LIBRARIA

Building Drugs from Knowledge

Capturing Chemistry

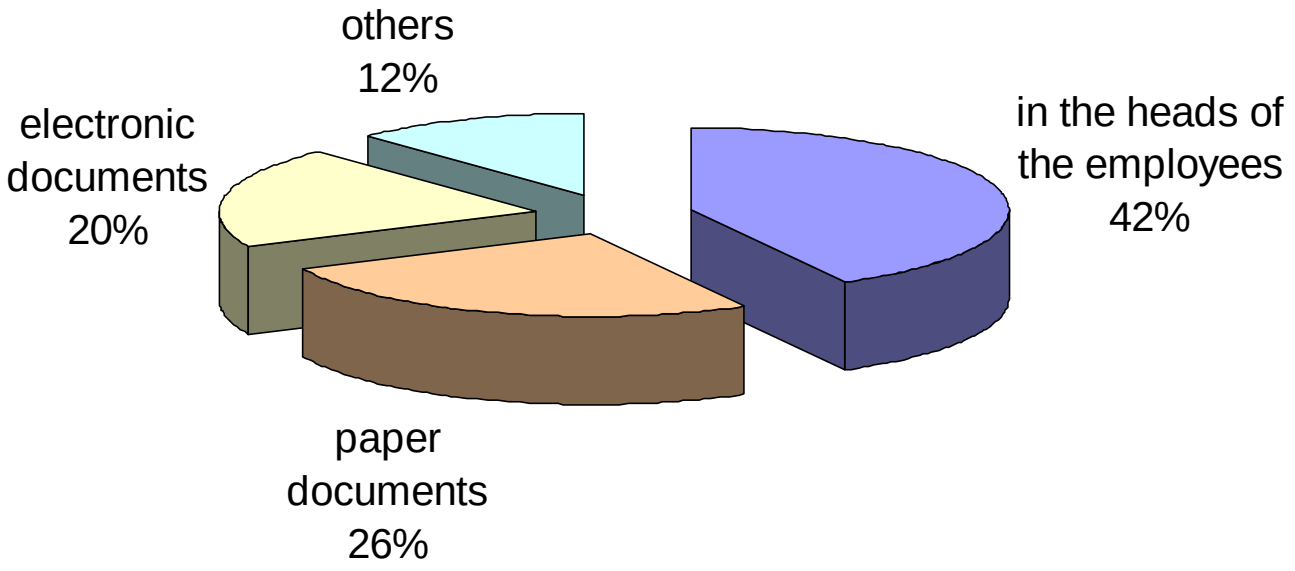
"What you see is what you get"
In the world of mechanism and chemical transformations

Dr. Stephan Schürer

Head of Intl. Sci. Content
Libraria, Inc.

