Integrated Cheminformatics to Guide Drug Discovery

Matthew Segall, Ed Champness, Peter Hunt, Tamsin Mansley

CINF Drug Discovery Cheminformatics Approaches

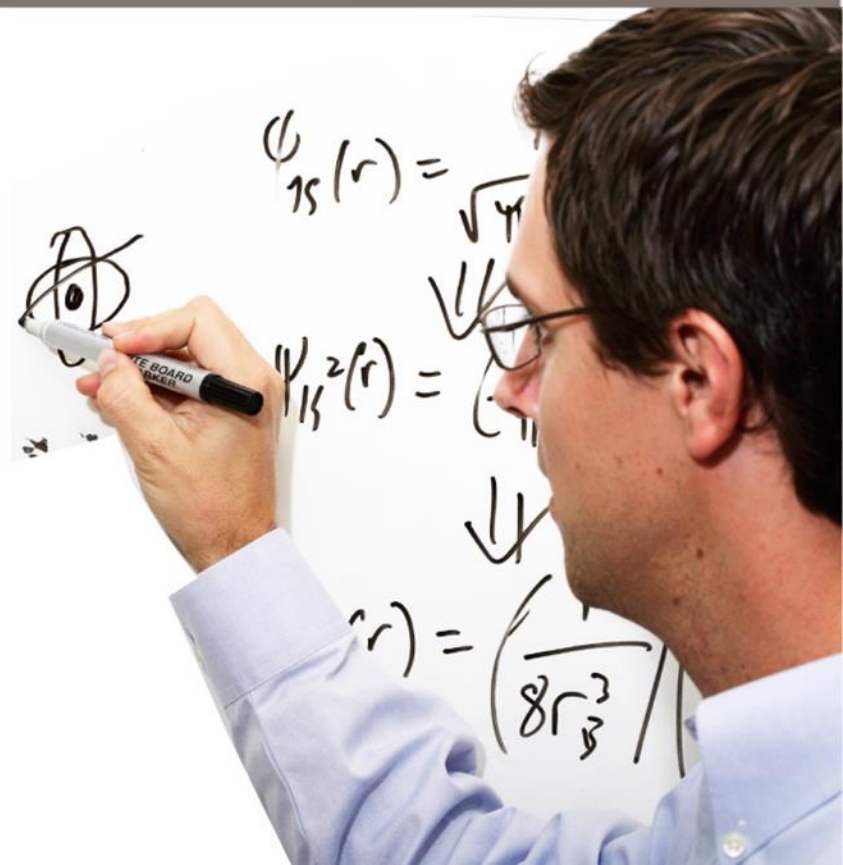
August 23rd 2017

Overview

- The impact of fragmentation of cheminformatics tools
- Challenges of integration
  - Examples: Data access and docking/alignment
- Illustrative application
- Conclusions

Fragmentation

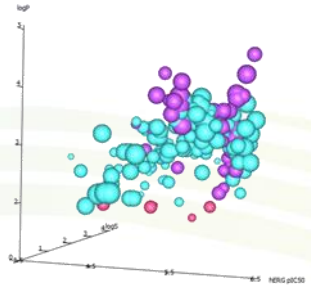
optibrium



Fragmentation of Cheminformatics Tools

- Many capabilities are required to drive drug discovery projects
  - Database access
  - Visualisation and analysis of data
  - 2D predictive modelling (QSAR)
  - 3D structure-based design
  - …
- Disparate tools for these features create bottlenecks and inefficiencies
  - Reformatting of data to move between applications
    - Time consuming and lost information
  - Different user interfaces
    - Training burden
  - ‘Expert’ tools can be impenetrable to non-computational chemists
    - Support from experts even for mundane tasks
    - Delay and distraction of experts from adding scientific value

