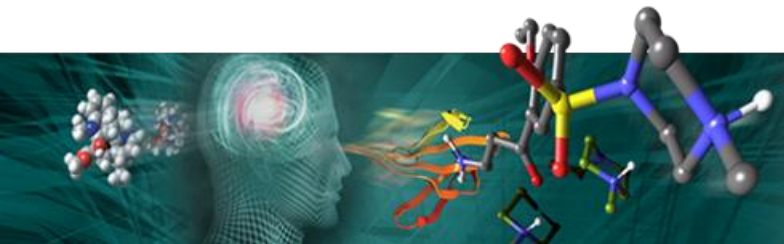
The Schrödinger KNIME extensions

Computational Chemistry and Cheminformatics in a workflow environment

Jean-Christophe Mozziconacci

Volker Eyrich



SCHRÖDINGER.

Topics

- What are the Schrödinger extensions?
- Workflow application examples
- 2012 development plan: KNIME, Maestro, Canvas, PyMOL and Seurat integration

What are the Schrödinger extensions?

Access to Schrödinger tools via KNIME

- Modelling and computational chemistry in a workflow environment
- Graphical assembly of workflows
- Generate, manipulate, analyze and visualize chemical data
- Interactive and automated analysis of datasets
- Presentation and communication of results

Access to Schrödinger tools via KNIME

Run on Linux and Windows

~150 nodes

- Ligand/protein preparation
- LBDD (cheminformatics, pharmacophore modeling)
- SBDD (Protein structure prediction, docking and scoring)
- Molecular and quantum mechanics

Molecular Mechanics

- MacroModel Single Point Energy
- MacroModel Minimization
- MacroModel Coordinate Scan
- ConfGen Standard
- ConfGen
- Conformational Search
- Conformational Search and Cluster
- Premin
- Impref
- Uffmin

Quantum Mechanics

- Jaguar Single Point Energy
- Jaguar Minimization
- NMR Shielding Constants
- Jaguar Charges

Cheminformatics

Fingerprint Based Tools

- Fingerprint Generation
- Generate Pairwise Matrix
- Generate Pairwise Matrix (2 Inputs)
- Similarity Matrix (from Molecules)
- Dissimilarity Selection (from Matrix)
- Build Report for Clustering (from Matrix)
- Hierarchical Clustering (from Matrix)

Filters and Mining Tools

- Maximum Common Substructure Search
- Substructure Search
- REOS Filter
- Structure Filter

Utilities and Converters

- Principal Components
- Multi-dimensional Scaling
- Combine Fingerprints
- Concatenate Bitvectors
- Convert Fingerprint to Bitvector
- Convert Fingerprint to Table
- Convert Matrix to Table
- Convert Table to FingerPrint
- Convert Table to Matrix
- Convert Bitvector to FingerPrint

Modeling

- Bayes Classification Model Building
- Bayes Classification Model prediction
- PLS Classification Model Building
- PLS Classification Model Prediction

Pharmacophore Modeling

- Phase Shape
- Phase DB Query
- Phase File Query
- Phase DB Creation
- Phase Hypothesis Identification

Docking and Scoring

- Glide Grid Generation
- Glide Ligand Docking
- Glide Multiple Ligand Docking
- XP Visualizer

Post-processing

- Prime MM-GBSA

Protein Structure Prediction

- BLAST
- Prime Build Homology Model
- Prime Side Chain Sampling
- Prime Minimization

Schrödinger nodes

Workflows

Protein Preparation

- Protein Preparation Wizard
- Protein Assignment

Induced-fit docking

- IFD and individual steps

Ligand Preparation

- LigPrep
- Ligprep individual tools (Ionizer, Desalter, Neutralizer…)
- Epik

Property Generation

- QikProp
- Molecular Descriptors
- Calculate properties

Filtering

- Ligfilter
- Ligparse
- Property Filter (Propfilter)

Scripting

- **Run Maestro Command**
- **Chemistry External Tool** 0:1, 1:0, 1:1, 1:2, 2:1 and 2:2 nodes
- **Python Script** 0:1, 1:0, 1:1, 1:2, 2:1 and 2:2 nodes

Desmond

- System builder
- Molecular Dynamics
- Trajectory extract frames and manipulation
- Trajectory reader, CMS reader

Reporting

- Run Maestro
- Run Canvas
- View CSV (open xls/ooffice)
- Text Viewer

Tools

Combinatorial Libraries

- CombiGlide Library Enumeration
- CombiGlide Reagent Preparation

Fragments

- Fragment Joiner
- Fragments from Molecules

Data Manipulation

- **Compare Ligands**
- Lookup and Add Columns
- **Group MAE**
- Ungroup MAE

Structure Manipulation

- Add Hydrogens
- **Delete Atoms**
- **Split by Structure**
- **MAE Parser**
- Extract MAE Properties
- Delete MAE Properties
- Set MAE Properties
- Set Molecule Title
- Set MAE Index

Utilities

- Get PDB
- Align Binding Sites
- Protein Structure Alignment
- Prime Fix
- **RMSD**
- Assign Bond Orders
- Unique Title Check
- PDB Name
- SD Format Checker
- Generate Smarts
- Unique Smiles
- Entropy Calculation
- RRHO Entropy
- Boltzmann Population
- Volume Overlap Matrix

Readers/Writers

- CSV Reader
- Molecule Reader
- SD, PDB, Mol2 Reader nodes
- Sequence Reader
- Alignment Reader
- Fingerprint Reader
- Hypothesis Reader
- Glide Grid Reader
- Glide Multiple Grid Reader
- Variable Based Glide Grid Reader
- Molecule Writer
- Sequence Writer
- Alignment Writer
- Hypothesis Writer
- Fingerprint Writer

Converters

- Molecule-to-MAE
- MAE-to-Pdb, to-SD, to-Smiles and to-Mol2
- SD-to-Smiles
- PoseViewer-to-Complexes
- Complexes-to-PoseViewers
- String-to-Type
- Hartree-to-kcal/mol Converter
- kJ-to-kcal Converter

Schrodinger Suite 2011- some new features

Infrastructure

- Canvas **2D renderer on 64 bit** KNIME workbench, **Windows** installers
- New **start-up script options** to pass user/machine/OS-specific parameters (memory limit, temporary directory…)