sschurer@libraria.com

Distribution of Knowledge



Nachr. Chem. **2002**, *12*, 1416

in many companies only 20 to 40 % of existing knowledge is utilized

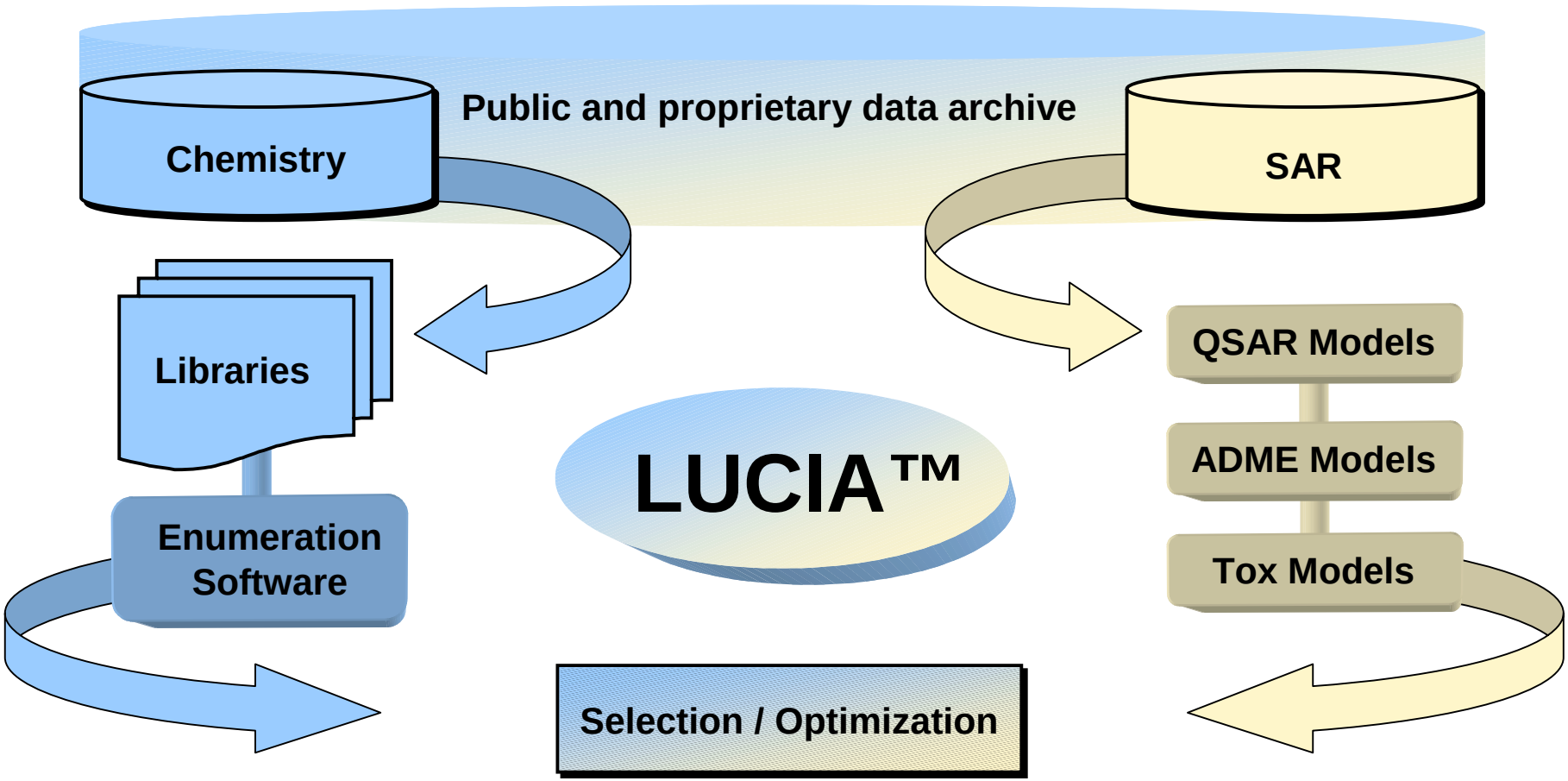
Pharmaceutical Industry

Challenges :

- Patent expiration of many drugs
- Increased R&D costs, but fewer drugs in pipeline
- Uncertainty of success - high attrition rate
- Retirement of researchers and job fluctuation

How to activate and utilize (and preserve) knowledge as a resource

Knowledge-Based Drug Discovery

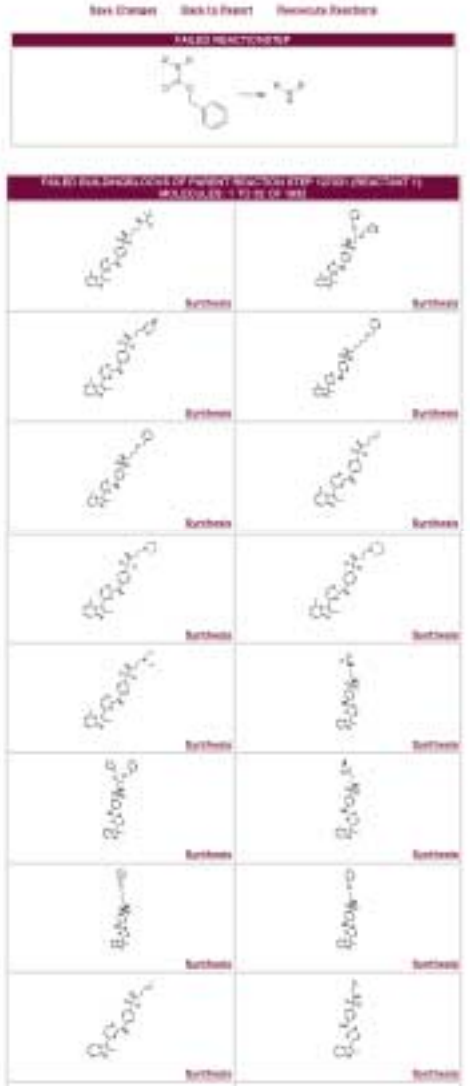


"I think having breadth in chemistry is the way to go" (with respect to identifying multiple lead series)

Chris Lipinski, DDT Jan 2003 interview

Enumeration Engine

- Interprets reaction mechanisms
  - Multi-component reactions
  - Rearrangement reactions
  - Intramolecular reactions – cyclizations
- Complex multi-step transformations
- Reaction step subsequences
- Generic stereochemistry
- Multiple reaction sites
- Multiple matches in building blocks
- Product mixtures
- Formation of stereoisomers
- Salt forms of reactants / products

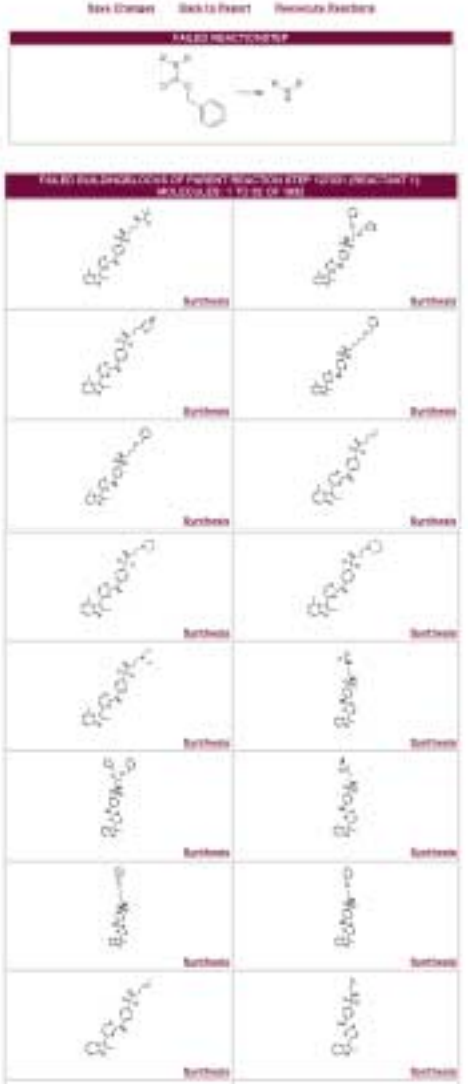


"I think having breadth in chemistry is the way to go" (with respect to identifying multiple lead series)

Chris Lipinski, DDT Jan 2003 interview

Enumeration Engine

- Easy to use
  - All interactions via intuitive GUI
  - Web-browser and drawing tool
  - Interactive graphical debugging functionality
  - Failed BB analysis
  - Synthetic sequences of individual compounds
- Fast
  - 10K library – 2 minutes
  - 100K library – 7 minutes



