Virtual Synthesis and Metabolism
New Developments

György Pirok
Reactants

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{C}_6\text{H}_5)</td>
<td>(\text{C}_6\text{H}_4)</td>
<td>(\text{C}_6\text{H}_4\text{O}_2)</td>
<td>(\text{C}_7\text{H}_5\text{N})</td>
<td>(\text{C}_6\text{H}_5\text{N}_2)</td>
<td>(\text{C}_6\text{H}_5\text{Cl})</td>
</tr>
</tbody>
</table>

| \(\text{C}_6\text{H}_5\text{COOCH}_2\text{CH}_3\) | \(\text{C}_6\text{H}_5\text{CHCl}_3\) | \(\text{C}_6\text{H}_5\text{CH}_2\text{Br}\) | \(\text{C}_6\text{H}_5\text{CH}_2\text{Cl}\) | \(\text{C}_6\text{H}_5\text{CH}_2\text{Br}\) | \(\text{C}_6\text{H}_5\text{Cl}\) |

ChemAxon Reaction Library

- Baeyer-Villiger ketone oxidation
- Baylis-Hillman vinyl alkylation
- Beckmann rearrangement
- Bischler-Napieralski isoquinoline synthesis
- Friedel-Crafts reaction
- Friedlander quinoline synthesis
- Gabriel synthesis
- Grignard reaction
- Hell-Volhardt-Zelinski halogenation

Products

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{C}_6\text{H}_5\text{COOCH}_2\text{CH}_3)</td>
<td>(\text{C}_6\text{H}_5\text{CHCl}_3)</td>
<td>(\text{C}_6\text{H}_5\text{CH}_2\text{Br})</td>
<td>(\text{C}_6\text{H}_5\text{CH}_2\text{Cl})</td>
<td>(\text{C}_6\text{H}_5\text{CH}_2\text{Br})</td>
<td>(\text{C}_6\text{H}_5\text{Cl})</td>
</tr>
</tbody>
</table>
Reactor the Application
Reactor Module in Instant JChem

The Reactor module of Instant JChem provides an easy to use virtual synthesis application for the scientists working in local or corporate database environments.
Reactor Module in Instant JChem

The reactor module features multimolecular reactions with sequential and combinatorial reactant combination outputting products or specific reactions into database tables. Beside ChemAxon’s reaction library, reactions can be saved as templates in a user library.
ChemAxon’s reaction library currently consists of 145, mostly named reactions. We have recently started to build a new, systematically designed reaction library that will contain several hundred classic preparative reactions. The old reactions are carefully reviewed and inserted in the new library that will be available in a next major release.
JChem for Excel is a new application bringing ChemAxon’s cheminformatics toolkit to Microsoft Excel. Reactor is integrated as one of the first features.
Reactor features

- High performance (supports parallel execution)
- Compatible
  - reactions: MRV, RXN, RDF, SMARTS/SMIRKS
  - compounds: MRV, MOL, SDF, SMILES, InChI, chemical name (IUPAC)
- Well integrated (Wizard, Instant JChem, Oracle Cartridge, JChem for Excel, command line, Java and .NET API)
- Sequential or combinatorial enumeration
- Multimolecular reactions
- Chemo-, regio- and stereospecific main product generation
- Generate products or specific reactions
- Various mapping styles including automapper
- Reverse reaction processing, reactant ratios
- Customizable reaction library, library editor
The KnowTox project

- hepatotoxicity
- metabolic stability
- metabolism
- toxicity

KnowTox
Metabolizer original goals

- Enumerate all possible metabolites of a given substrate
- Estimate metabolic stability class
- Predict major metabolites
- Operate with various biotransformation libraries
- Build human xenobiotic phase I. CYP450 biotransformation library
A generic transformation scheme specifies the structural changes. So the biotransformations of the library are classified by mechanism.
Transformation rules

The generic scheme can be applied specifically to the given substrate with the help of rules written in the form of Chemical Terms. Rules can involve physicochemical property calculations.
Literature examples

Each biotransformation contains a set of examples from the literature for illustration and test purposes. Additional data is stored with the biotransformations including data related to the example reactions.
A biotransformation library