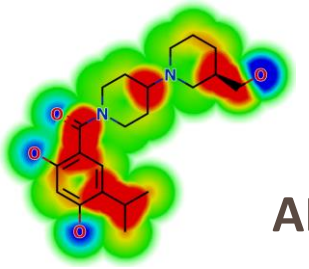
Common Situation



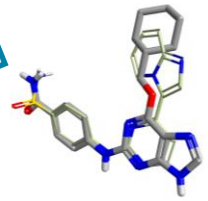
Visualisation

Medicinal Chemist

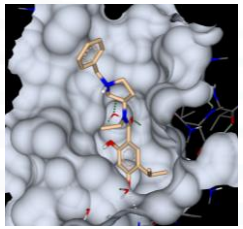
Computational Chemist



ADMET predictions



Ligand-based design



Structure-based design

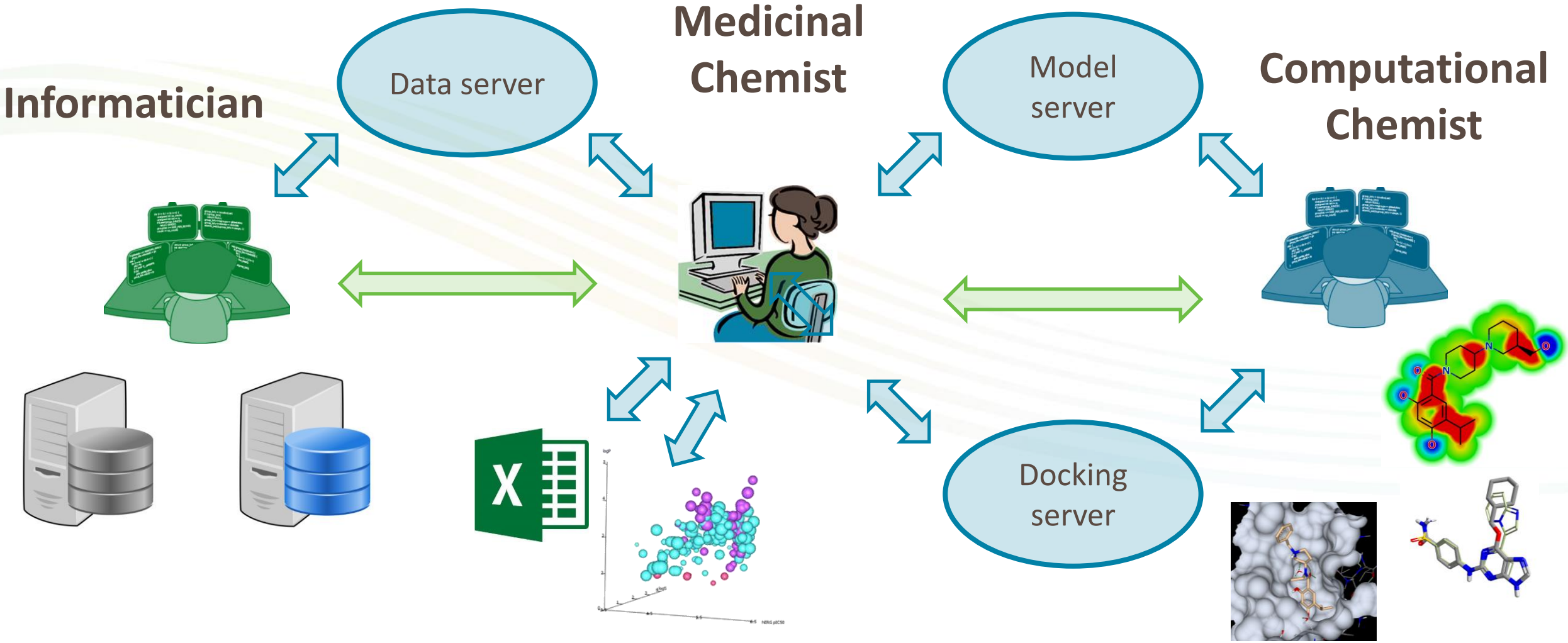
Excel

Informatician

Database(s)
(Internal and external)