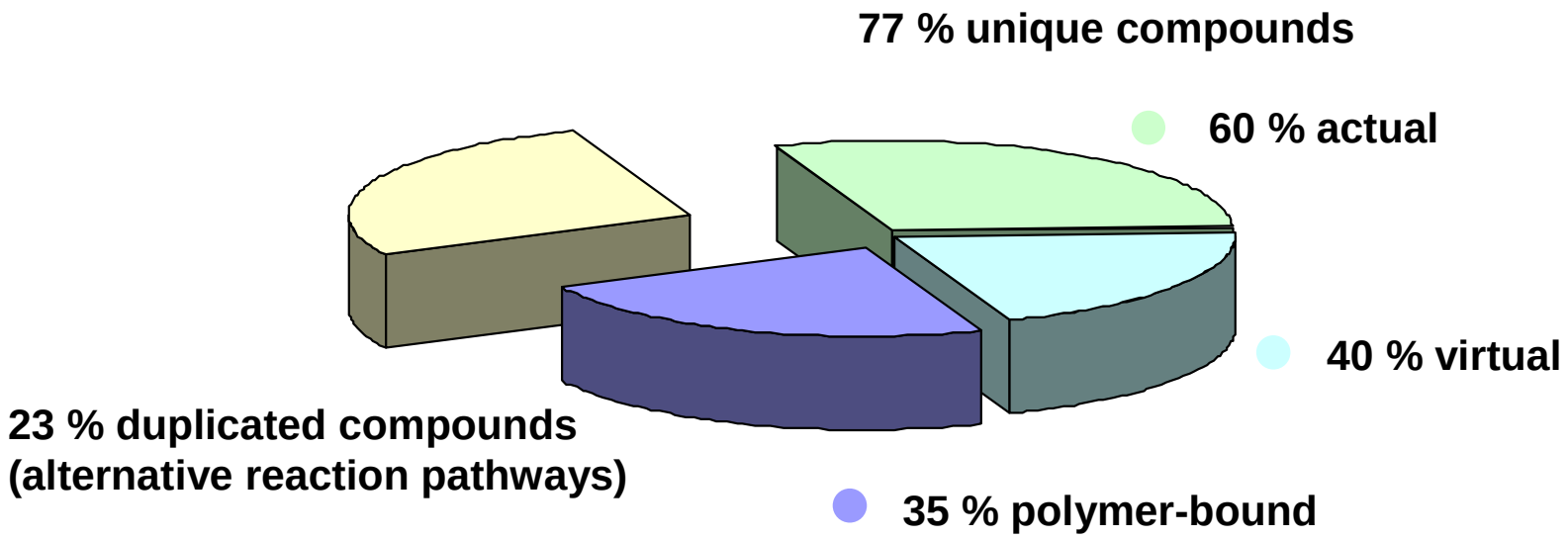
"Where computational chemistry tools really shine is capturing the types of functionality you want to avoid and they are pretty good at filtering by property, but they do a poor job of taking into account, chemical feasibility."

Chris Lipinski, DDT Jan 2003 interview

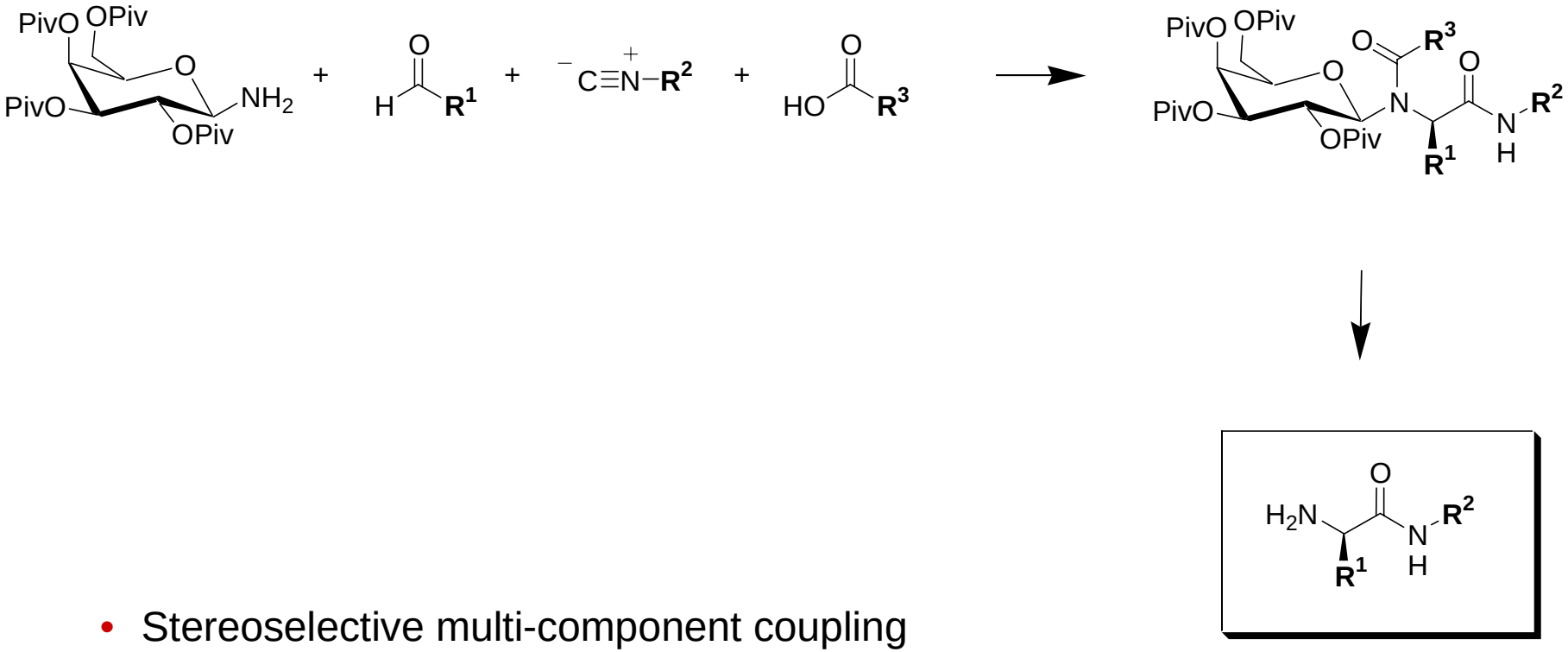
Chemistry Archive

- 14.600 Reaction transforms with detailed experimental procedures
- Reactions are categorized by mechanism and product type
  - High throughput chemistry
  - Kinase-related heterocyclic chemistry

