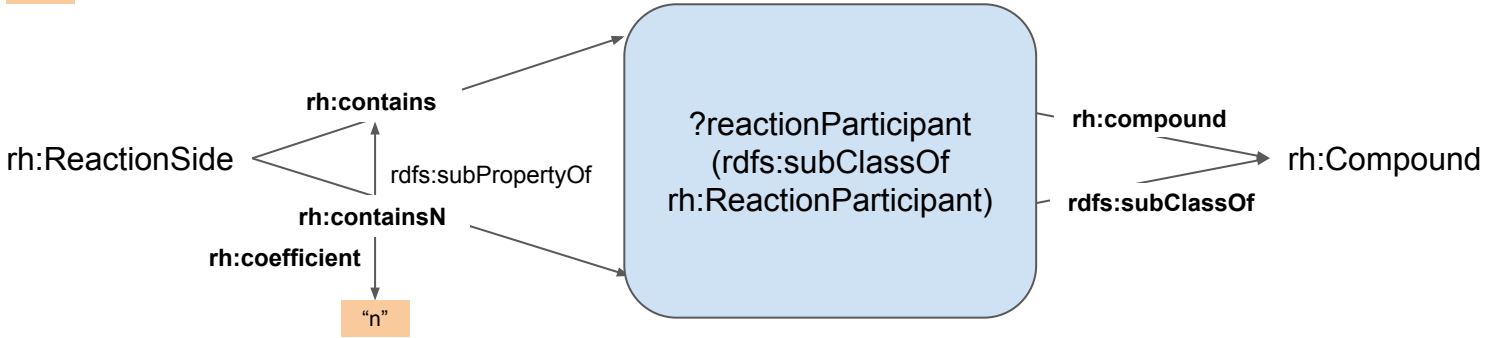
Definition:
A chemical or transport reaction (rh:Reaction) is composed of two reaction sides (rh:ReactionSide), sometimes denoted as left and right reaction sides.
*In a **rh:Reaction**, the rh:ReactionSide are arbitrarily defined. They are linked by the predicate **rh:side**.*
*In a **rh:DirectionalReaction**, one rh:ReactionSide represents the substrate(s) (reactant(s), input) of the reaction, whereas the other represents the product(s) (output). In that case, the predicates **rh:substrates** and **rh:products** are used to denote the links.*
*In a **rh:BidirectionalReaction**, depending on the context, one rh:ReactionSide represents either the substrate(s) or the product(s). In that case, the predicate **rh:substratesOrProducts** is used to denote the links.*
The link between the two reaction sides of a given reaction is denoted by the rh:transformableTo predicate.
A rh:ReactionSide gathers all the reaction participants (rh:ReactionParticipant) that are transformed into another rh:ReactionSide.

Reaction and reaction sides (example)



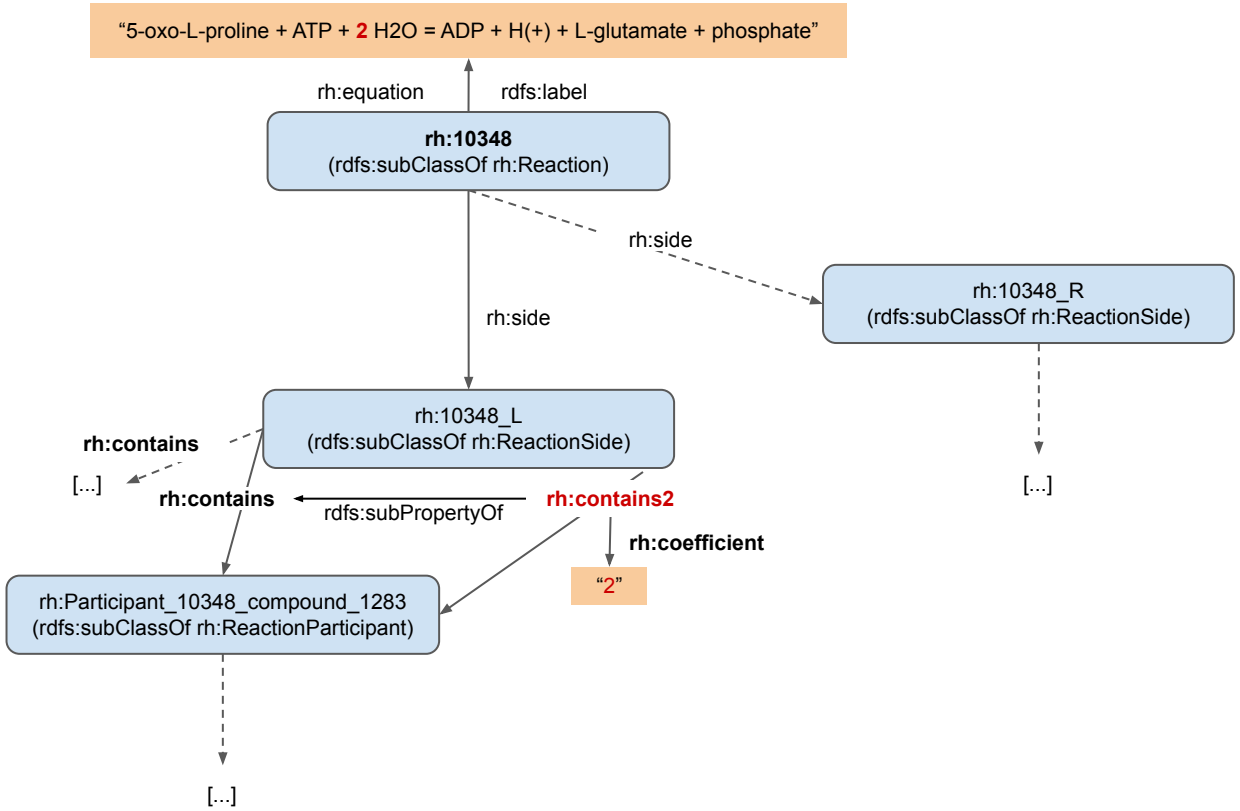
rh:ReactionParticipant

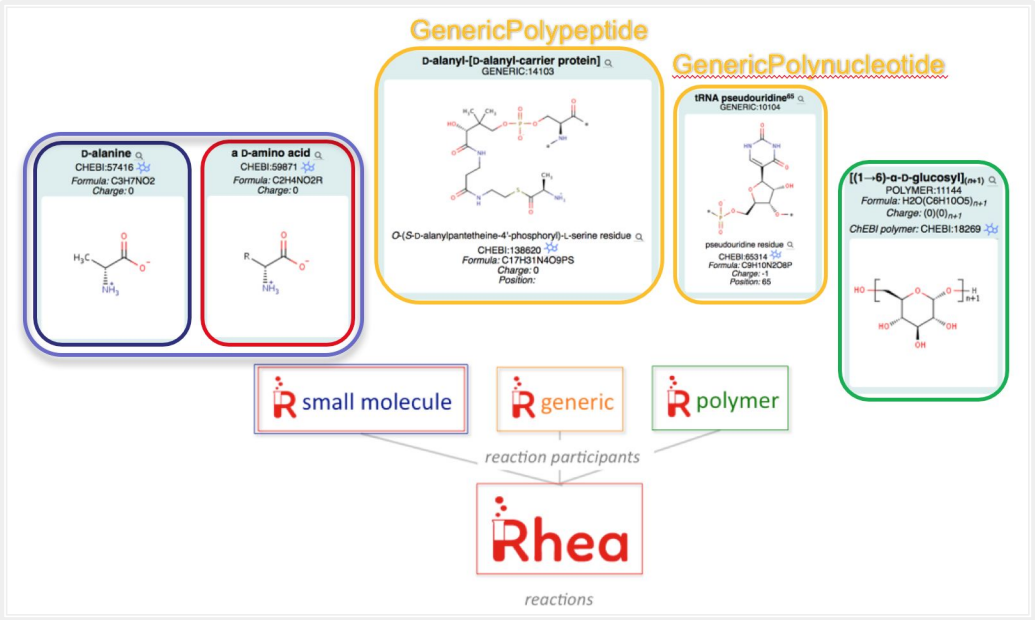


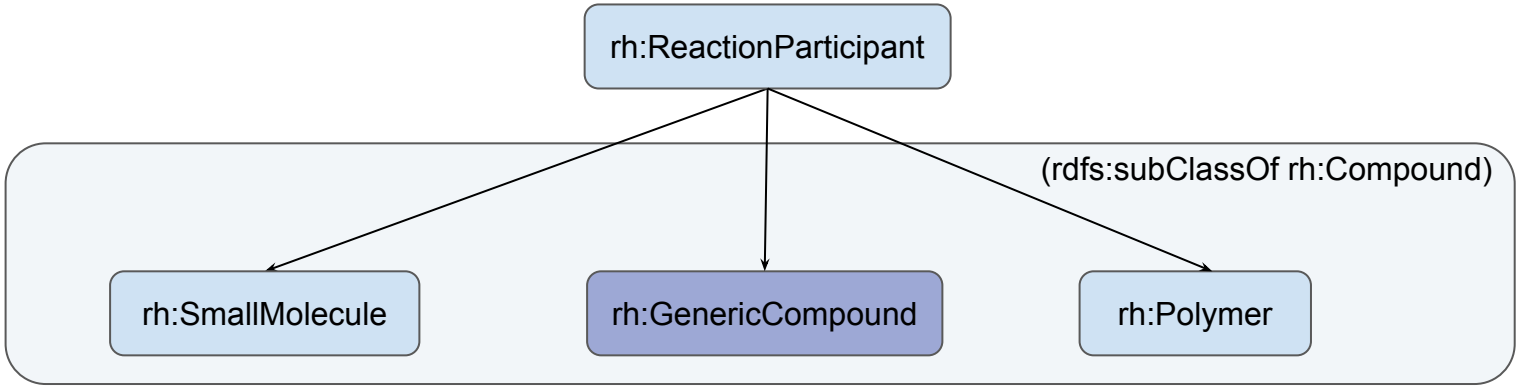


Definition: A chemical entity (subclass of rh:Compound) that is transformed in a chemical reaction or translocated in a transport reaction. There are three kind of participants: small molecules (rh:SmallMolecule), macromolecules (rh:GenericCompound) and polymers (rh:Polymer).
A reaction side (rh:ReactionSide) contains (rh:contains) participants (rh:ReactionParticipant).
The stoichiometric coefficient of a participant is defined in a subproperty of rh:contains (rh:contains1, rh:contains2, rh:containsN,...).

