TBI

universität wien

# Creating an Enzyme Database with Rewrite Rules

35th TBI Winterseminar 2020

13.2.2020

By Bojana Ristivojcevic

Supervisor: ao. Univ.-Prof. Mag. Dr. Christoph Flamm

# Content

- Aim of my master thesis
- Graph Rewrite Rule
- Causal Analysis
- Benefits and useful application of the approach
- Workflow
- References

# Aim

- To create an online database, which stores enzymatic reaction graph rewrite rules
- Build on-demand chemical networks

# Network

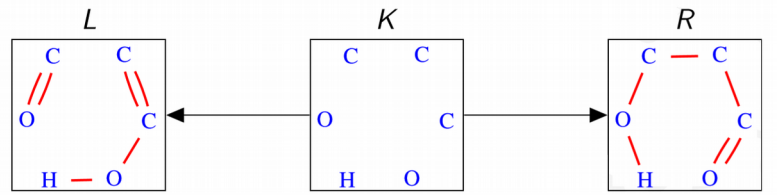
- Metabolism can be regarded as a network of chemical reactions activated by enzymes and connected via their substrates and products
- Chemical reaction networks can be represented as directed (multi)hypergraphs whose vertices are the molecules and whose hyperedges represent chemical reactions

# Graph Rewrite Rule

- Chemical reaction is expressed as a graph transformation rule

- A reaction pattern is described as a double pushout formalism

$$p = (L \xleftarrow{l} K \xrightarrow{r} R),$$



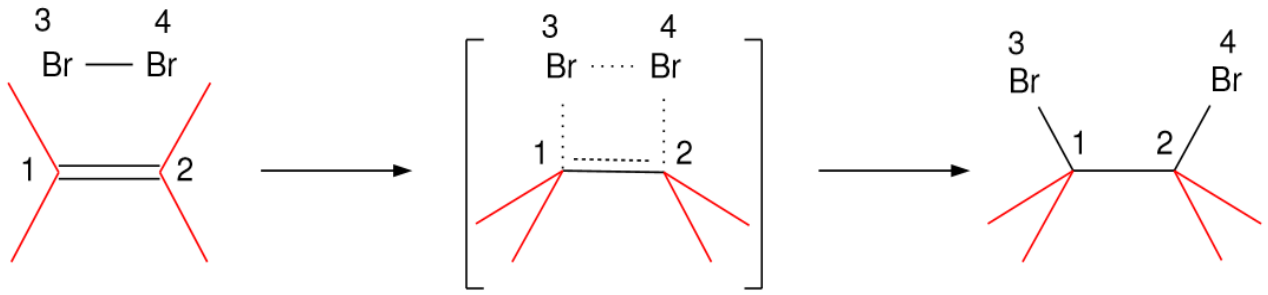

- Graph Grammar Rules are encoded using the GML format (Graph Modelling Language)

```
key1 [
 key2 value2
 key3 [
  key4 value4
  key5 value5
 ]
 key6 value6
]
```

***GML format***


1. Andersen, J. L., et al (2018). Rule composition in graph transformation models of chemical reactions. *Match*, *80*(3), 661-704.
2. Andersen, J. L., et al (2013). Generic strategies for chemical space exploration. arXiv preprint arXiv:1302.4006.
3. https://www.tbi.univie.ac.at/software/GGL/Tutorials/tutorial-rules.pdf

# Graph Rewrite Rule



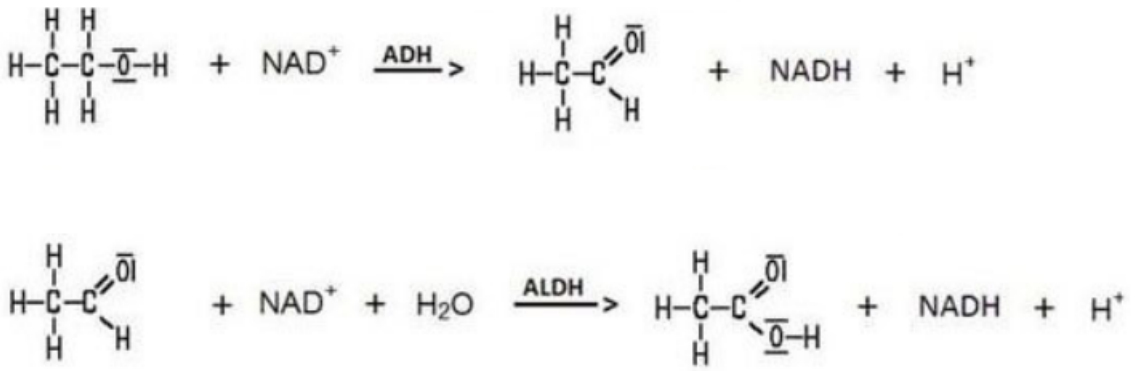

https://www.tbi.univie.ac.at/software/GGL/Tutorials/tutorial-rules.pdf

```
key1 [
 key2 value2
 key3 [
  key4 value4
  key5 value5
 ]
 key6 value6
]
```

```
rule [
 ruleID "Double bond bromination"
 left [
  edge [ source 1 target 2 label "=" ]
  edge [ source 3 target 4 label "-" ]
 ]
 context [
  node [ id 1 label "C" ]
  node [ id 2 label "C" ]
  node [ id 3 label "Br" ]
  node [ id 4 label "Br" ]
 ]
 right [
  edge [ source 1 target 2 label "-" ]
  edge [ source 1 target 3 label "-" ]
  edge [ source 2 target 4 label "-" ]
 ]
]
```