2.3 Million Compounds

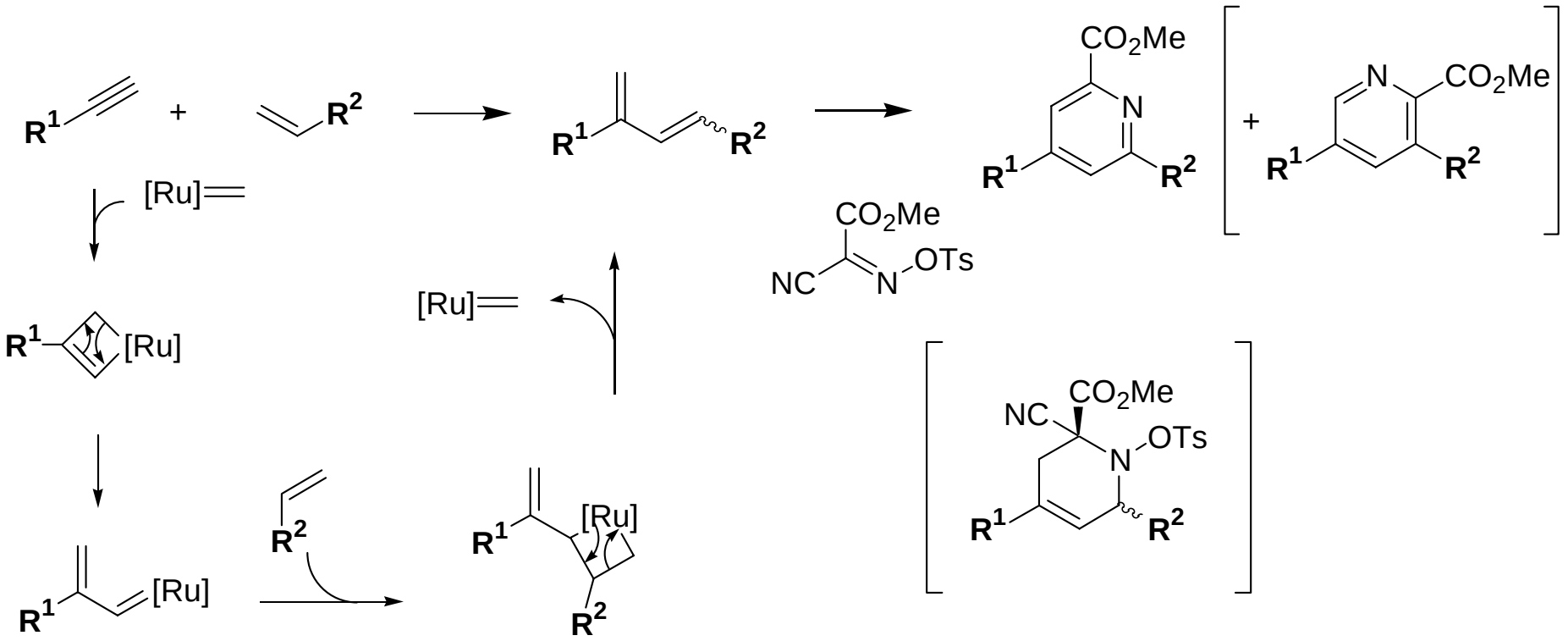


Demo Enumeration Sequence



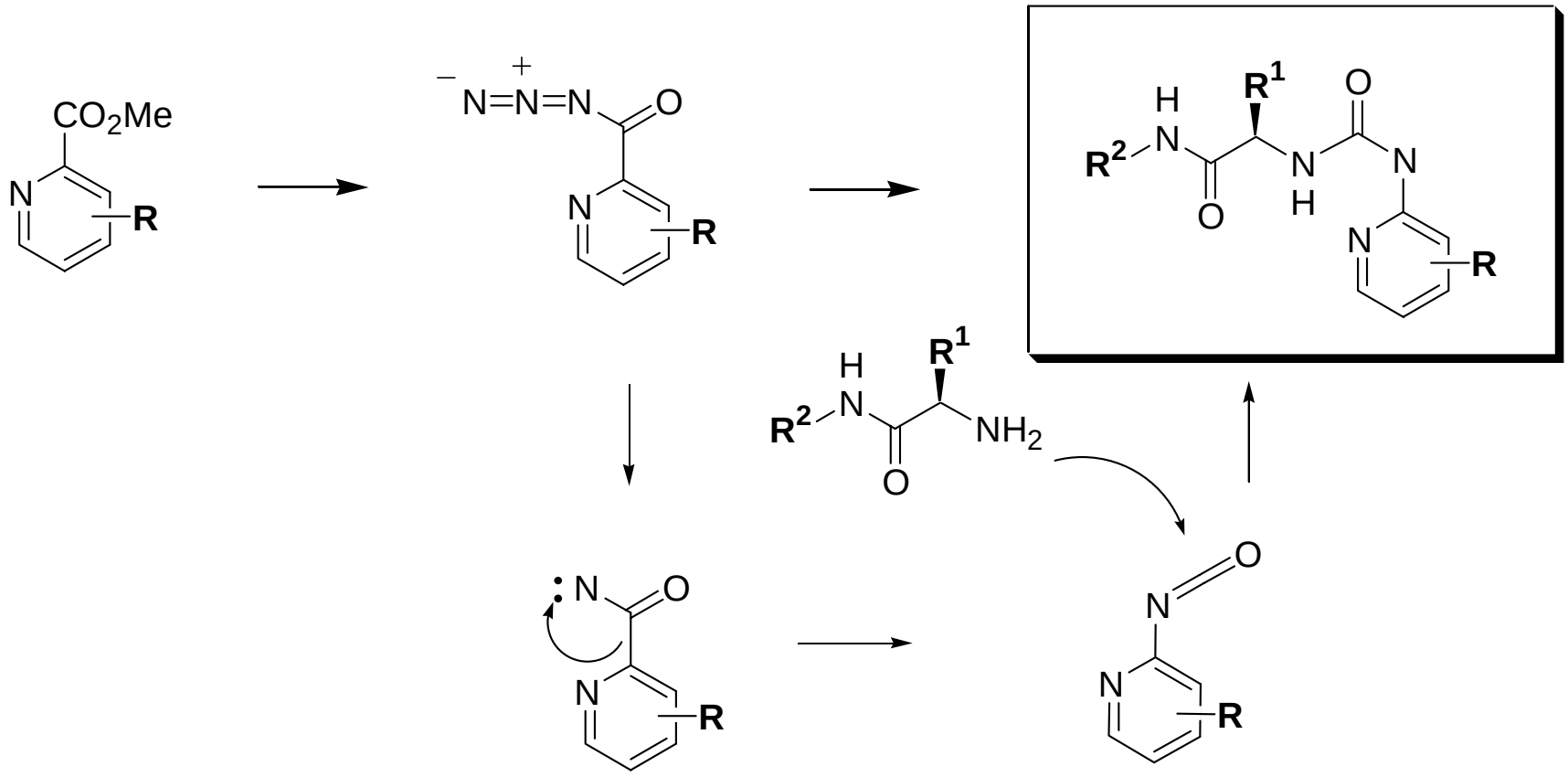
- Stereoselective multi-component coupling (stereochemical induction)