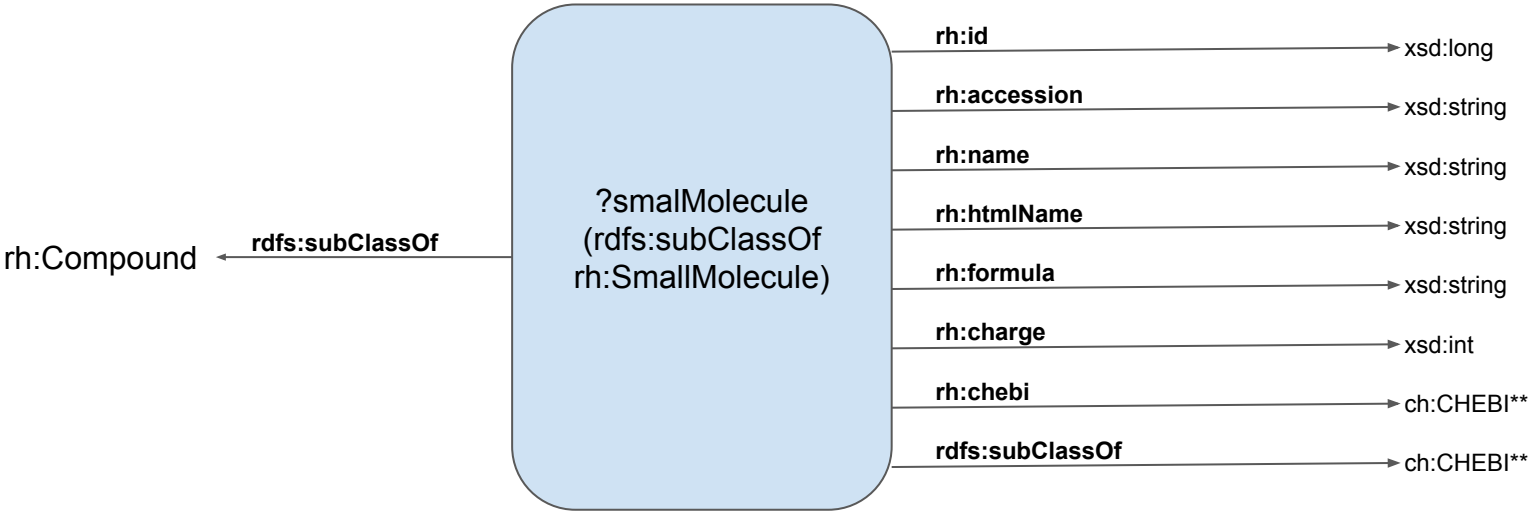
rh:contains & rh:contains2 (example)




GenericPolypeptide
D-alanyl-[D-alanyl-carrier protein]
GENERIC:14103
O-(S-D-alanylpantetheine-4'-phosphoryl)-L-serine residue
CHEBI:138620
Formula: C17H31N4O9PS
Charge: 0
Position:
GenericPolynucleotide
tRNA pseudouridine65
GENERIC:10104
pseudouridine residue
CHEBI:65314
Formula: C9H10N2O8P
Charge: -1
Position: 65
D-alanine
CHEBI:57416
Formula: C3H7NO2
Charge: 0
a D-amino acid
CHEBI:59871
Formula: C2H4NO2R
Charge: 0
[(1→6)-α-D-glucosyl](n+1)
POLYMER:11144
Formula: H2O(C6H10O5)n+1
Charge: (0)(0)n+1
ChEBI polymer: CHEBI:18269
small molecule
generic
polymer
reaction participants
Rhea
reactions

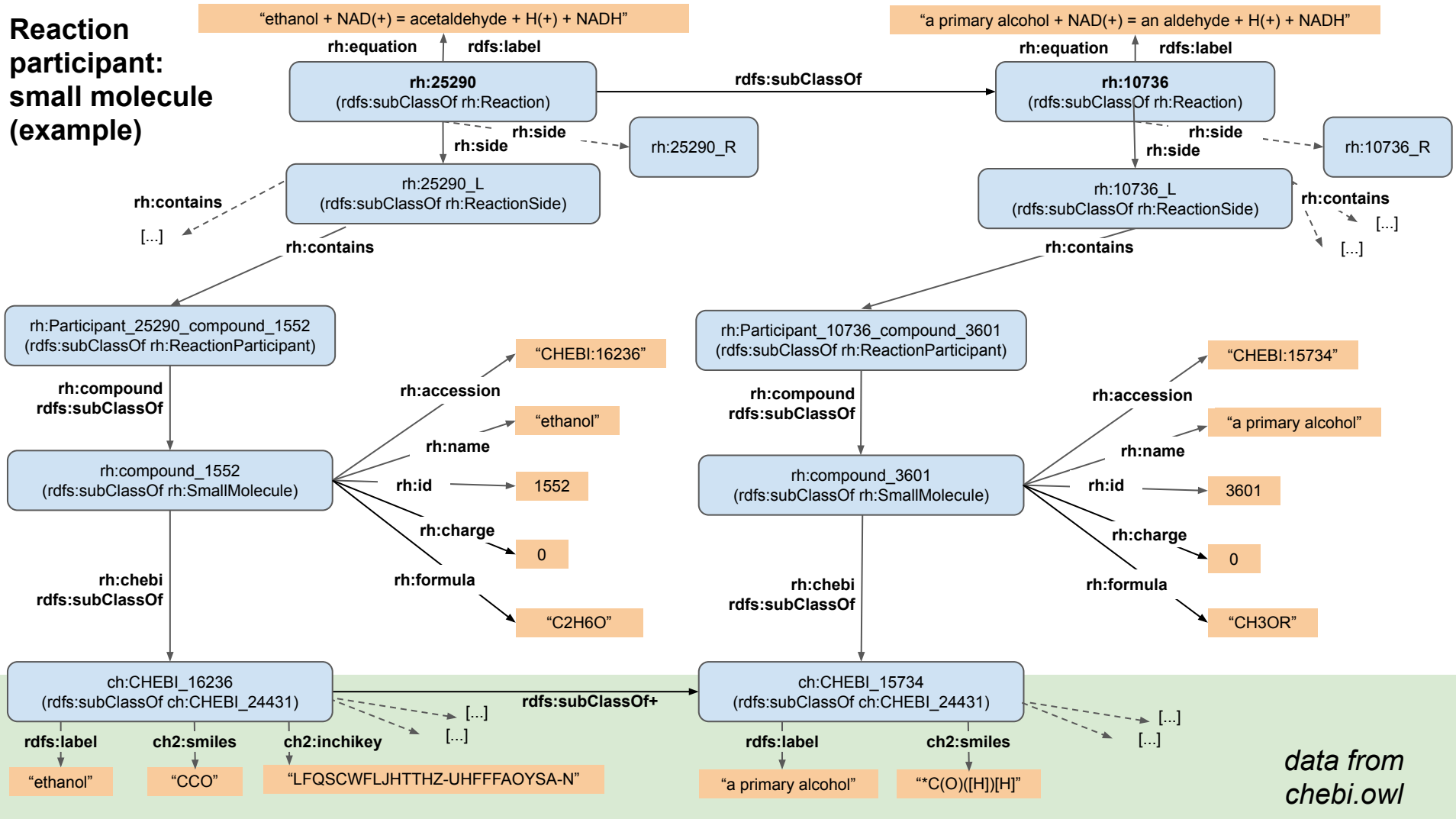

rh:ReactionParticipant
(rdfs:subClassOf rh:Compound)
rh:SmallMolecule
rh:GenericCompound
rh:Polymer