Pipelines

Better Situation

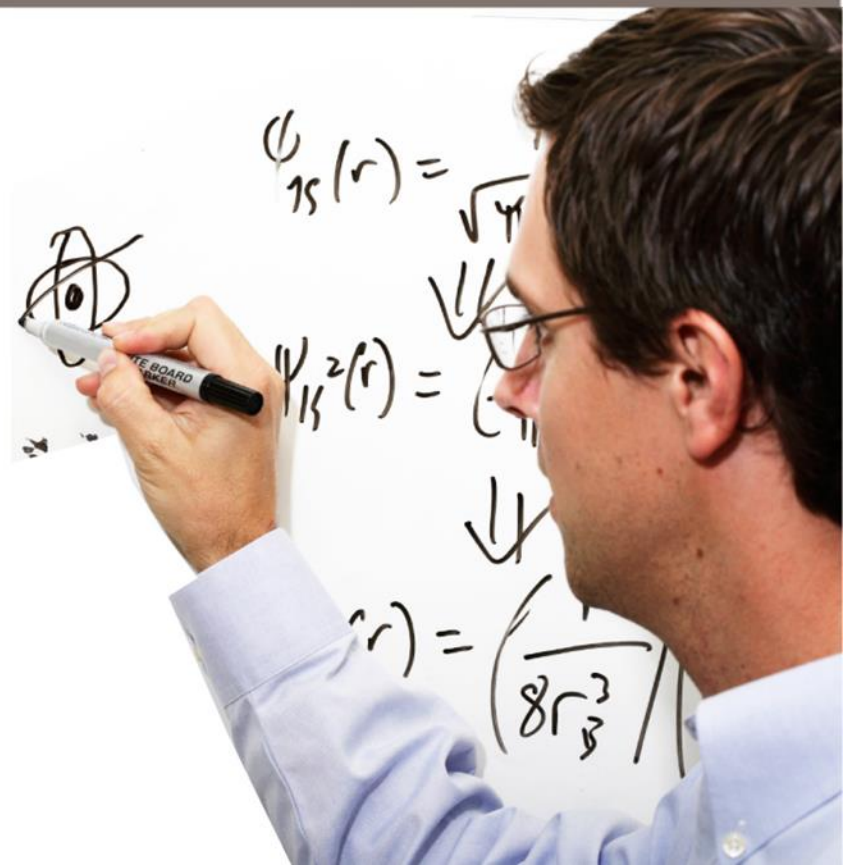


Advantages

- Project leader can have all information available in one place
  - (Ideally) single, common interface
  - Can get instant feedback on new design ideas
- Computational and data experts do not waste time on routine, mundane calculations
  - Focus on scientifically interesting and valuable activities
- Encourages closer collaboration
  - Uninteresting ideas are quickly eliminated
  - Collaborations can focus on high-value ideas for more detailed investigation