New nodes

- **Desmond** nodes (System builder and Molecular dynamics, Trajectory manipulation, extract frames, Desmond trajectory and CMS Readers)
- **Phase query on files** (already possible on a Phase database)
- **Canvas model** building and prediction nodes (PLS and Bayes classification)
- **Run Canvas**
- Filter structures based on an ASL expression, Calculate properties

Options and functionalities

- Prime Build homology model- include ligand and cofactors
- Run Maestro- import as groups, read Hypotheses, specify a project to which structures are added

…

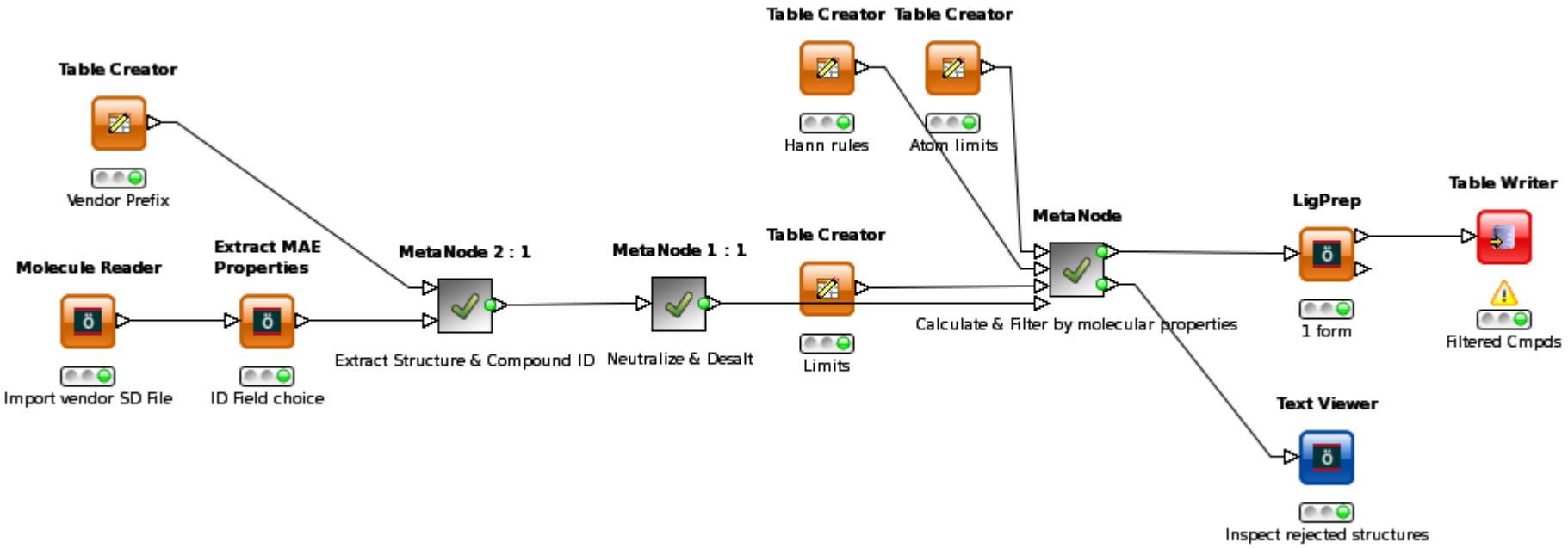
SCHRÖDINGER.

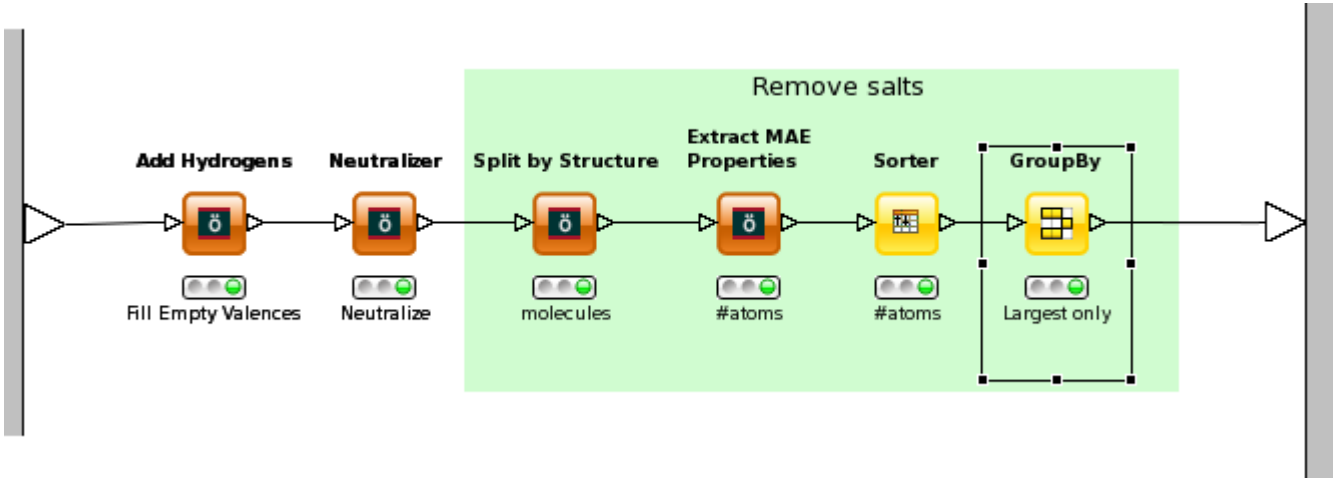
KNIME workflow page- http://www.schrodinger.com/knimeworkflows/