rh:SmallMolecule



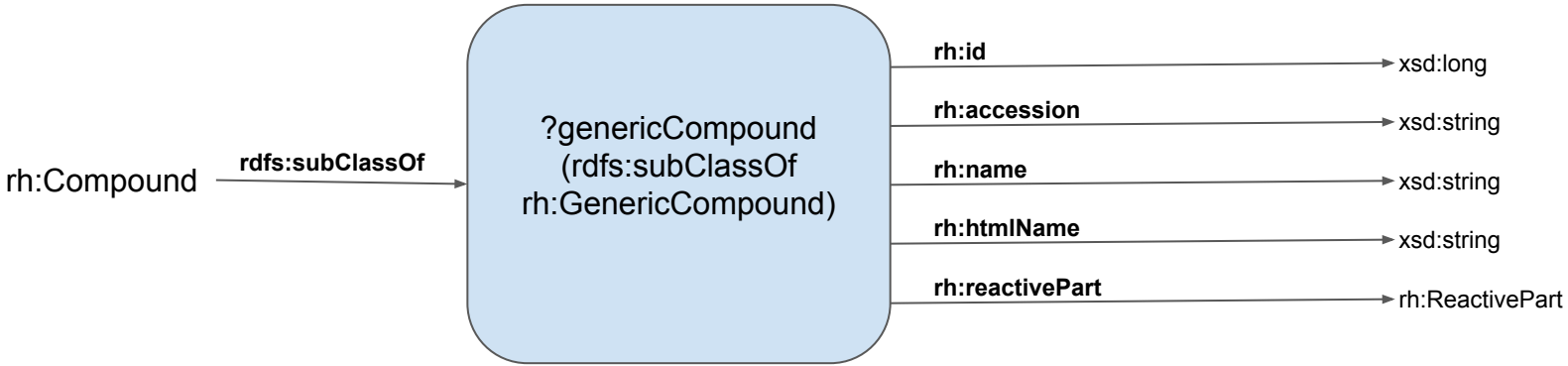
Definition:
A chemical compound that is a low molecular weight (MW) molecule involved as participant, in a chemical or transport reaction.
It is a subclass of rh:Compound.
It is a subclass of a given ChEBI chemical entity.

Reaction participant: small molecule (example)

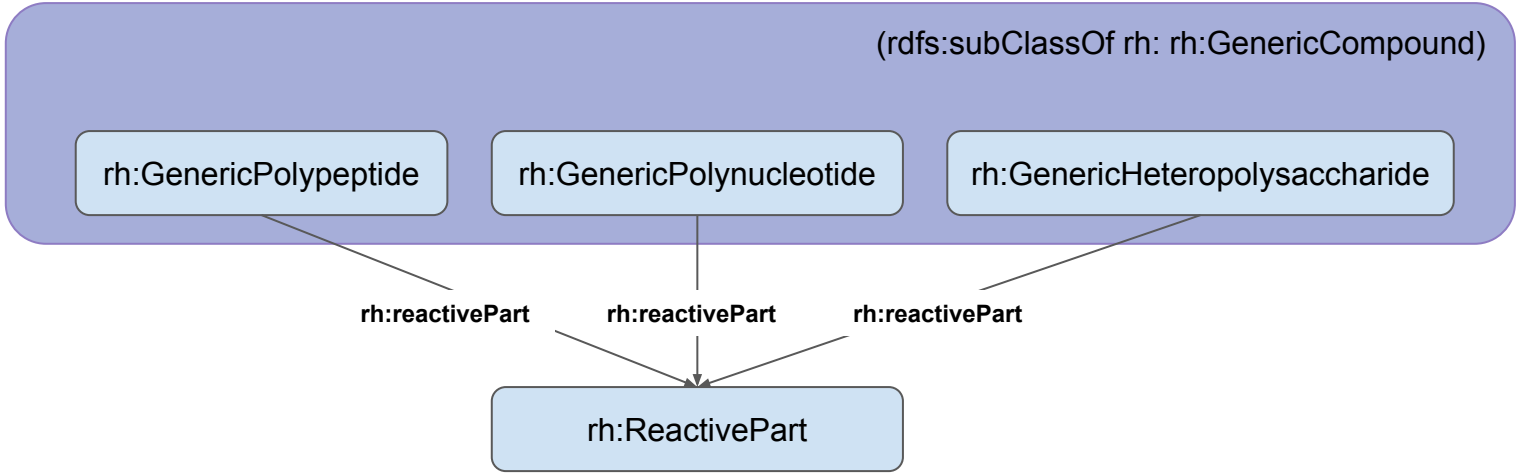


data from chebi.owl

rh:GenericCompound

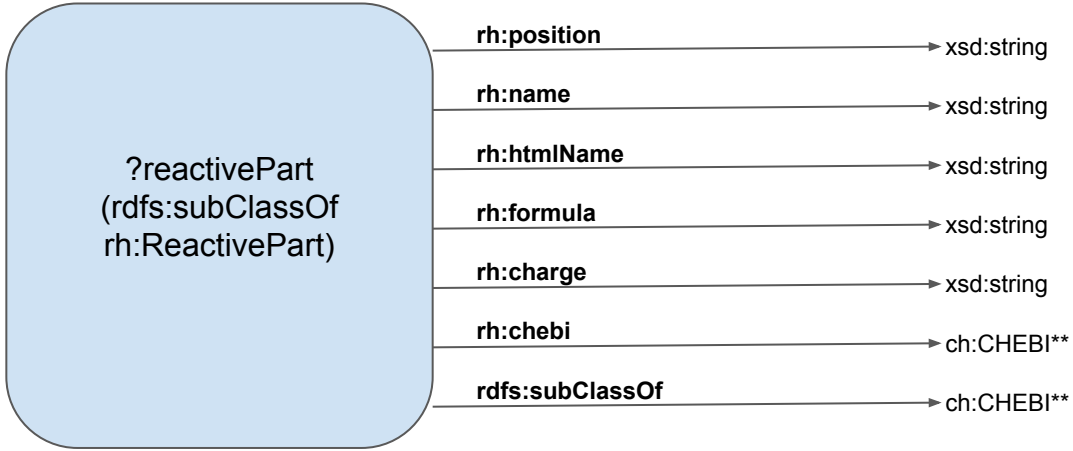


Definition:
A macromolecule involved as participant, in a chemical or transport reaction. Its chemical structure is not fully described but simplified to the functional group(s) involved in the reaction. These reactive part(s) are chemically defined.
rh:GenericCompound represents three kind of macromolecules: proteins (rh:GenericPolypeptide), nucleic acids (rh:GenericPolynucleotide) and heteropolysaccharides (rh:GenericHeteropolysaccharide).
rh:GenericCompound is a subclass of rh:Compound and is subclassed in rh:GenericPolypeptide, rh:GenericPolynucleotide, rh:GenericHeteropolysaccharide.



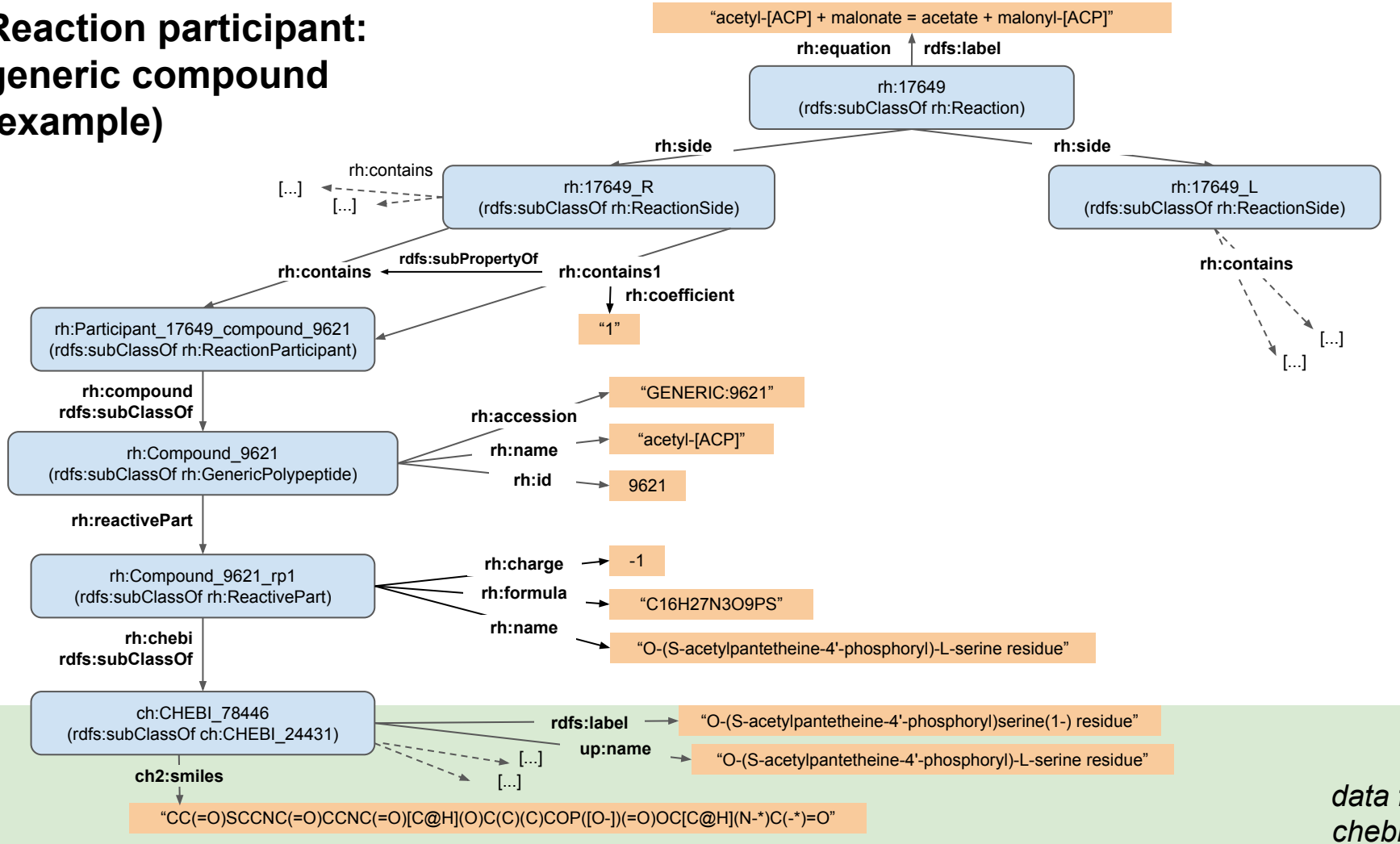
rh:GenericPolypeptide	A protein (polypeptide sequence). It is a subclass of rh:GenericCompound.
rh:GenericPolynucleotide	A DNA or RNA nucleic acid (polynucleotide sequence). It is a subclass of rh:GenericCompound.
rh:GenericHeteropolysaccharide	An heteropolysaccharide. It is a subclass of rh:GenericCompound.