Integration

optibrium



The Challenges of Integration

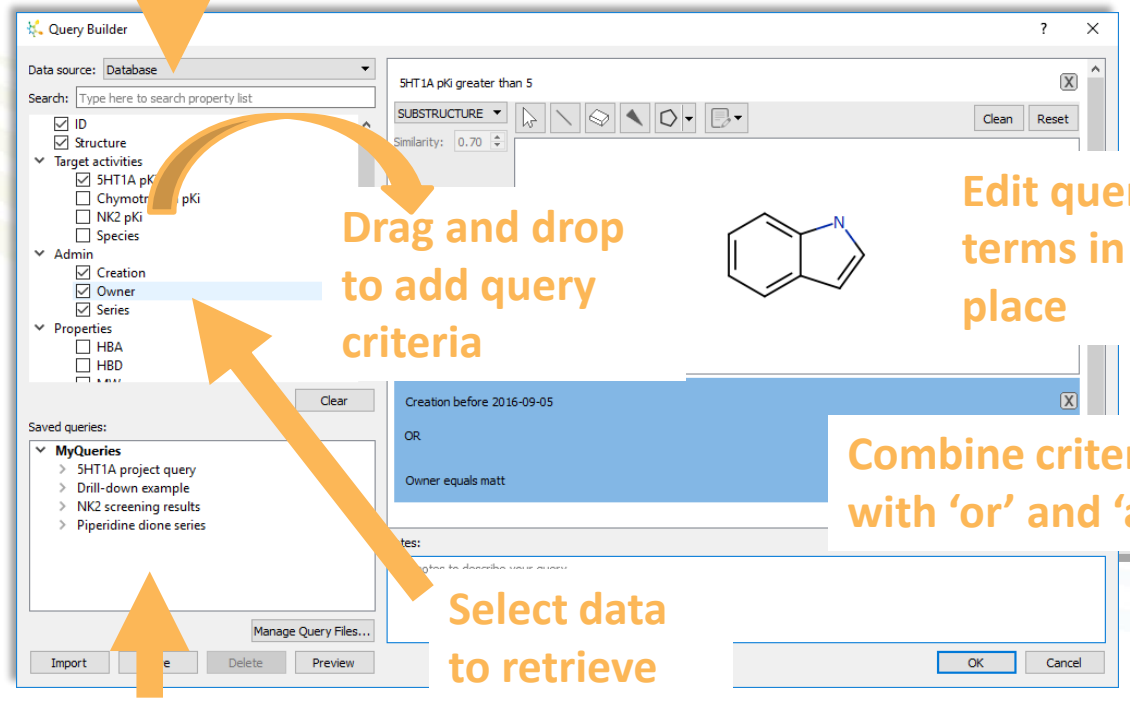
- Different software preferences for many tasks
  - Database/ELN providers
  - ‘Favourite’ docking or alignment platforms
  - Must be agnostic to the source of data and predictions
- Different architectures
  - Web, web services, desktop, command line…
- Interaction is key
  - Stimulates new ideas and strategies
  - Black boxes are not trusted
- Must be user-friendly and intuitive
  - Low barrier to use
  - ‘Gluing’ lots of different software together leads to a poor user experience

Example: Database access
Challenge

- Provide access to multiple data sources, e.g.
  - SQL via ODBC
  - Flat files
  - Web services…
- User-friendly definition of search criteria and fields
  - Support for criteria based on chemical structure, numerical, date, textual and categorical fields
- Save, share, edit and execute pre-defined queries
  - Stored individually or by project
- Support for multiple data aggregation levels
  - Drill-down to underlying data

Example: Database access
Solution

StarDrop™



Select data source

Drag and drop to add query criteria

Edit query terms in place

Select data to retrieve

Load and save queries

Combine criteria with ‘or’ and ‘and’