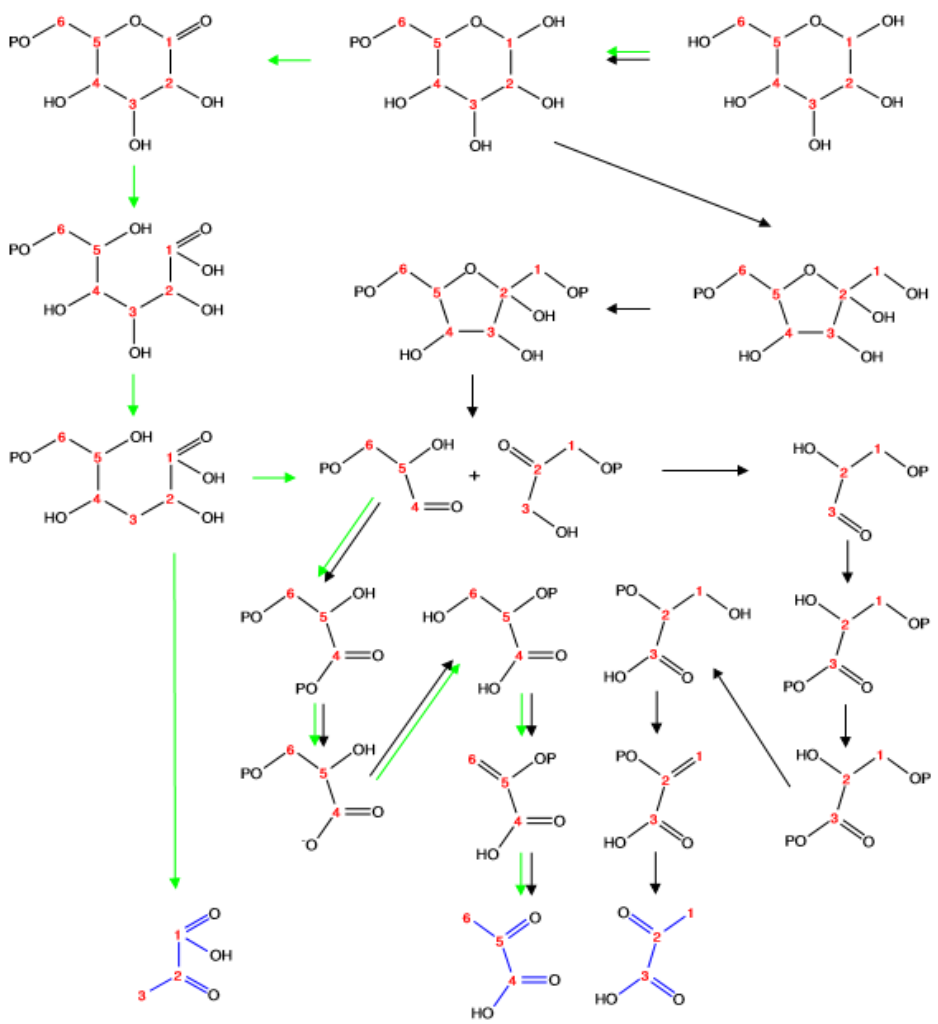
# Causal analysis

- The advantage of Rule- based models lies in their suitability for causal analysis that takes the logically nature of interaction
- The order of reaction can be investigated




Cristescu, I., et al (2019). Interactions between causal structures in graph rewriting systems. arXiv preprint arXiv:1901.00592.

- Databases usually not list atom maps
- Chemical Reaction based on transformation rules support the analysis of isotope labeled experiments
- Isotope labeling experiments in glycolysis are commonly used to analyze the activity of the different pathway variations
- With chemistry model based on DPO transformation rules enable the automatic inference of atom traces for complete pathways



Embden Meyerhof Parnas EMP (black) and Entner Doudoroff ED (green) pathway

Andersen, J. L., et al. (2019, June). Graph Transformations, Semigroups, and Isotopic Labeling. In *International Symposium on Bioinformatics Research and Applications* (pp. 196-207). Springer, Cham.

# Benefits

- The whole framework can be used by people who have no chemical experience
- Because of the atom mapping it is possible to track atoms in reactions
- A useful application of my approach is in planning synthesis of enzyme catalyzed pathways
- The database has the Graph Rewrite Rules stored

# Workflow

- Reaction Atom Mapping Smiles were downloaded from Metacyc Database (Total number:  about 14 000 Atom Mapping Smiles)

- Thiosulfate Transferase:
  Thiosulfate + Cyanide → Sulfite + Thiocyanate

- Atom Mapping Reaction Smile:
  [C:2]#[N:3].[O:1]=[S:5](=[O:6])([O-:7])[S:4]>>[C:2](#[N:3])[S-:4].[O-:1][S:5]([O-:7])=[O:6]

# Workflow

- Fixed some invalid Reaction Smiles and the rest, which could not be repaired,  were removed
- From the valid Reaction Smiles the maximum and minimum Graph Rewrite Rule were extracted
- Currently building the database and also the frontend website that exposes the database
- And also in the process of making the automatic specificity determination

# Acknowledges

Thanks to:

TBI-Team,

- Christoph Flamm

Thanks for your attention !!!