rh:ReactivePart



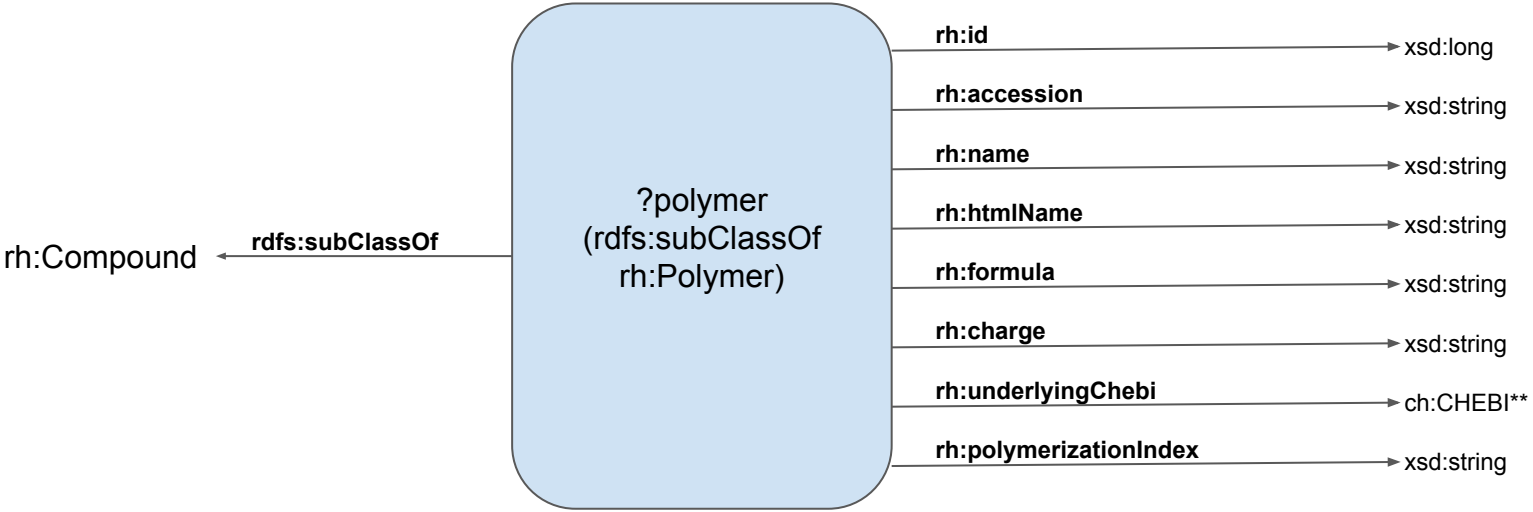
Definition:
The reactive part of a macromolecule (rh:GenericCompound). It has a defined chemical structure.
It is a subclass of a given ChEBI chemical entity.

Reaction participant: generic compound (example)



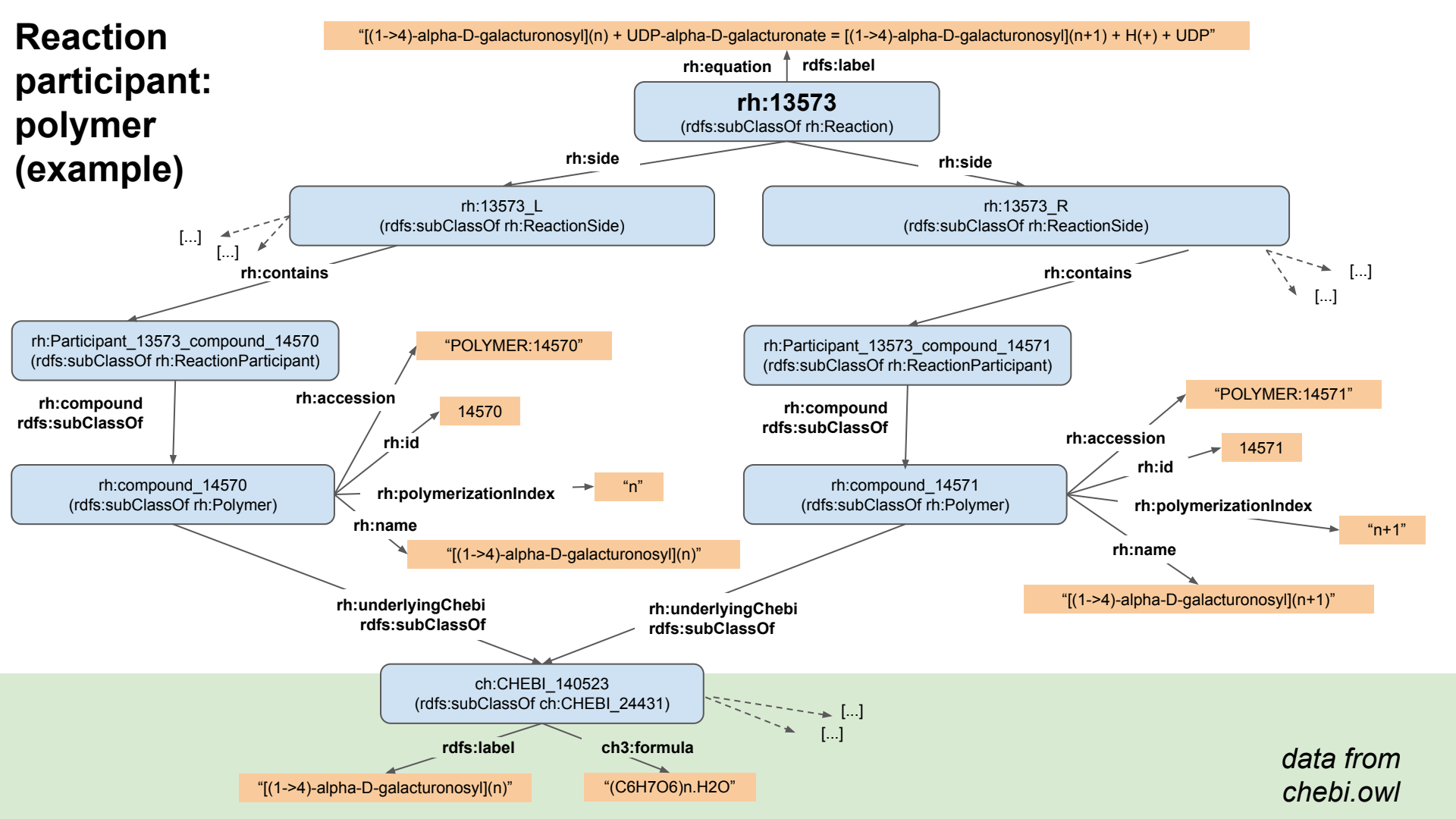
data from chebi.owl

rh:Polymer



Definition:
A simple polymer (a molecule with defined repeated unit(s) and polymerization index (PI) involved as participant, in a chemical or transport reaction. It is defined by an underlying ChEBI polymer and a PI (n, n+1, n-1, etc).
It is a subclass of rh:Compound.