- **Cheminformatics**
  - Cluster by fingerprint
  - **Database analysis**
  - **Maximum Common Substructure**
  - Select diverse molecules
  - Similarity search
  - **Substructure Search**
- **Docking and post-processing**
  - **Protein preparation and Glide grid generation** ◄
  - Docking and scoring
  - Ensemble docking
  - Induced Fit Docking protocol ◄
  - **Loop over docking parameters**
  - **Validate docking parameters**
  - **Virtual screening**
- **Pharmacophore modeling**
  - **Phase Shape screening**
  - Phase database screening ◄
  - Phase hypothesis identification
- **Molecular Mechanics**
  - **Compare conformational search methods**
- **Quantum mechanics**
  - **Conformational search and QM optimization Using the Report designer**
- **Homology modeling**
  - Model building and refinement ◄
- **KNIME workbench**
  - Group by use-cases
- **General tools**
  - **Python script node use-cases**
  - **Chemistry external tool node use-cases**
  - Ensure molecule title uniqueness
  - Group Looper
  - Output column structure option philosophy
  - Run maestro command node use-cases
  - Protein structure alignment

Simplest, **most exciting**, **new** and improved ◄ workflows

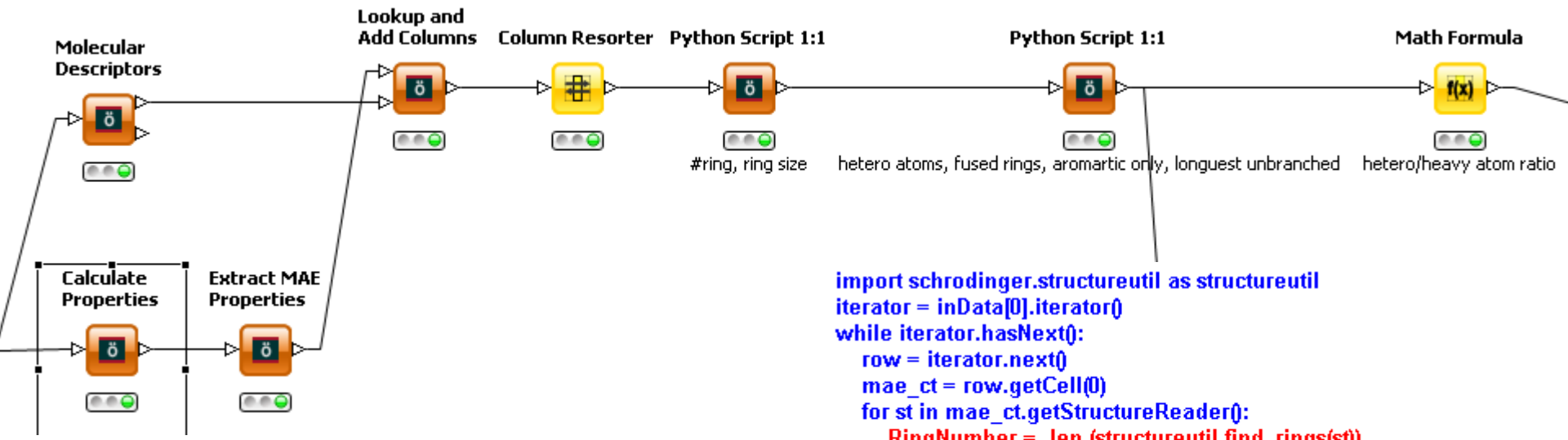
Vendor database preparation



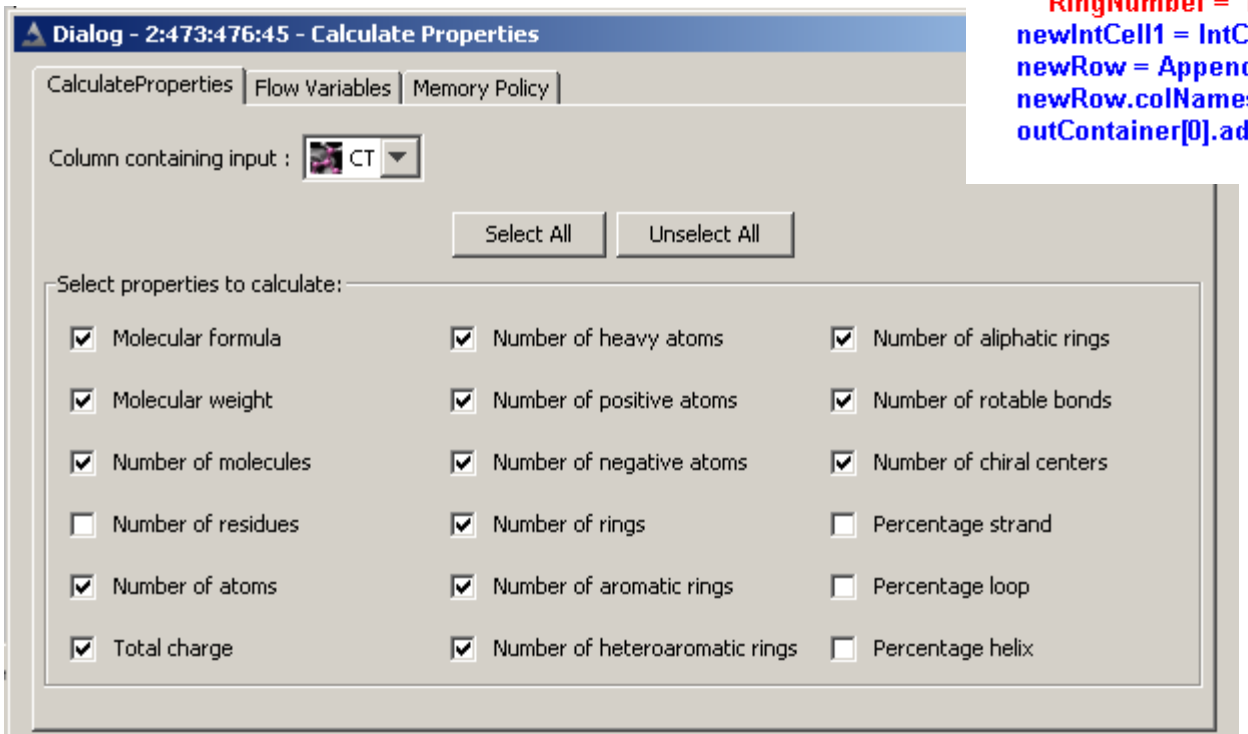


SCHRÖDINGER.