Flat files
e.g. SD, CSV

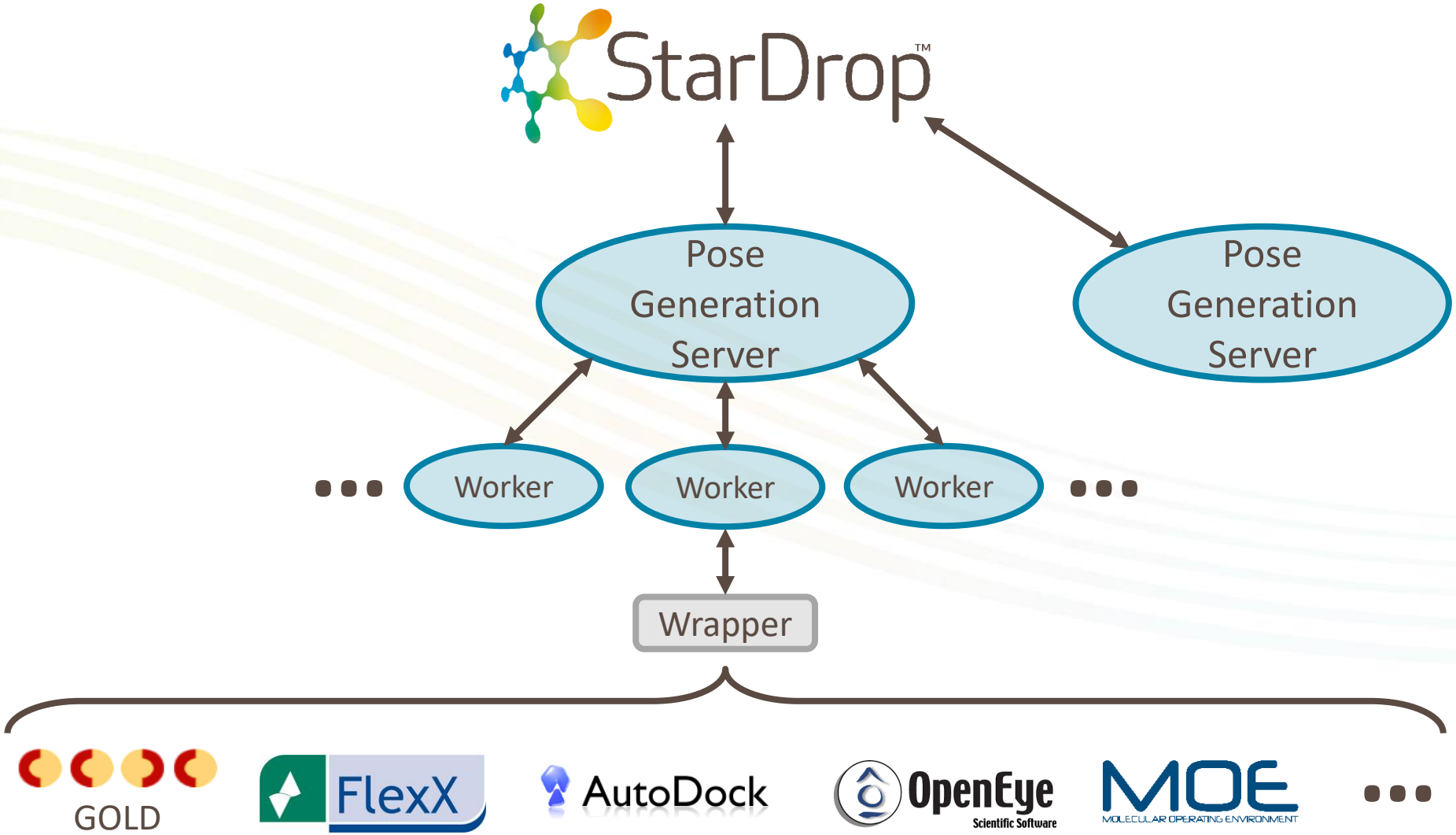
Simple databases
e.g. Oracle, MySQL

Custom data architectures

Example: Docking and Alignment
Challenge

- Enable non-expert users to run docking, pharmacophore or conformation generation calculations
  - Quick feedback on new compound designs
  - Link information from 3D models with other analyses and data
- Support for all major 3D modelling platforms
  - Flexibility to run models using preferred methods
- Expert computational chemists can easily publish new 3D models
  - Configuration files for new targets ‘drop in’ on server
- Server to enable queuing and batch processing of long calculations