Reaction participant: polymer (example)

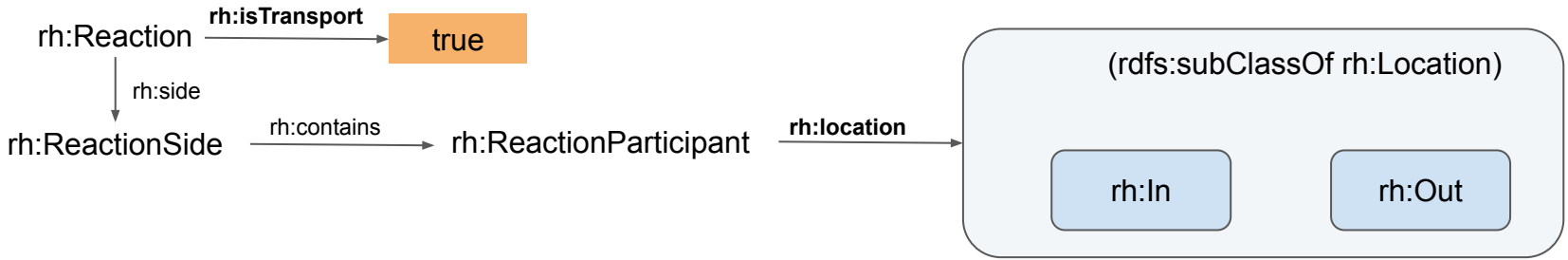


data from chebi.owl

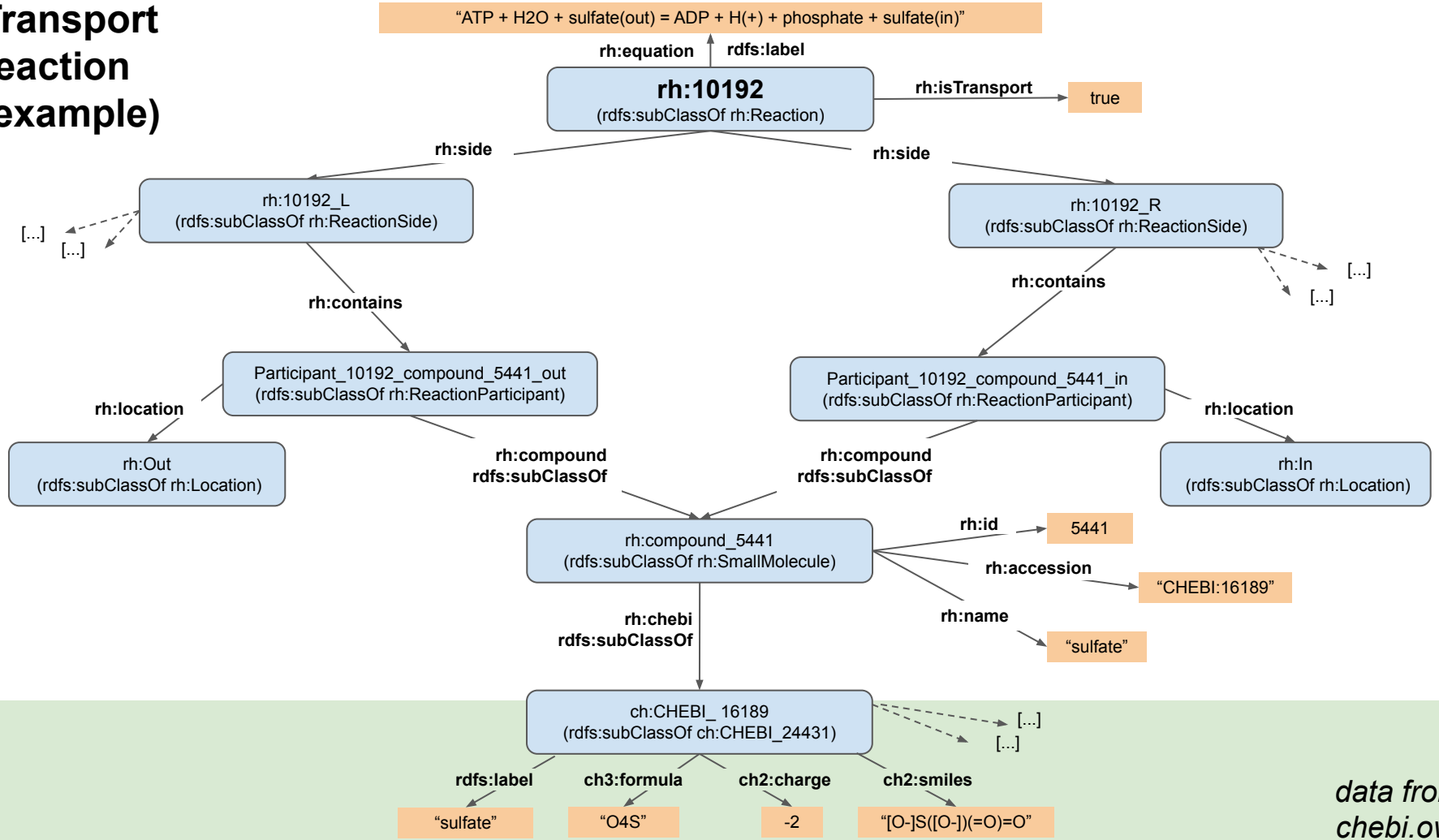
Transport reaction and Location

rh:isTransport	Boolean flag indicating if a reaction is a transport reaction. In transport reaction, the same compound is found in both reaction sides but with a different location.

rh:Location	The location of a reaction participant (rh:ReactionParticipant) in a transport reaction.
rh:In	It is a subclass of rh:Location.
rh:Out	It is a subclass of rh:Location.



Transport reaction (example)