Vendor database preparation- descriptor generation

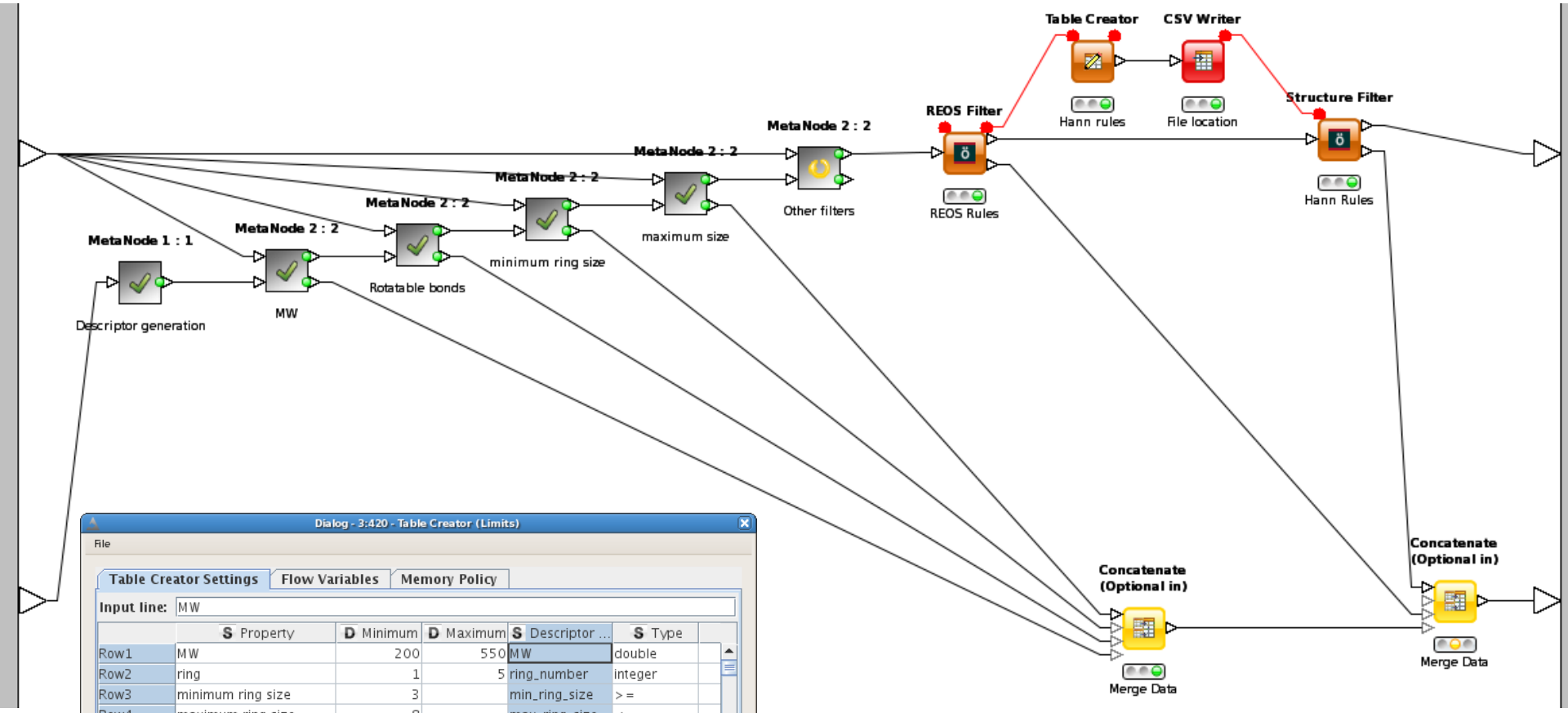


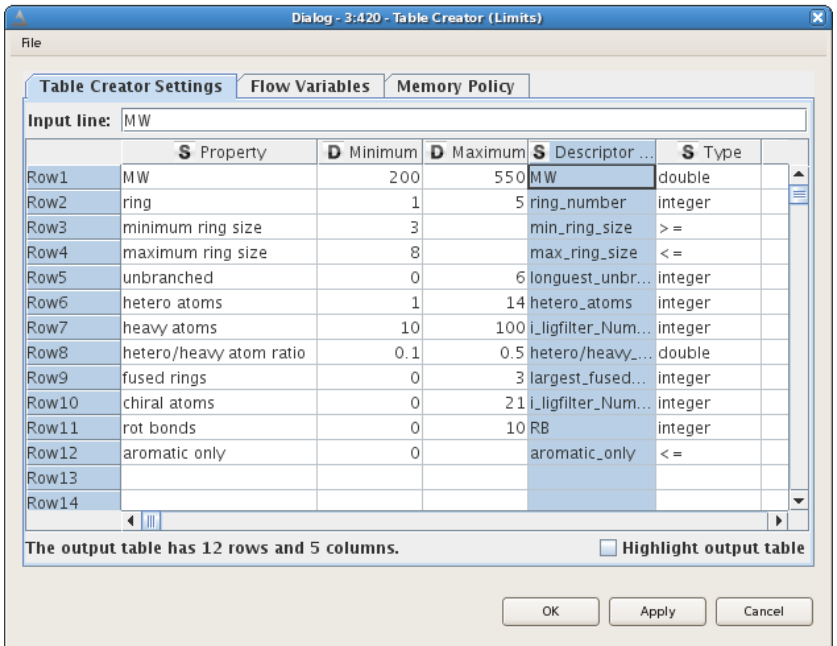

```
import schrodinger.structureutil as structureutil
iterator = inData[0].iterator()
while iterator.hasNext():
    row = iterator.next()
    mae_ct = row.getCell(0)
    for st in mae_ct.getStructureReader():
        RingNumber =  len (structureutil.find_rings(st))
    newIntCell1 = IntCell(RingNumber)
    newRow = AppendedColumnRow(row, [newIntCell1])
    newRow.colNames = [ "ring_number"]
    outContainer[0].addRowToTable(newRow)
```

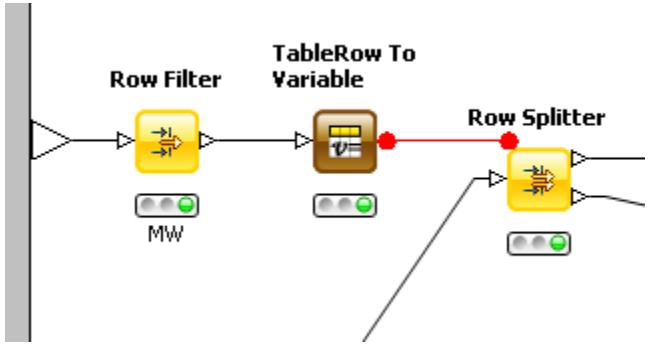


SCHRÖDINGER.