- Generic stereochemistry (chiral centers do not influence the reaction)

Demo Enumeration Sequence



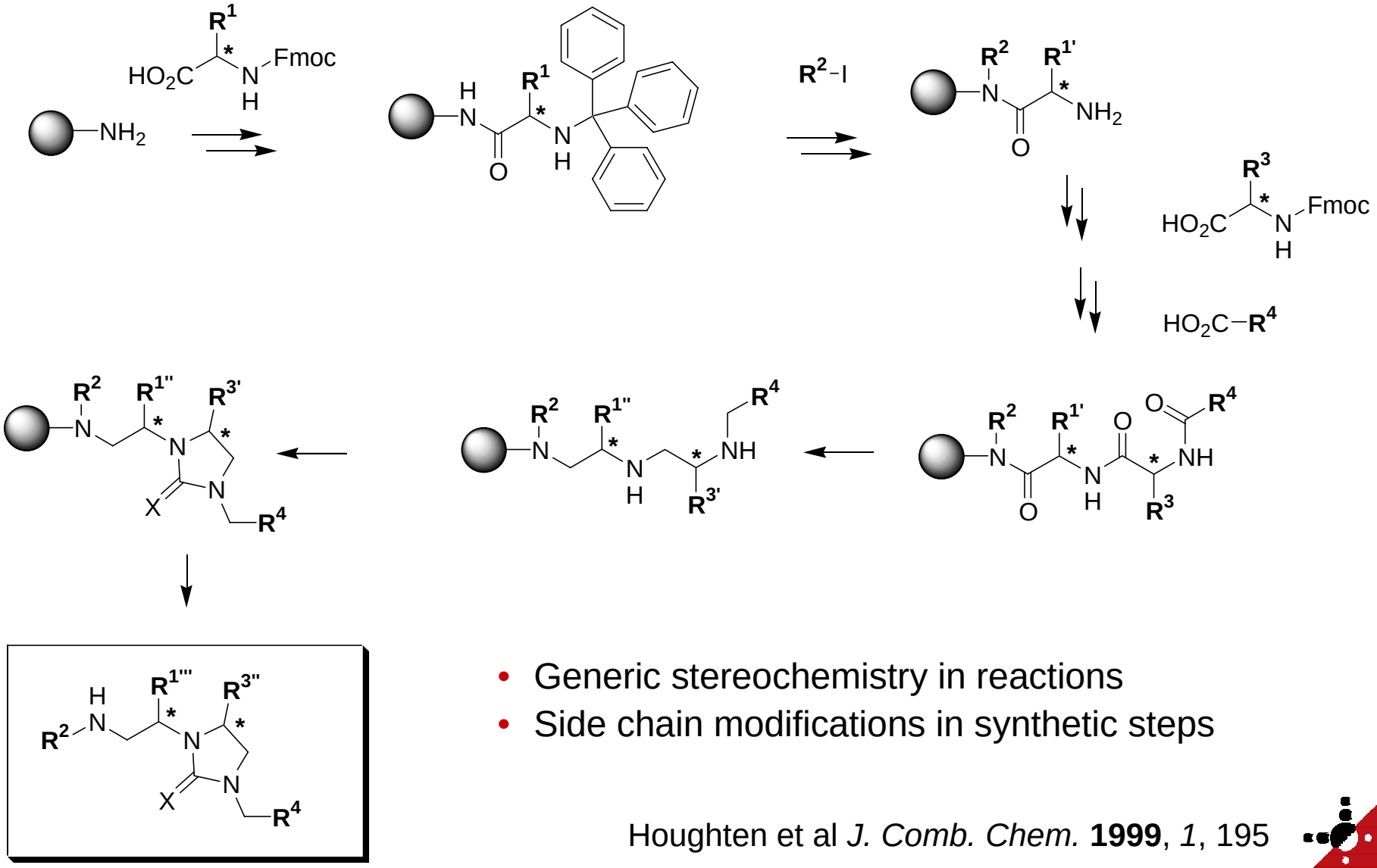
- Yne-ene cross metathesis
- Diels-Alder and double elimination
- Formation of isomeric product mixtures