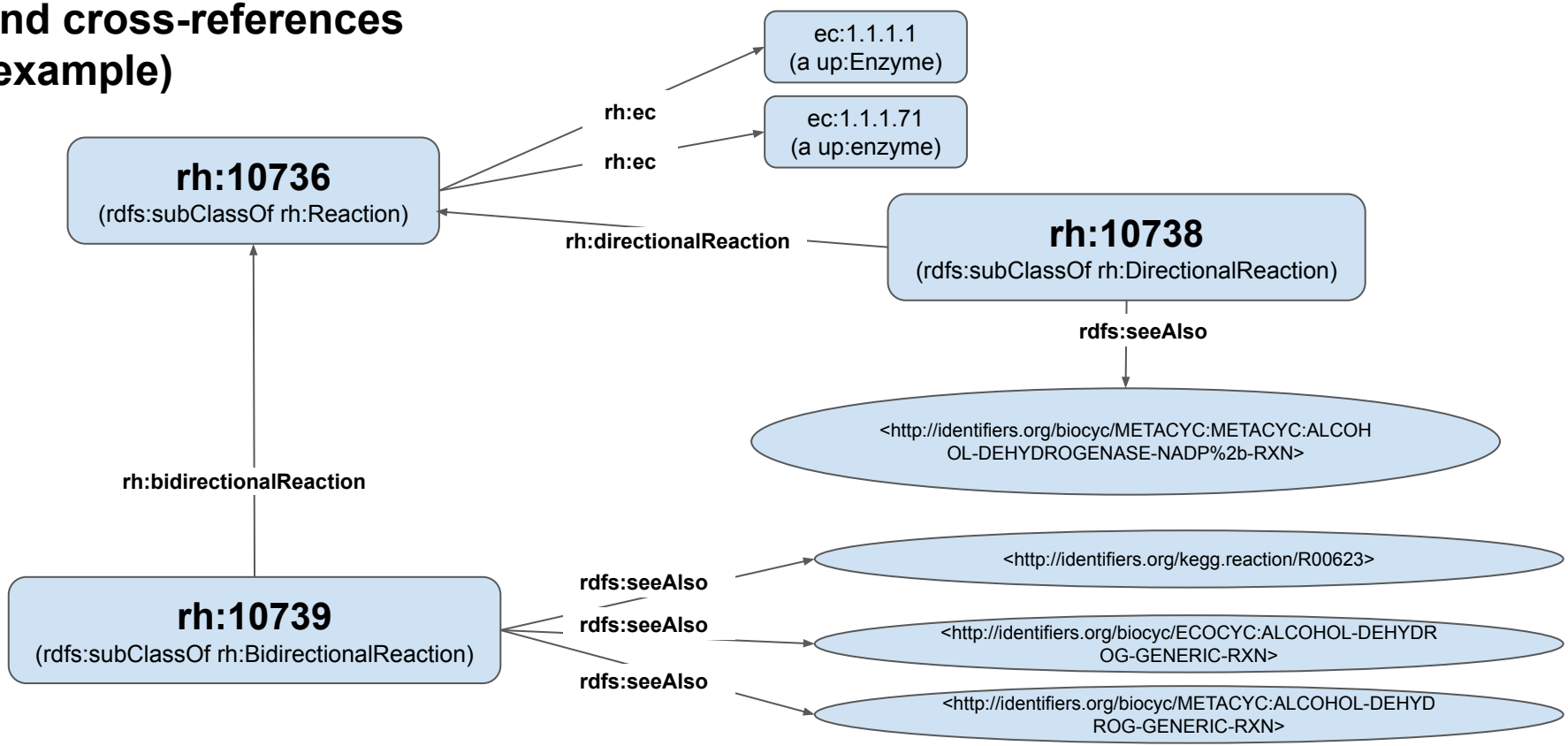


data from chebi.owl

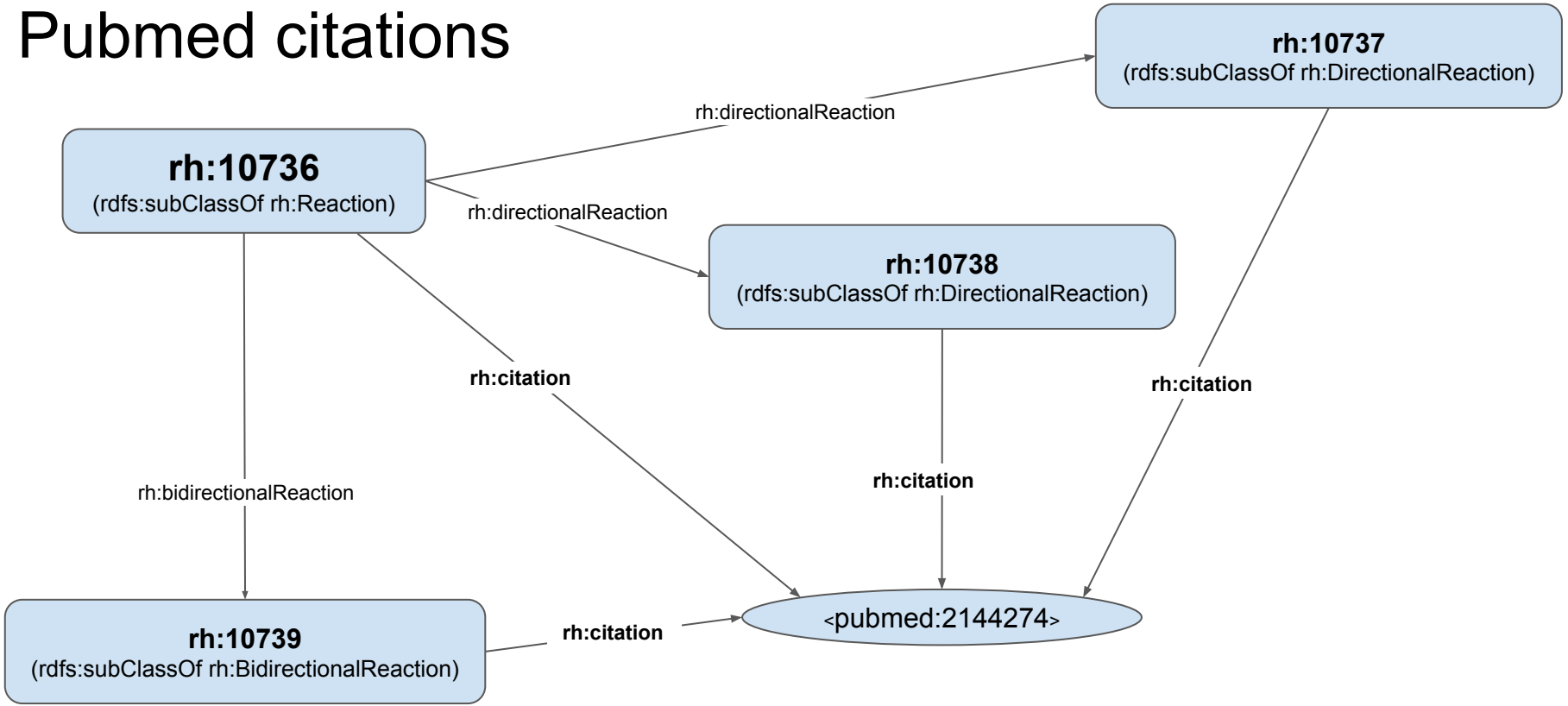
Links to EC numbers and cross-references

rh:ec	The nomenclature committee of the IUBMB defines enzyme classes based, among others, on the catalyzed reaction(s). The enzyme classes are identified by an Enzyme Commission number (EC number). rh:ec is used between a rh:Reaction (the catalyzed reaction) and a complete EC number (the enzyme class).
rdfs:seeAlso	The cross-references to other metabolic/protein resources. There are two kind of cross-references: - Manually curated (KEGG reaction, MetaCyc reaction, EcoCyc reaction) - Computed based on ChEBI participants (Reactome (hsa reaction), MACiE reaction) The MetaCyc/EcoCyc reactions are linked to either rh:Reaction (no define direction), rh:DirectionalReaction or rh:BidirectionalReaction. The KEGG reactions are linked to rh:BidirectionalReaction. The cross-references to UniProt protein entries are not part of the rhea.rdf distribution. They are available from the uniprot.rdf distribution.

EC numbers and cross-references (example)

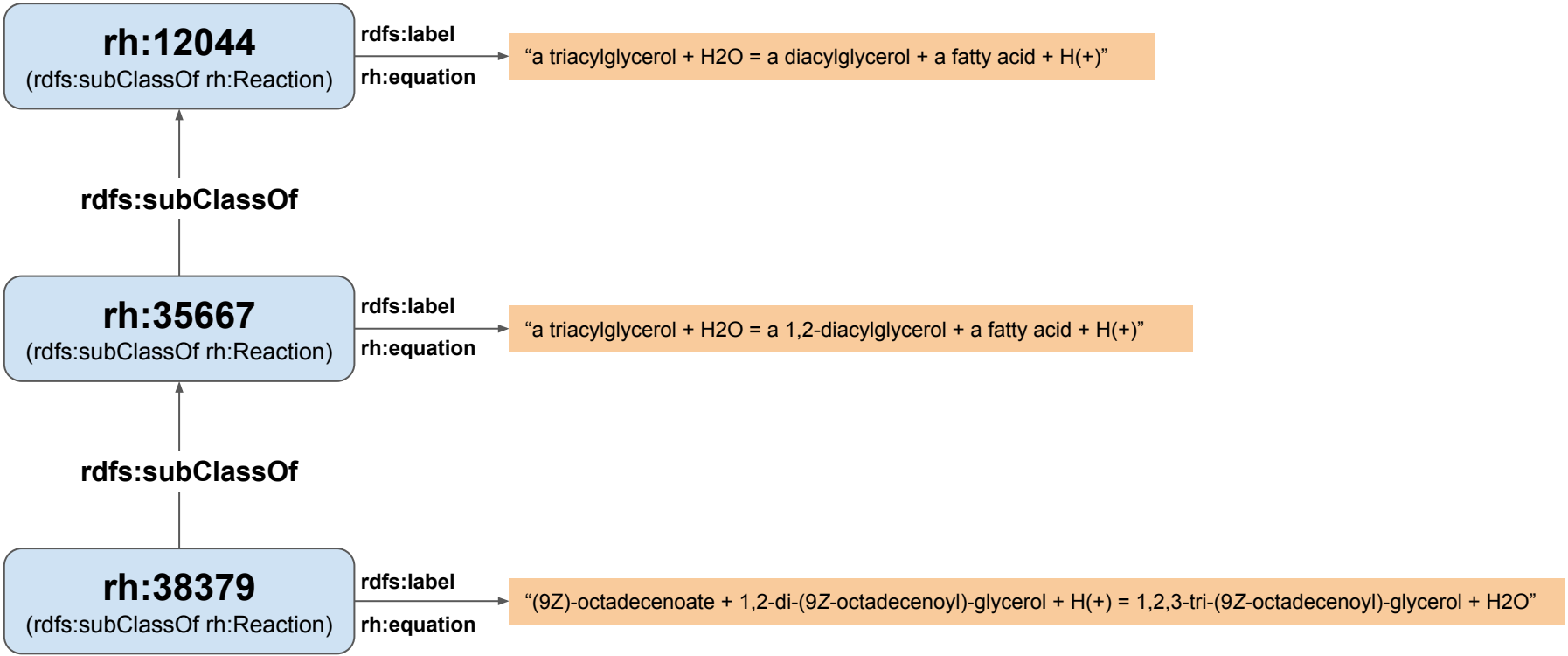


Pubmed citations



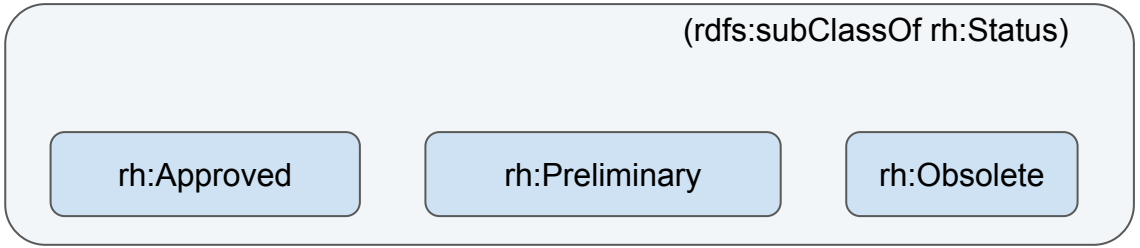
rh:citation	The relation between a rh:Reaction, a rh:DirectionalReaction or rh:BidirectionalReaction and a PubMed URI.