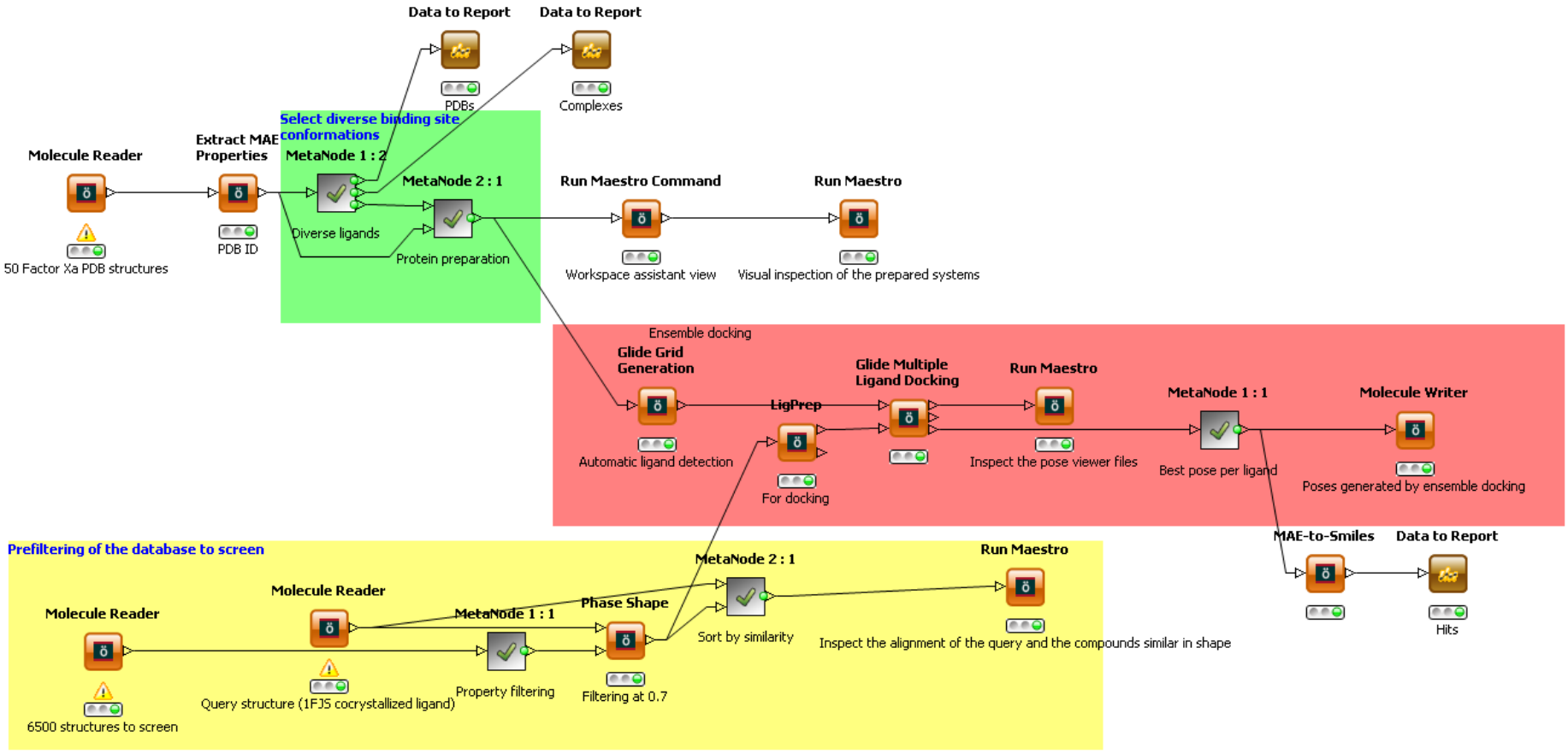
Vendor database preparation- filtering on various criteria