Demo Enumeration Sequence



- Rearrangement via nitrene
  (Stereochemical configuration is preserved)

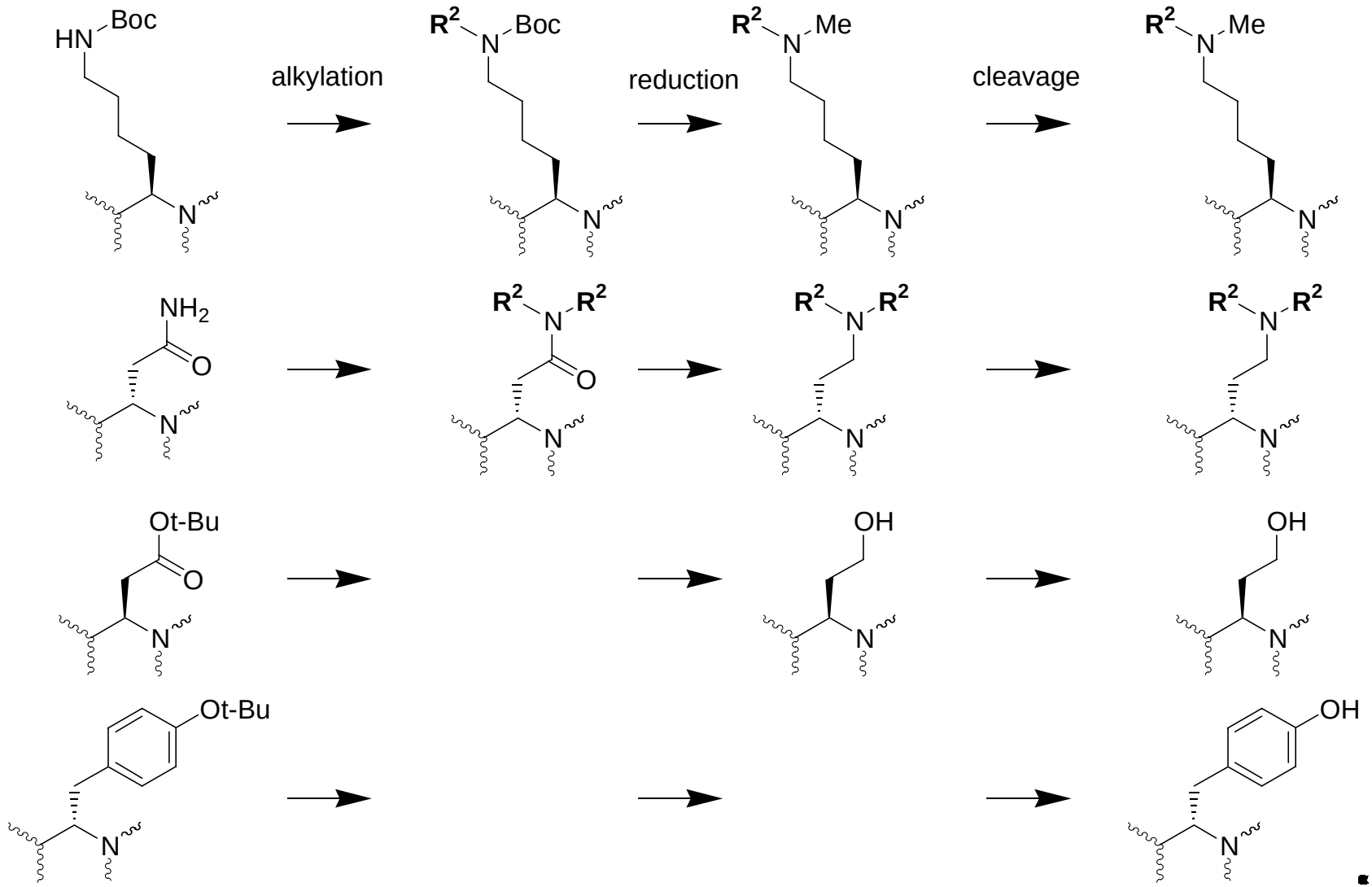
Another Enumeration Sequence



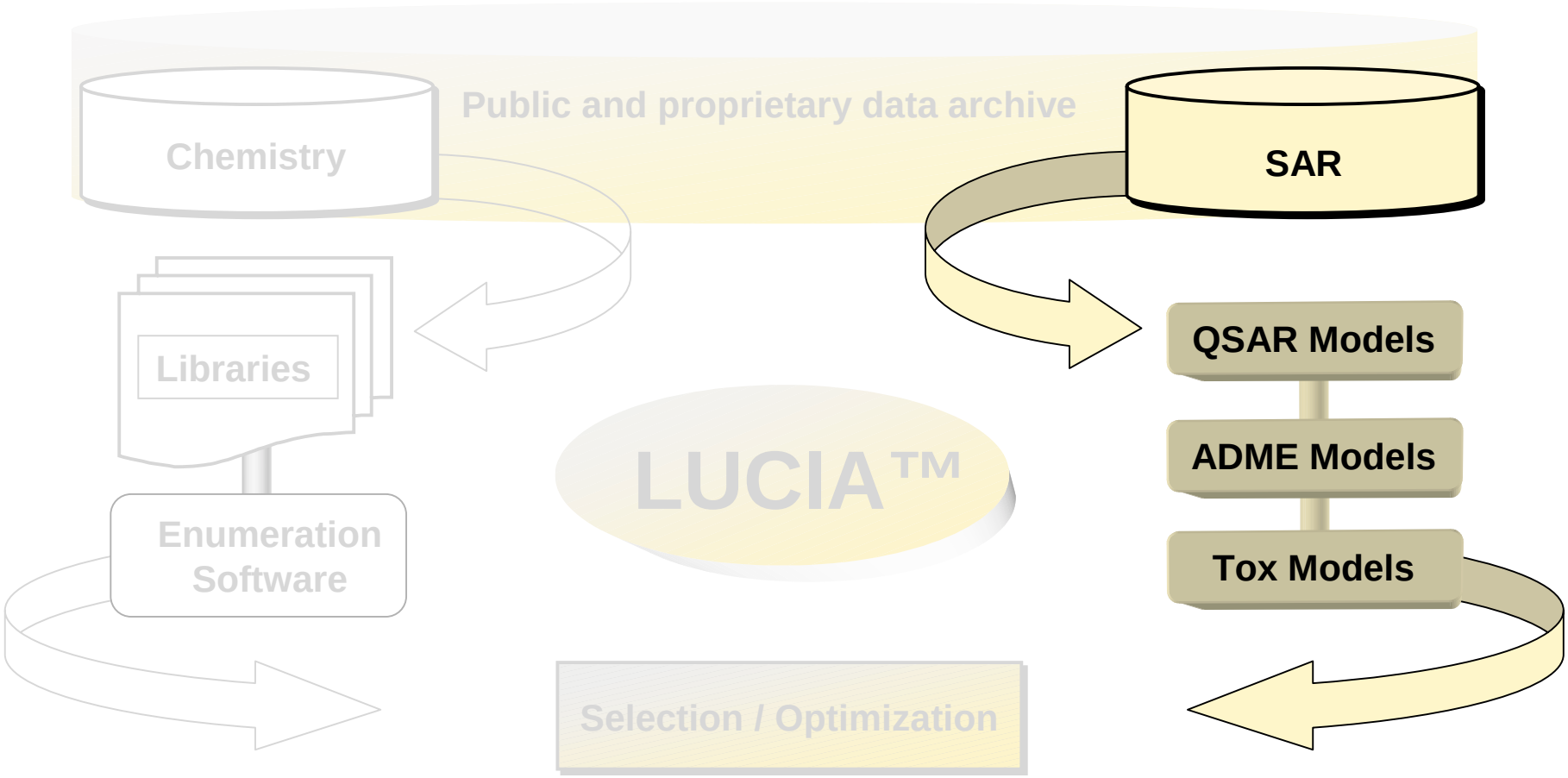
- Generic stereochemistry in reactions
- Side chain modifications in synthetic steps