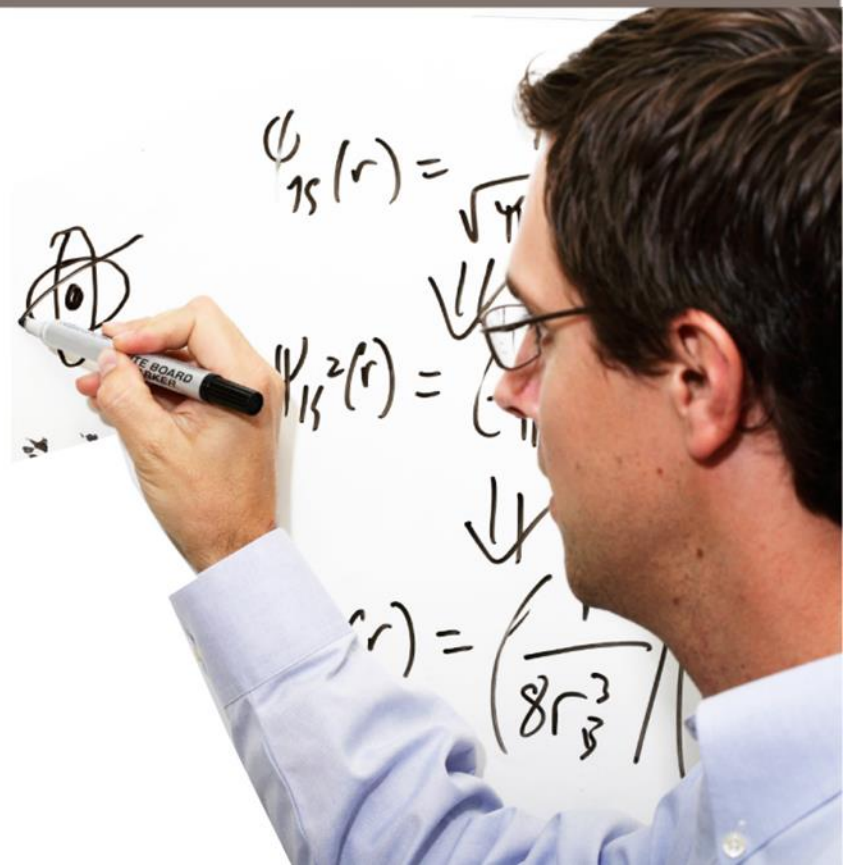
Example: Docking and Alignment
Solution



http://www.optibrium.com/community/videos/introduction-to-stardrop-modules-and-features/375-pgi