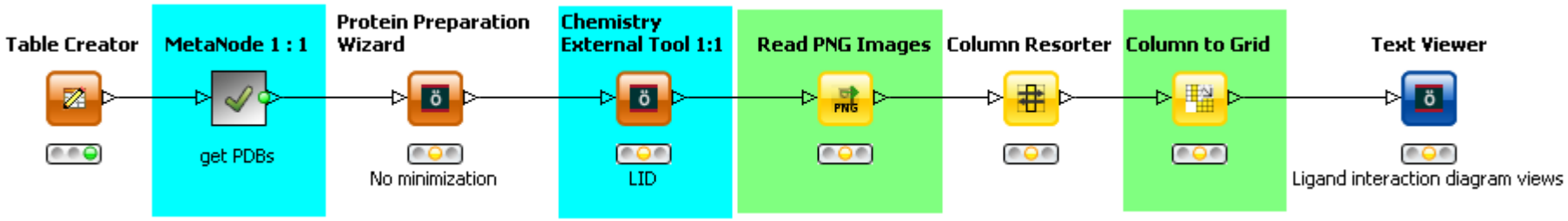


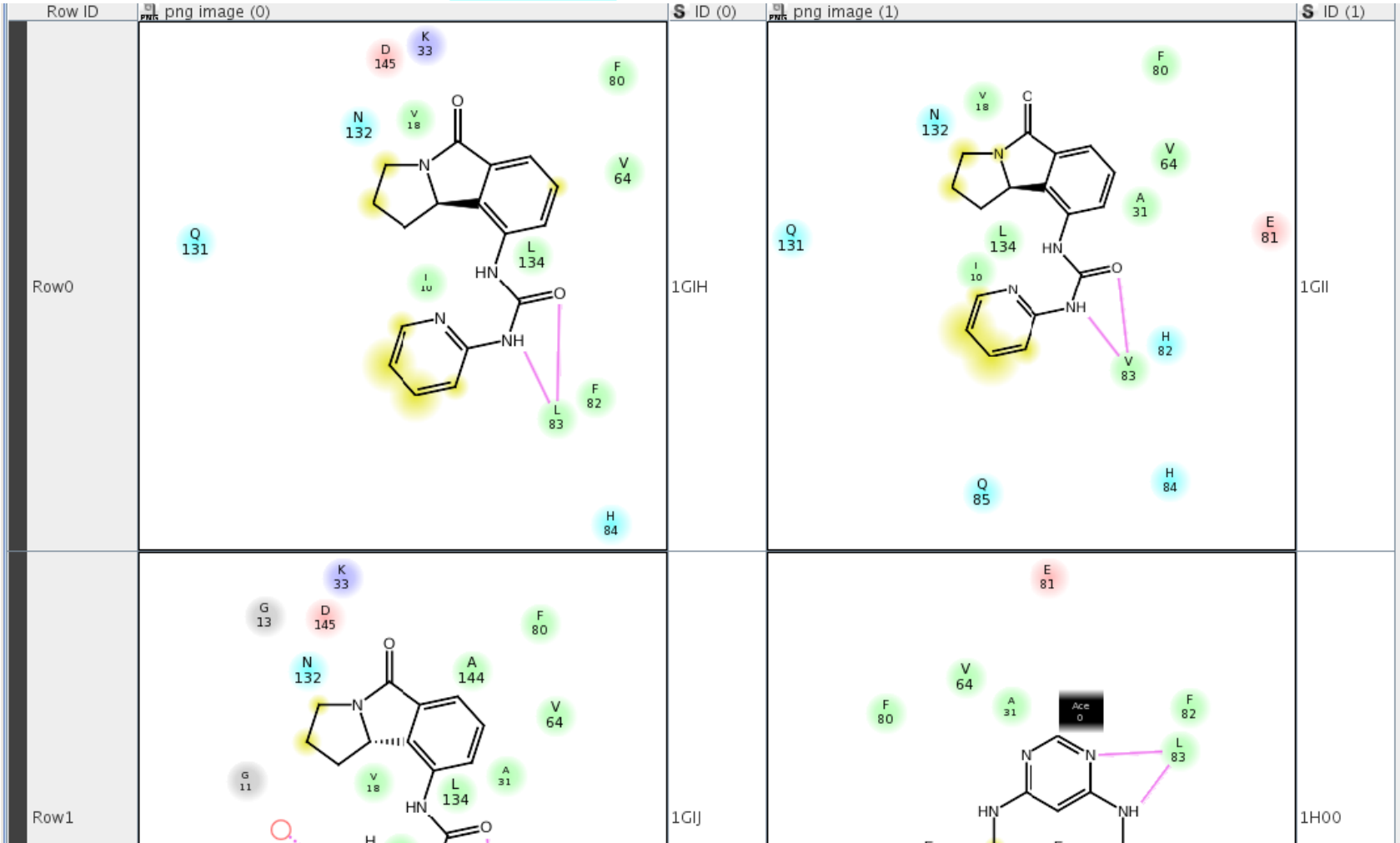
Database shape screening and ensemble docking



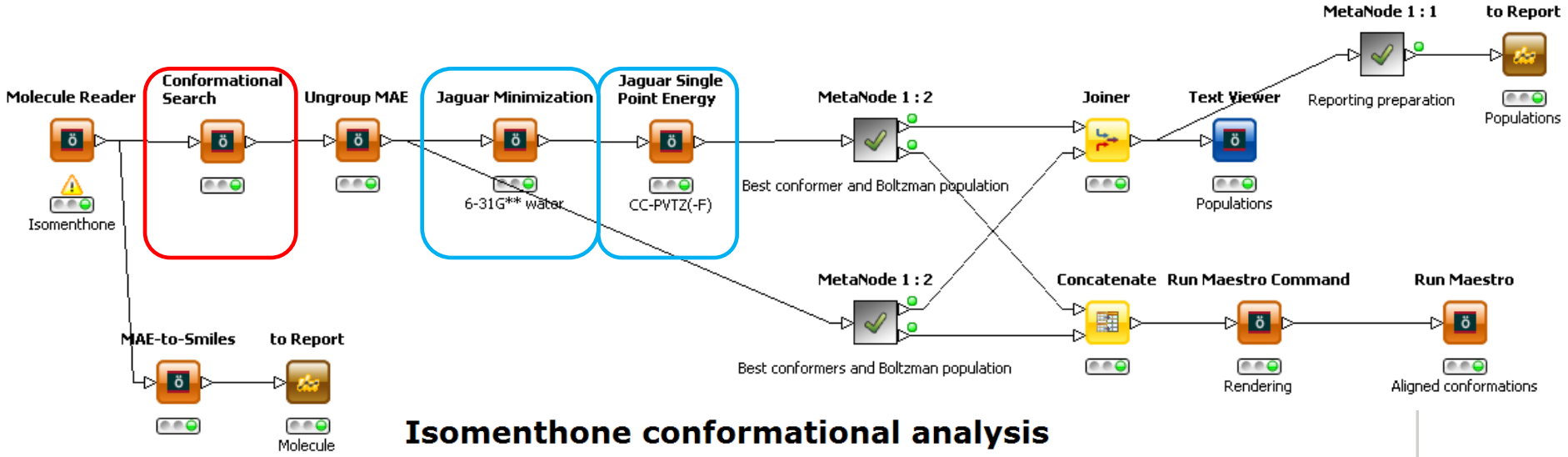
SCHRÖDINGER.