Houghten et al *J. Comb. Chem.* **1999**, *1*, 195

Side-Chain Modifications



Knowledge-Based Drug Discovery



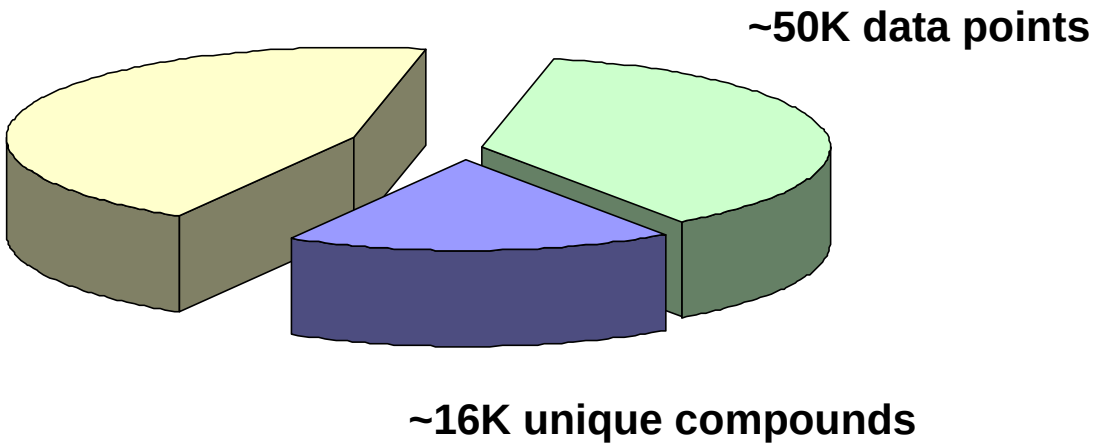
"Another recent theme is how to efficiently capture chemical data from the literature, especially in terms of constructing quantative structure-activity relationship datasets."