1. Aliphatic epoxidation
2. Alkenyl dioxy compound degradation
3. Aromatic epoxidation
4. Aromatic hydroxylation (CP450)
5. Arylalkyl Phosphate Cyclization Step1
6. Arylalkyl Phosphate Cyclization Step2
7. Benzodiazepine formation (Cyclization)
8. Benzimidazole formation (Cyclization)
9. Benzodiazepine formation (Cyclization)
10. Benzocolchicine formation (Cyclization)
11. Benzocinone to hydroquinone (CP450 mediated reduction) I
12. Benzocinone to hydroquinone (CP450 mediated reduction) II
13. Benzoxazine formation (Cyclization)
14. Catechol oxidation
15. Cleavage of 1,2,3-oxadiazoles
16. Cleavage of N-nitroso pyrroline
17. Cleavage of morpholine
18. Cyclic aromatization of pyridylmethyl aldehyde
19. Cyclohexadienones to phenols (CP450 mediated reduction)
20. Cyclohexadienones to phenols (CP450 mediated reduction) II
21. Dealkylation on Phosphate function
22. Decarboxylation
23. Dehalogenation II (reductive)
24. Dehalogenation II (oxidative carboxylic group formation)
25. Dehydration I (undetectable aliphatic hydroxylation step)
26. Dehydration II (from aliphatic hydroxylated compound)
27. Dehydrohalogenation I (Step 1, hydroxyl intermediate)
28. Dehydrohalogenation I (Step 2, germinal dehydrohalogenation)
29. Dehydrohalogenation II (cumulenic dehydrohalogenation)
30. Dehydrohalogenation III (epoxide formation)
31. Denitration and denitrosation
32. Dephenylation
33. Desaturation of aziridine to alkene
34. Desaturation of cyclic alkef I
35. Desaturation of cyclic alkene II
36. Desaturation of thiol and thioester
37. Desulfuration (thioate oxidation)
38. Dimine to cyclic imine
39. Dihydrodiol formation from aromatic compound I
40. Dihydrodiol formation from aromatic compound II
41. Dihydrodiol formation from epoxide
42. Disulfide reduction
43. Furan formation (Cyclization) Step1
44. Furan formation (Cyclization) Step2
45. hetero-dealkylation and -deformylation (CP450)
46. Hydrazine, hydrazide cleavage to free radical compound
47. Hydroxyalkene to tetrahydrofurane (Cyclization)
48. Imidazolidine from chloroethylurea (Cyclization)
49. Imine formation
50. Iminium ion formation
51. Lactam formation I (Cyclization)
52. Lactam formation II (Cyclization)
53. Long aliphatic chain degradation Step 1.
54. Long aliphatic chain degradation Step 2.
55. Long aliphatic chain degradation Step 3.
56. Metabolic hydrolysis of aziridines
57. Hydrolysis of dioxolane
58. Metabolic hydrolysis of halogenated compound
59. Metabolic hydrolysis of nitrate
60. Metabolic hydrolysis of nucleoside
61. Metabolic hydrolysis of phosphate
62. Metabolic hydrolysis of sulphate
63. Metabolic hydrolysis of thiolactone
64. Metabolic hydrolysis on carbonyl function
65. Methylation of Pyridine
66. Methylation
67. N-dealkylation of alkyl nitroaniline
68. N-dealkylation via N-hydroxylation (Step I)
69. N-dealkylation via N-hydroxylation (Step II)
70. N-dealkylation via N-hydroxylation (Step III)
71. N-formylation
72. N-hydroxylation (Oxidation of amine I)
73. N-oxidation of aromatic compound with N heterocatom
74. NH shift
75. Nitrene formation

The first biotransformation library contains more than two hundred generic phase I human xenobiotic CP450 biotransformations.
The metabolism model

- Fast
- Medium
- Slow

substrate
Substrates consumed by fast biotransformations are unstable.

$max(v)$ is the speed category of the fastest consumption reaction of the given substrate (1: very slow, 2: slow, 3: medium, 4: fast, 5: very fast)

$v_{max}$ is the fastest speed category (5)
Major metabolites are the ones that are accumulated in higher concentrations than others. They are produced by fast transformation routes and consumed by slow reactions.

\[ f(x) = ? \]
Reaction speed prediction problems

We originally had the following batch reaction speed estimation approaches in mind:

- **Calculations from the given substrate**
  - Problem: it is applicable for very few reaction types only

- **The similarity analysis of the same reaction with other substrates**
  - Problem: measurements are available for few reaction types only and the published results are not consistent

- **Estimated for each reaction type**
  - Problem: very raw substrate and site independent approach
Reaction speed prediction solution

• A base speed category value is assigned to each biotransformation.
• These speed categories can be directly used for metabolic stability estimation and can be converted to speed values for major metabolite prediction.
• The base speed values of a biotransformation library can be trained automatically by a random evolutionary optimization tool to reproduce the known major metabolites of drugs. This tool can be used to train any future biotransformation libraries.
Metabolizer application features

- High performance
- Manual exploration and batch enumeration modes
- Exclude unwanted metabolites
- Customizable biotransformations
- Can export to and rebuild metabolic trees from flat SMILES or SDfiles
- Termination conditions to stop branches
- New biotransformation libraries can be plugged in (human phase I, phase II, rat, bacterial, plant, etc.)
- Various coloring options (accumulation, production, major pathways)
- Display and export exact mass (monoisotopic MS Mass) values and pathway codes
A fast method is provided for the prediction of major metabolites, that avoids the extensive enumeration of minor ones.
Future plans

- Training of the human xenobiotic phase one CYP450 biotransformation library
- Indication of metabolically sensitive functionalities
- Reverse metabolism
Acknowledgements

Nóra Máté  Zsolt Mohácsi  Jenő Varga

Attila Szabó  István Cseh  György T. Balogh
Find out more

- Product descriptions & links
  www.chemaxon.com/products.html

- Forum
  www.chemaxon.com/forum

- Presentations and posters
  www.chemaxon.com/conf

- Download
  www.chemaxon.com/download.html