Column to Grid, Read PNG images



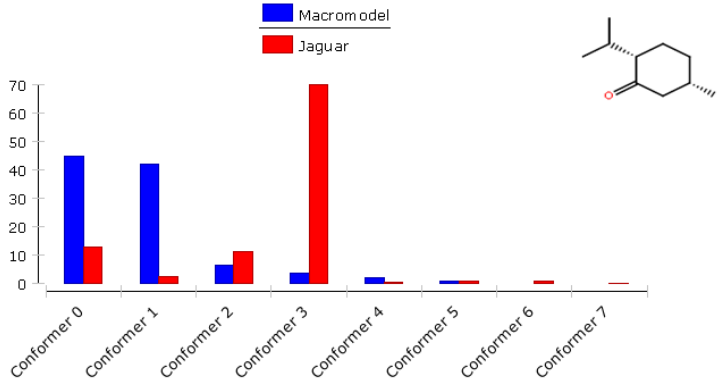


Conformational search and QM optimization



Isomenthone conformational analysis

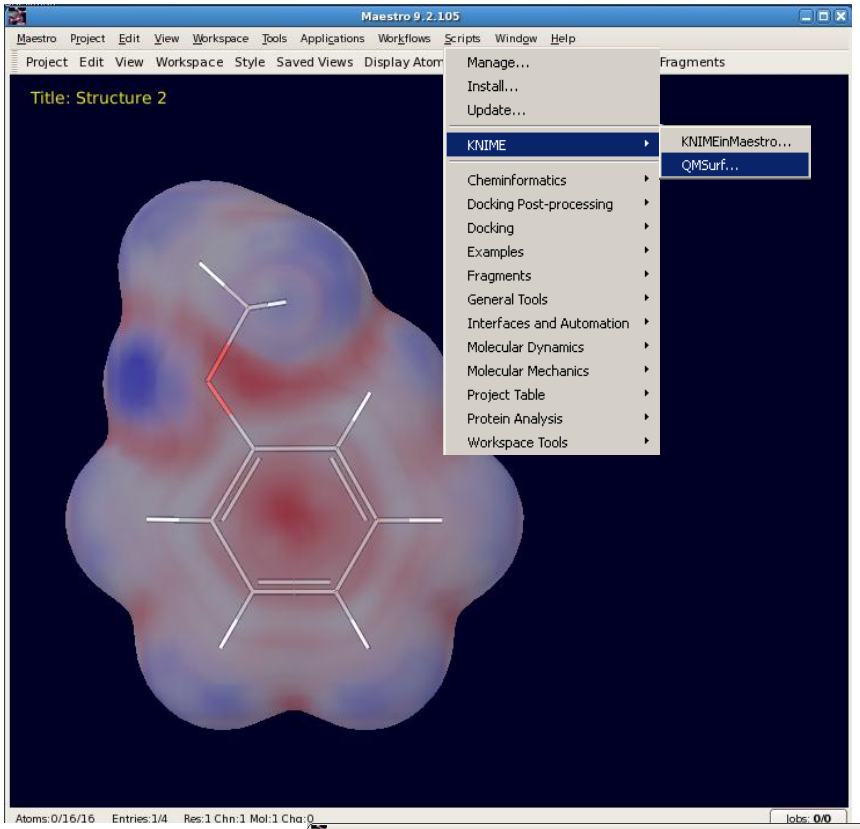
Conformer populations

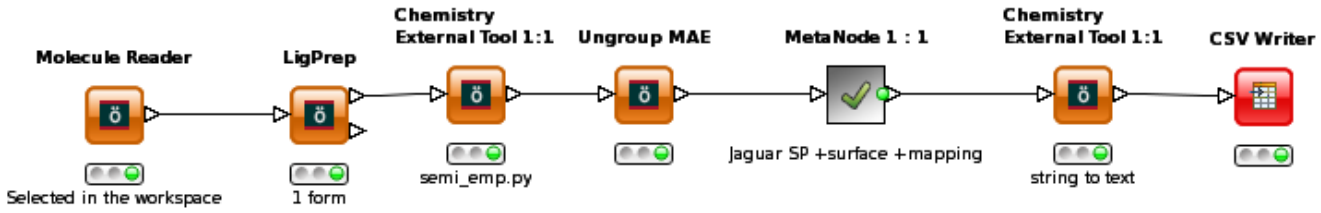


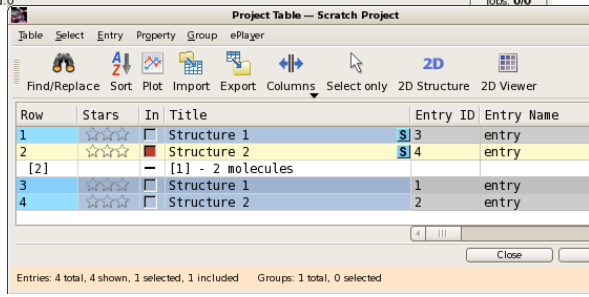
	Jaguar			MacroModel		
Row ID	Relative Energy	Ring conformation	Boltzmann population	Relative Energy	Ring conformation	Boltzmann population
Conformer 3	**0.0**	**-104.2**	**69.9**	6.32	-104.3	3.6
Conformer 4	11.65	-80.0	0.7	7.67	-85.9	2.1
Conformer 5	10.29	-68.9	1.1	9.9	-82.7	0.8
Conformer 7	15.96	-18.4	0.1	18.55	-20.4	0.0
Conformer 6	10.55	-18.3	1.0	17.95	-21.3	0.0
Conformer 0	4.17	-1.1	13.1	**0.0**	**-5.3**	**44.8**
Conformer 1	8.09	1.3	2.7	**0.15**	**-3.1**	**42.2**
Conformer 2	4.52	6.5	11.4	4.82	1.3	6.5

KNIME, Maestro, PyMOL, Canvas and Seurat integration

- Launch Maestro, Canvas and PyMOL from KNIME
  - Interactive visualization
  - Run Maestro command scripts
- Launch KNIME from Maestro
  - Through a simple python script
  - Dedicated menu in 2012
- Tighter integration in 2012 including Seurat