Chris Lipinski, DDT Jan 2003 interview

SAR Archive

- ~6K assay protocols with detailed experimental assay procedures
- ~300 kinase-related targets, categorized by mechanism
  - SAR organized by assay type, target, etc.
  - SAR grouped by binding-mode assay conditions
- Technology can capture SAR on any target / gene family

Kinase Gene Family



Integrated Computational Models

- E-screen / QSAR models
  - 12 quantitative e-screen models developed
  - 2 binary models
  - Data for 20 more models of kinase gene family targets

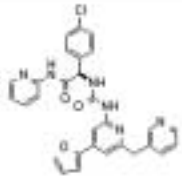
	Target	Value
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		

- Other calculated descriptors
  - Molecular properties
  - Lipinsky properties
  - ADME models

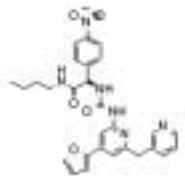
MR_ROTORS	0
MR_DONORHB	2
MR_ACCEPTHB	6.20
MR_CLOGP	5.3680
QPLOGPC16	15.4460
QPLOGPOCT	25.8730
QPLOGPW	22.24
QPLOGPO	5.5960
QPLOGS	-6.4270
STARS	3
Absorption	
BIPCACO	65.9490
AFFYPCACO	270.2650
AFFYPMDCK	215.4010
QPLOGKP	-3.2260
Distribution	
QPLOGKHSA	1.4580
Metabolism	
METABOLS	3
Blood-Brain-Barrier/CNS	
QPLOGBB	-0.2030
CNS	0
Functional groups	

- Tox

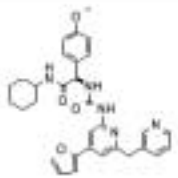
Demo Enumeration Results and Analysis



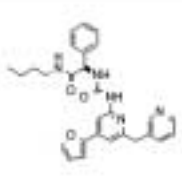
MoleculeId:1775742
MWT: 539.03
NAME:
CLOGP: 4.851
TPSA: 122.04
RotBonds: 3
HBA: 5
HBD: 3
Charge: 0
eScreen: eABL_11_7_02-pval (5.989581)
eScreen: eADK-binary (-.091224)
eScreen: eCDK1-B-pval (6.498156)
eScreen: eCDK2-A-pval (5.547401)
eScreen: eCDK2-E-pval (6.097973)
eScreen: eCDK4-D1-pval (5.241059)
eScreen: eCDK5-pval (4.322529)
eScreen: eEGFR-pval (3.990131)
eScreen: eLCK-pval (5.033617)
eScreen: eP38_alpha-pval (7.07538)
eScreen: ePKC-pval (4.742338)
eScreen: eSRC-pval (3.974903)
eScreen: eTYRK-binary (1.016216)



MoleculeId:1775745
MWT: 528.62
NAME:
CLOGP: 4.439
TPSA: 152.29
RotBonds: 3
HBA: 5
HBD: 3
Charge: 0
eScreen: eABL_11_7_02-pval (5.125718)
eScreen: eADK-binary (.030222)
eScreen: eCDK1-B-pval (6.252302)
eScreen: eCDK2-A-pval (5.592853)
eScreen: eCDK2-E-pval (5.700506)
eScreen: eCDK4-D1-pval (4.945835)
eScreen: eCDK5-pval (4.38652)
eScreen: eEGFR-pval (4.76778)
eScreen: eLCK-pval (4.663908)
eScreen: eP38_alpha-pval (7.090918)
eScreen: ePKC-pval (4.554293)
eScreen: eSRC-pval (4.368784)
eScreen: eTYRK-binary (.990398)



MoleculeId:1775747
MWT: 539.69
NAME:
CLOGP: 5.059
TPSA: 118.38
RotBonds: 3
HBA: 4
HBD: 3
Charge: 0
eScreen: eABL_11_7_02-pval (5.500013)
eScreen: eADK-binary (.015944)
eScreen: eCDK1-B-pval (6.246843)
eScreen: eCDK2-A-pval (5.257623)
eScreen: eCDK2-E-pval (5.792066)
eScreen: eCDK4-D1-pval (4.786525)
eScreen: eCDK5-pval (4.502575)
eScreen: eEGFR-pval (4.190692)
eScreen: eLCK-pval (4.742709)
eScreen: eP38_alpha-pval (7.10395)
eScreen: ePKC-pval (4.552693)
eScreen: eSRC-pval (4.417108)
eScreen: eTYRK-binary (.9665)



MoleculeId:1775753
MWT: 483.62
NAME:
CLOGP: 4.696
TPSA: 109.15
RotBonds: 3
HBA: 4
HBD: 3
Charge: 0
eScreen: eABL_11_7_02-pval (5.556639)
eScreen: eADK-binary (.008814)
eScreen: eCDK1-B-pval (5.921327)
eScreen: eCDK2-A-pval (5.268445)
eScreen: eCDK2-E-pval (5.642983)
eScreen: eCDK4-D1-pval (4.634898)
eScreen: eCDK5-pval (4.702122)
eScreen: eEGFR-pval (4.379848)
eScreen: eLCK-pval (4.687625)
eScreen: eP38_alpha-pval (7.100818)
eScreen: ePKC-pval (4.512421)
eScreen: eSRC-pval (4.088138)
eScreen: eTYRK-binary (1.020369)

- Enumerated products can be committed to the data base or exported
- Compounds are analyzed by variety of integrated models
- Results are automatically sent to the user by email with a life link to the server

Enabling Technologies

- DayCart 4.81 w/ program objects
- SMILES, SMARTS, SMIRKS
- Reaction toolkit 4.81
- Oracle 9i
- Weblogic 6.0
- Misc. tools and utilities
  - QuickProp fromSchrödinger
  - Linux RDF automapper (Infochem)
  - CSFC (SMILES depiction)
  - MDL Chime Pro plugin

Libraria's Discovery Platform

- Based upon accepted industry-standard database platform and software (Oracle/Daylight)
- Fully web-based to make technology available to bench chemists
- Scalable, open and enterprise-wide architecture
- Integrated, searchable archive of chemistry and SAR data with increasingly valuable knowledge content
- Intuitive technology for leveraging the medicinal chemist's conceptual capabilities in drug discovery

The knowledgebase-architecture is a growing repository of chemistry, biology, and derived computational models – a learning machine – that automatically develops and utilizes its predictive capability.

Acknowledgements

Thank you for listening!

Daylight Chemical Information Systems

Infochem GmbH

Wolf-Dietrich Ihlenfeldt

Peter Ertl