Reaction hierarchical classification

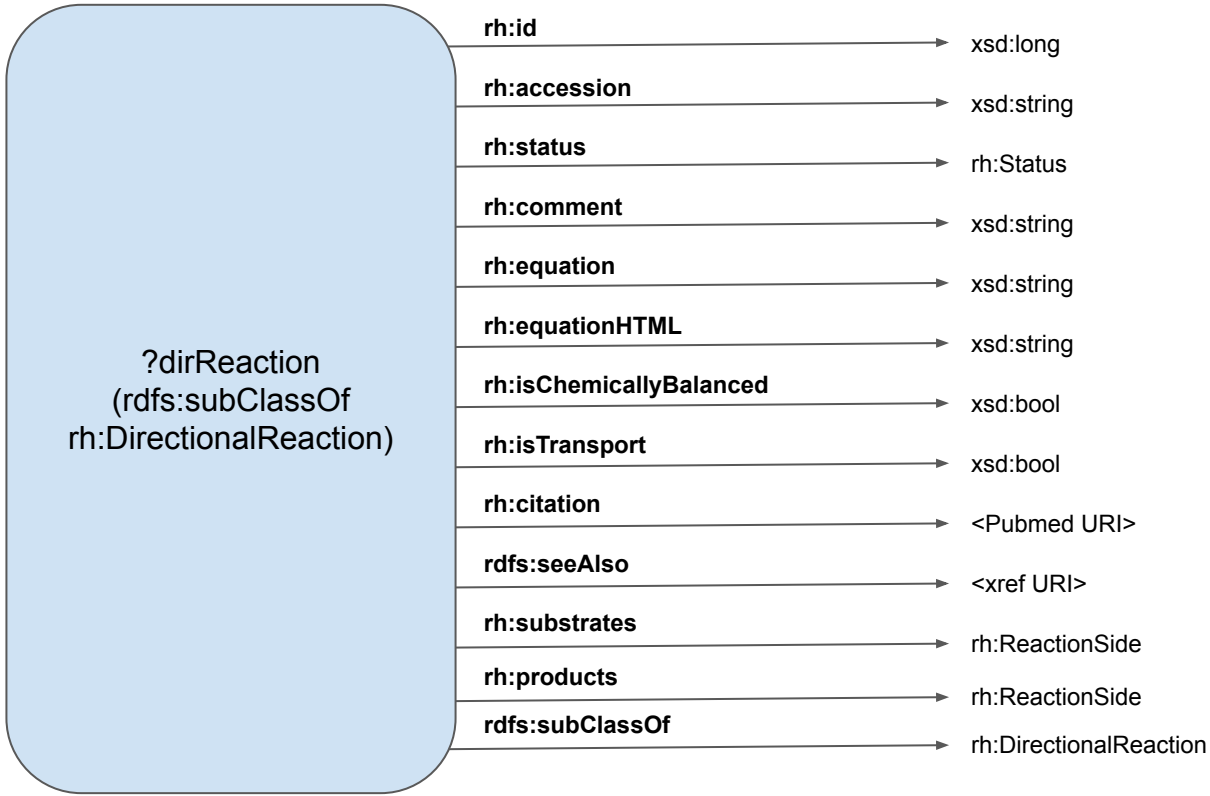


Rhea curation status

rh:Status	Curation status of a reaction (rh:Reaction).
rh:Approved	**Manually curated reaction**.The reaction is evidenced and is chemically balanced at the level of mass and charge. It is a subclass of rh:Status.
rh:Preliminary	**Manually curated reaction**. The reaction is indirectly evidenced and/or some knowledge is missing (e.g it has some undefined participant(s)). The reaction is possibly not balanced at the level of mass and charge. It is a subclass of rh:Status.
rh:Obsolete	The reaction is no more valid.