Schrodinger KNIME extensions

Jean-Christophe **KNIME Associate Product Manager**

jcmozzic@schrodinger.com

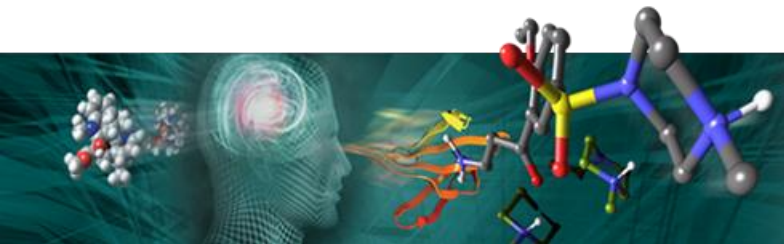
Volker Eyrich **VP of Technology**

volker@schrodinger.com

Schrödinger Developers

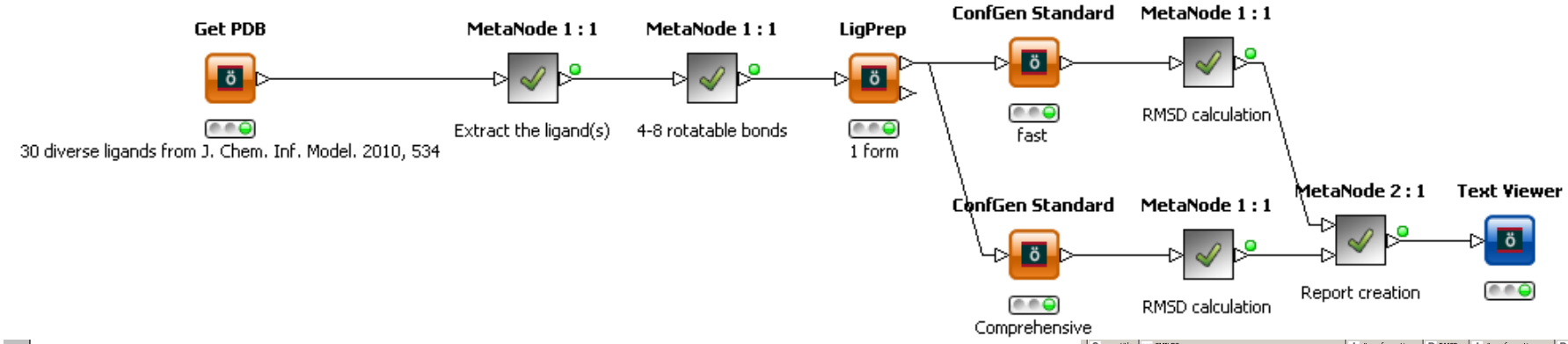
Woody Sherman **VP of Applications Science**

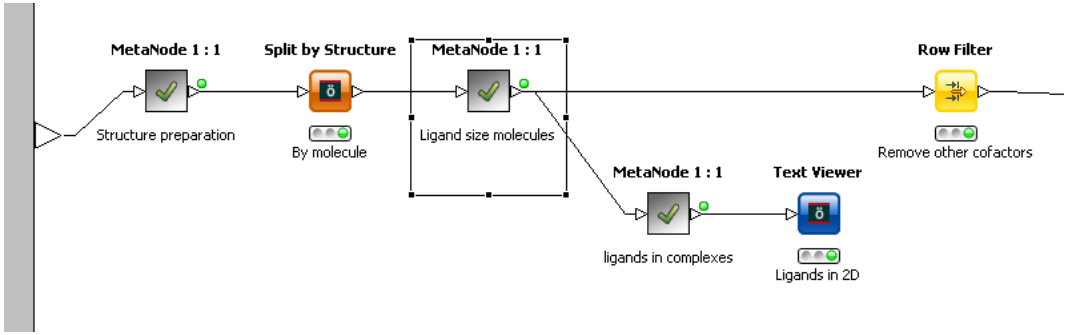
Applications scientists

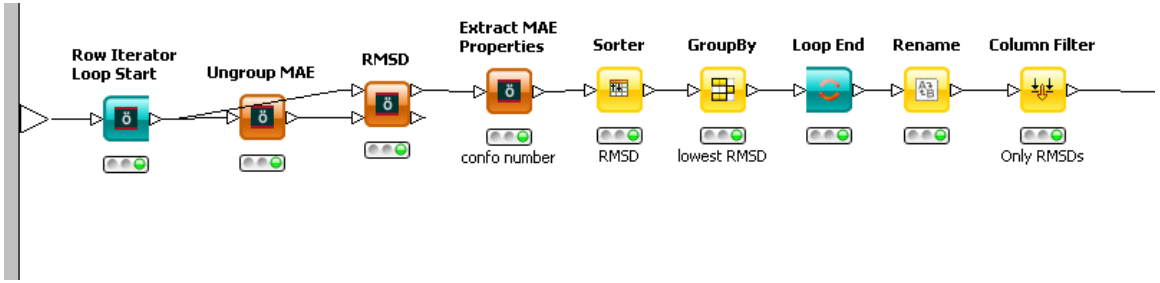


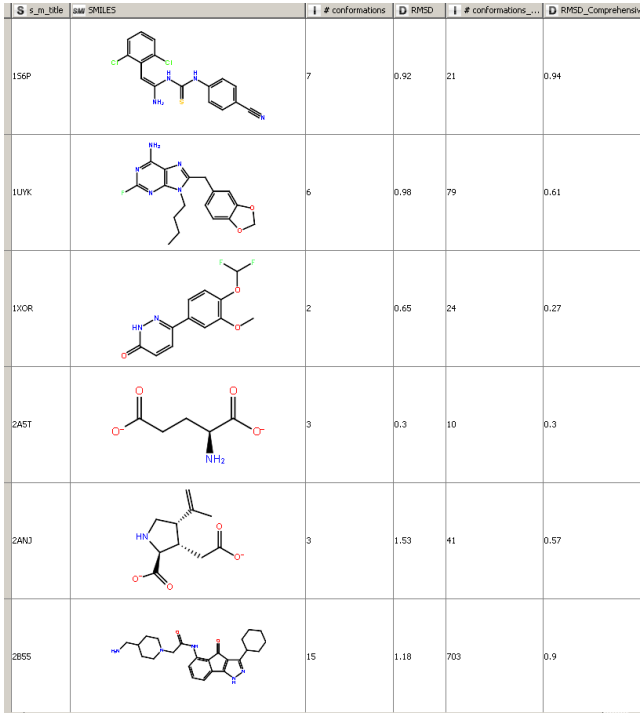
SCHRÖDINGER.

Conformational search- compare methods









SCHRÖDINGER.