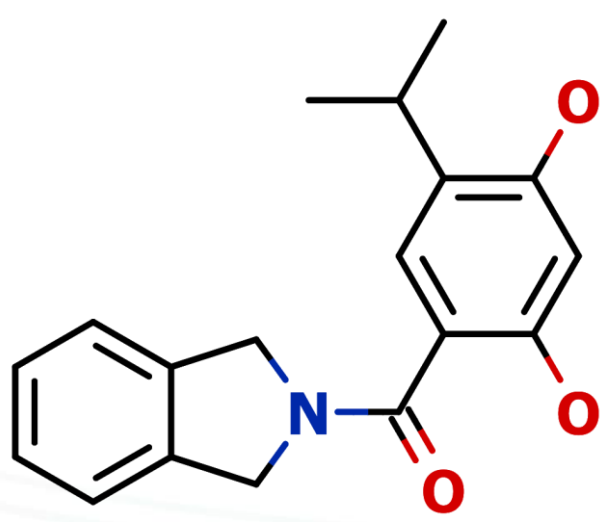
Example Application

optibrium



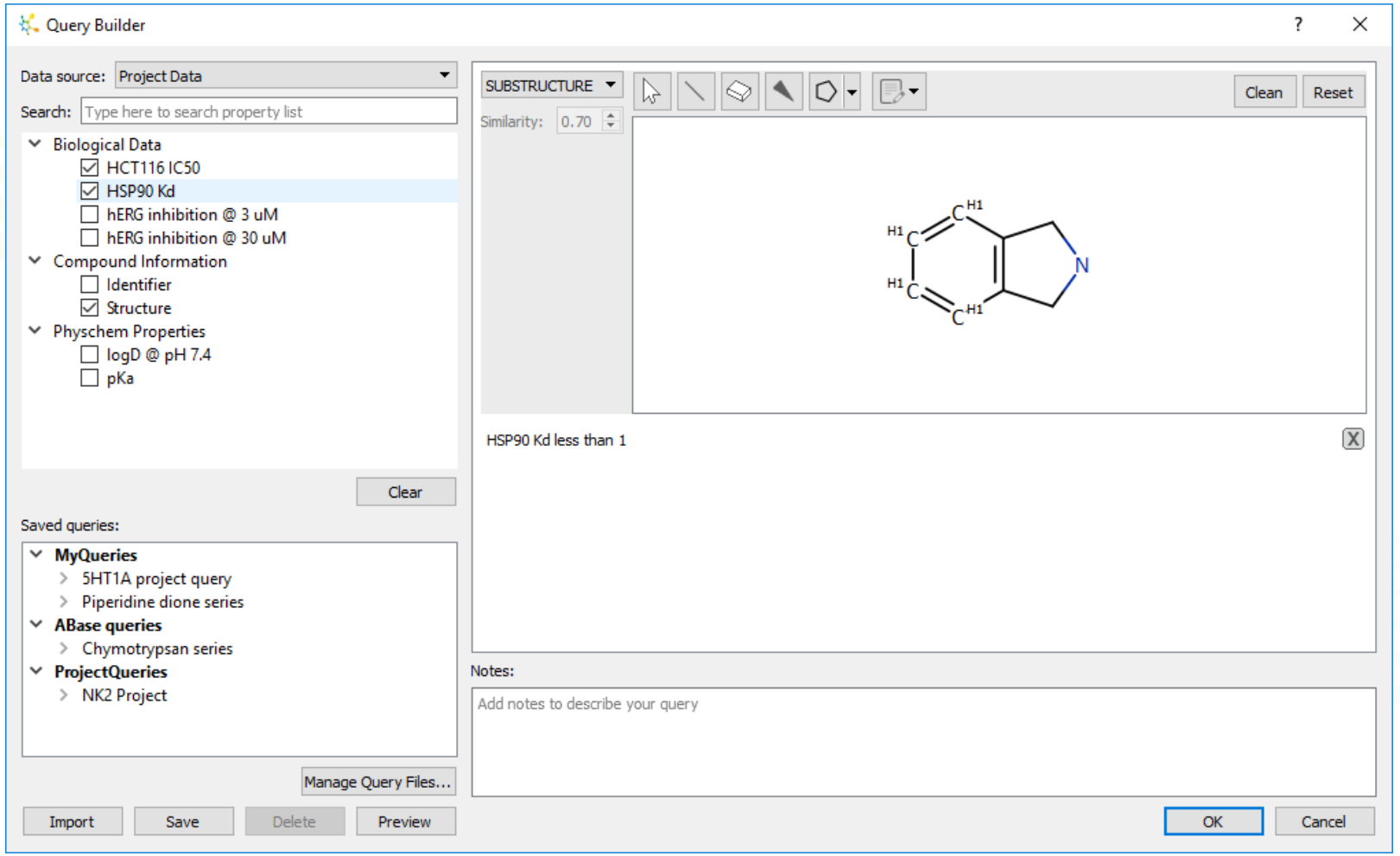
Background

- Target: Heat-shock protein 90 (HSP90)
- Initial hits based on fragment based drug discovery
  - PDB: 2XAB
  - SAR explored around resorcinol in isoindoline resorcinol series
- Objectives
  - Explore substitutions and replacements for isoindoline
  - Identify high quality compounds for oncology target
  - Decide on synthetic strategy
- Based on initial SAR from Woodhead *et al.* J. Med. Chem. 53 p. 5956 (2010)
  - N.B. This is not the process used in this project; structures and data used for illustrative purposes only