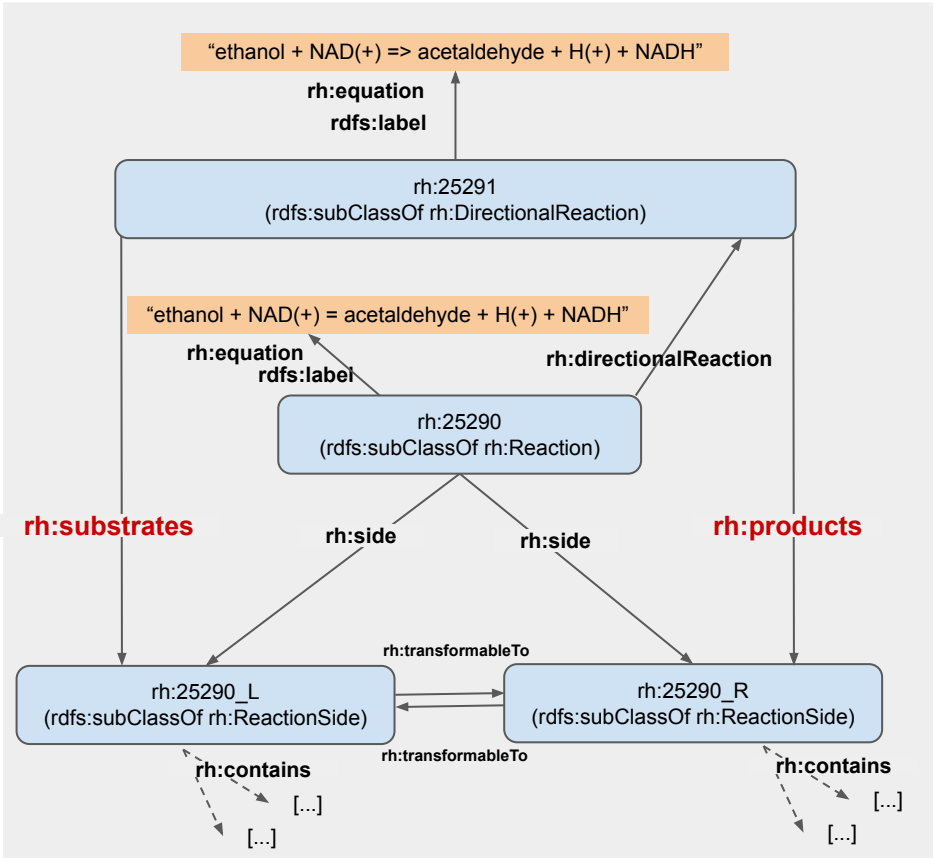


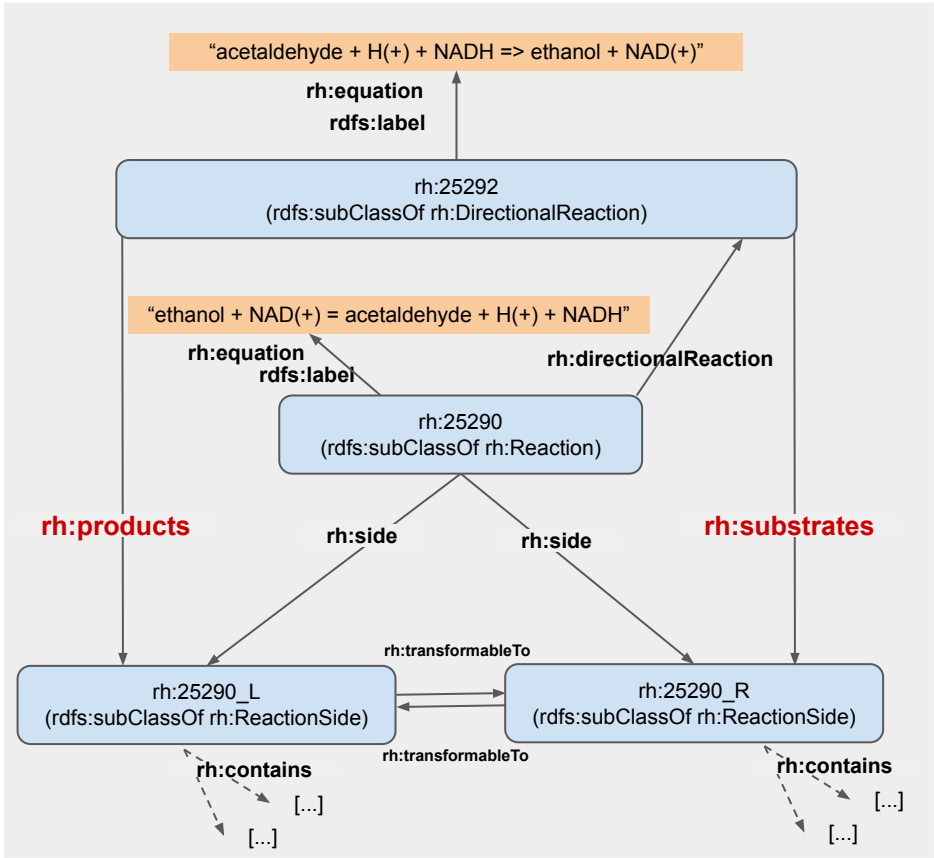
rh:DirectionalReaction



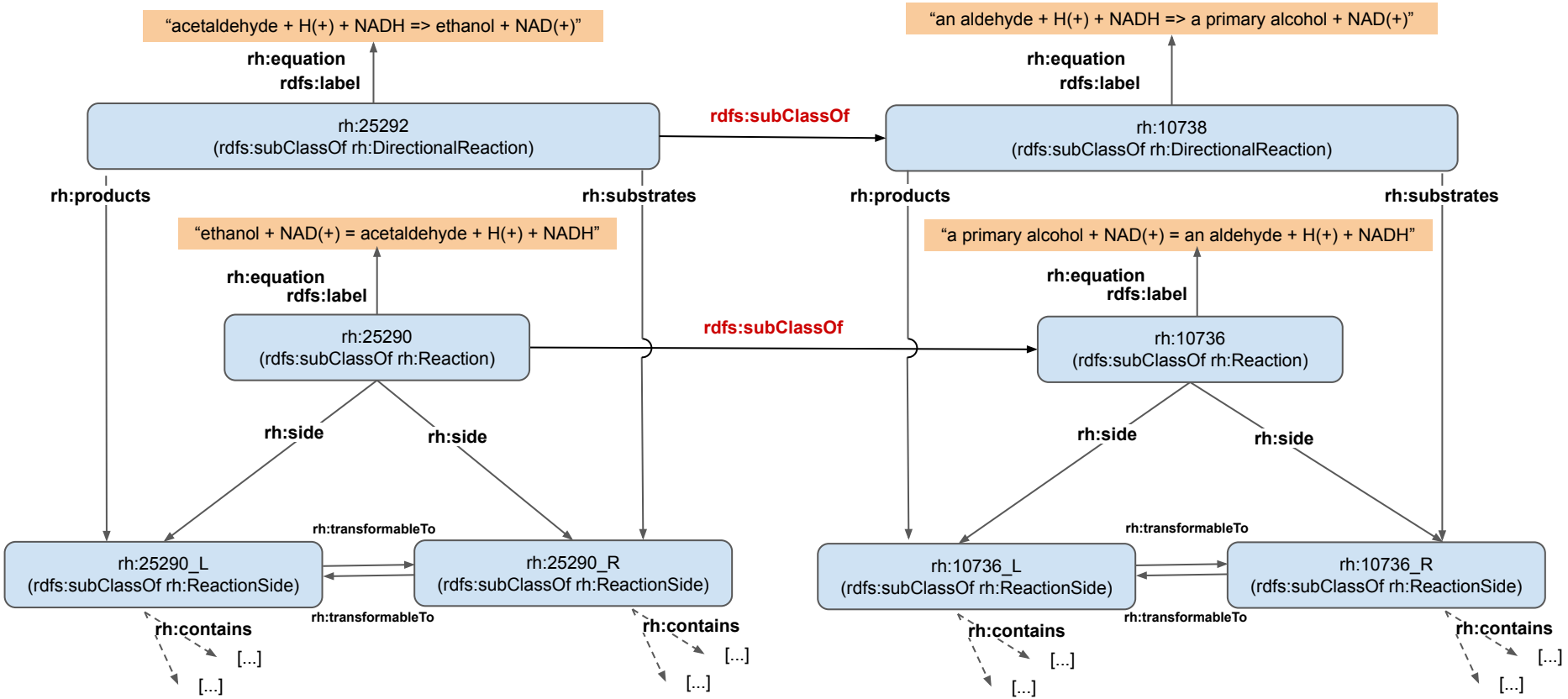
Definition:
A chemical or transport reaction of biological interest with a direction specified by substrates and products.
All the directional reactions published by Rhea are rdfs:subClassOf rh:DirectionalReaction.

rh:DirectionalReaction (example1)

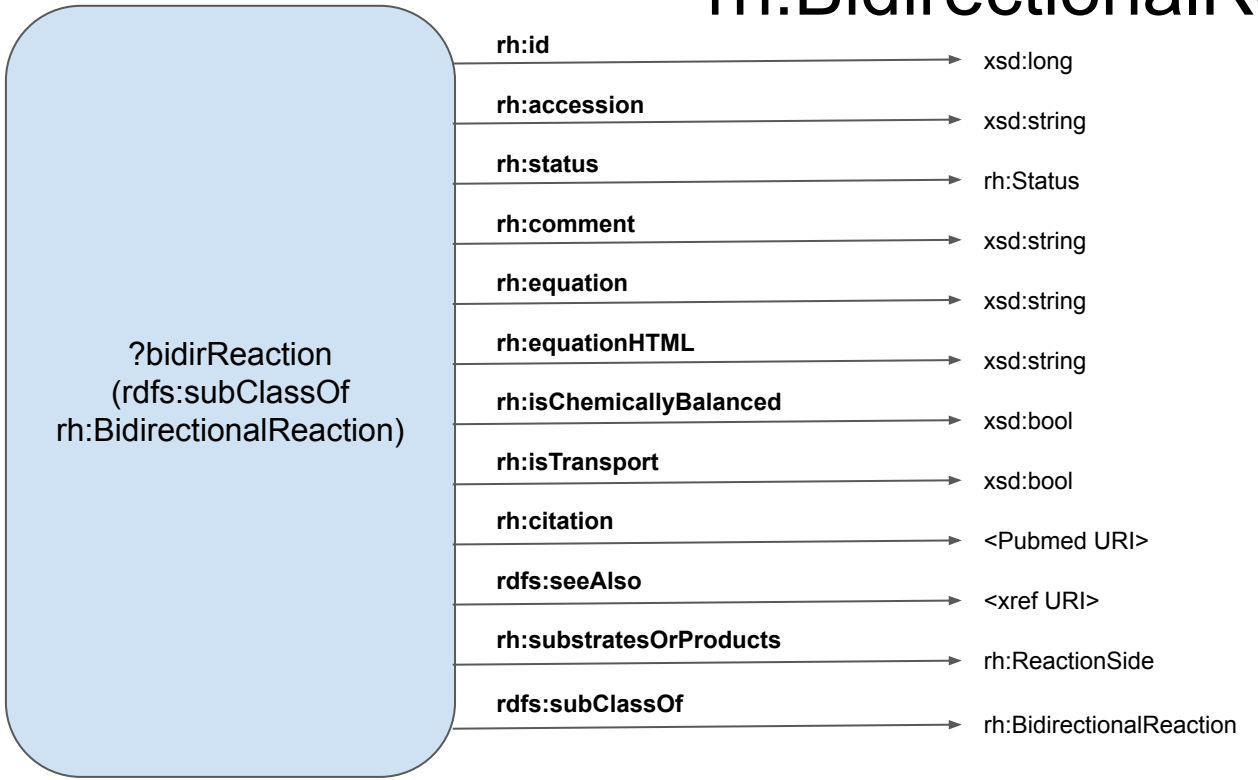




rh:DirectionalReaction (example2)

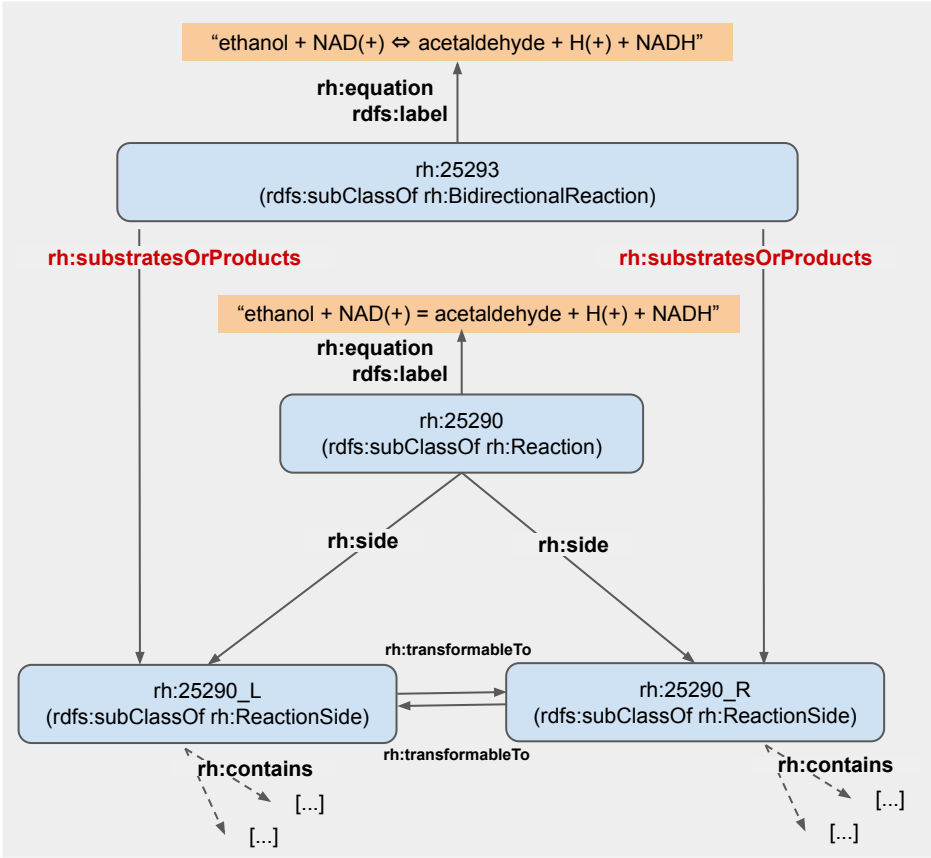


rh:BidirectionalReaction



Definition:
A chemical or transport reaction of biological interest which can happen in one direction or the other, depending on the physiological conditions.
All the bidirectional reactions published by Rhea are rdfs:subClassOf rh:BidirectionalReaction.

rh:BidirectionalReaction (example1)



rh:BidirectionalReaction (example2)