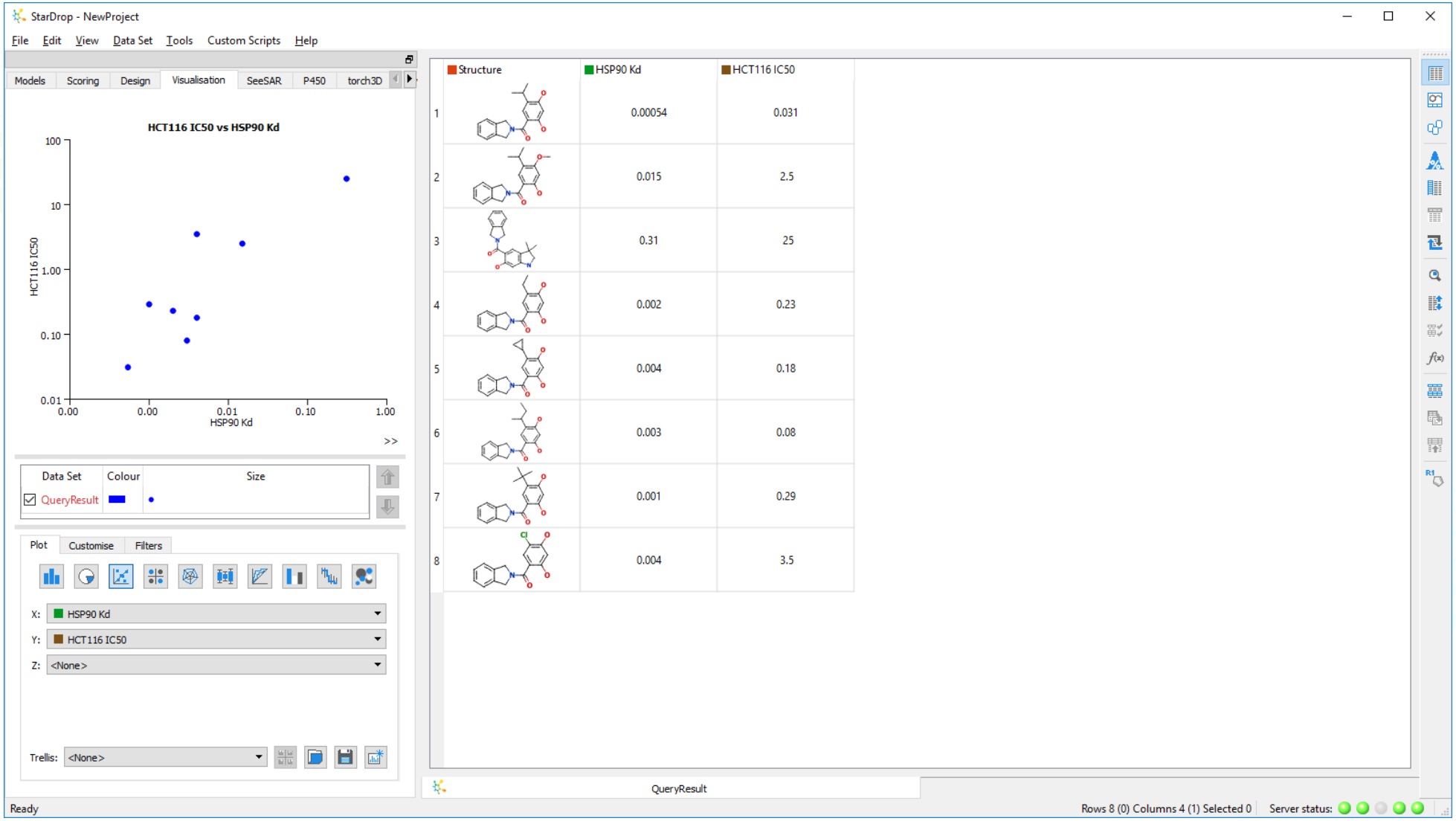


Retrieve Data for Latest Hits

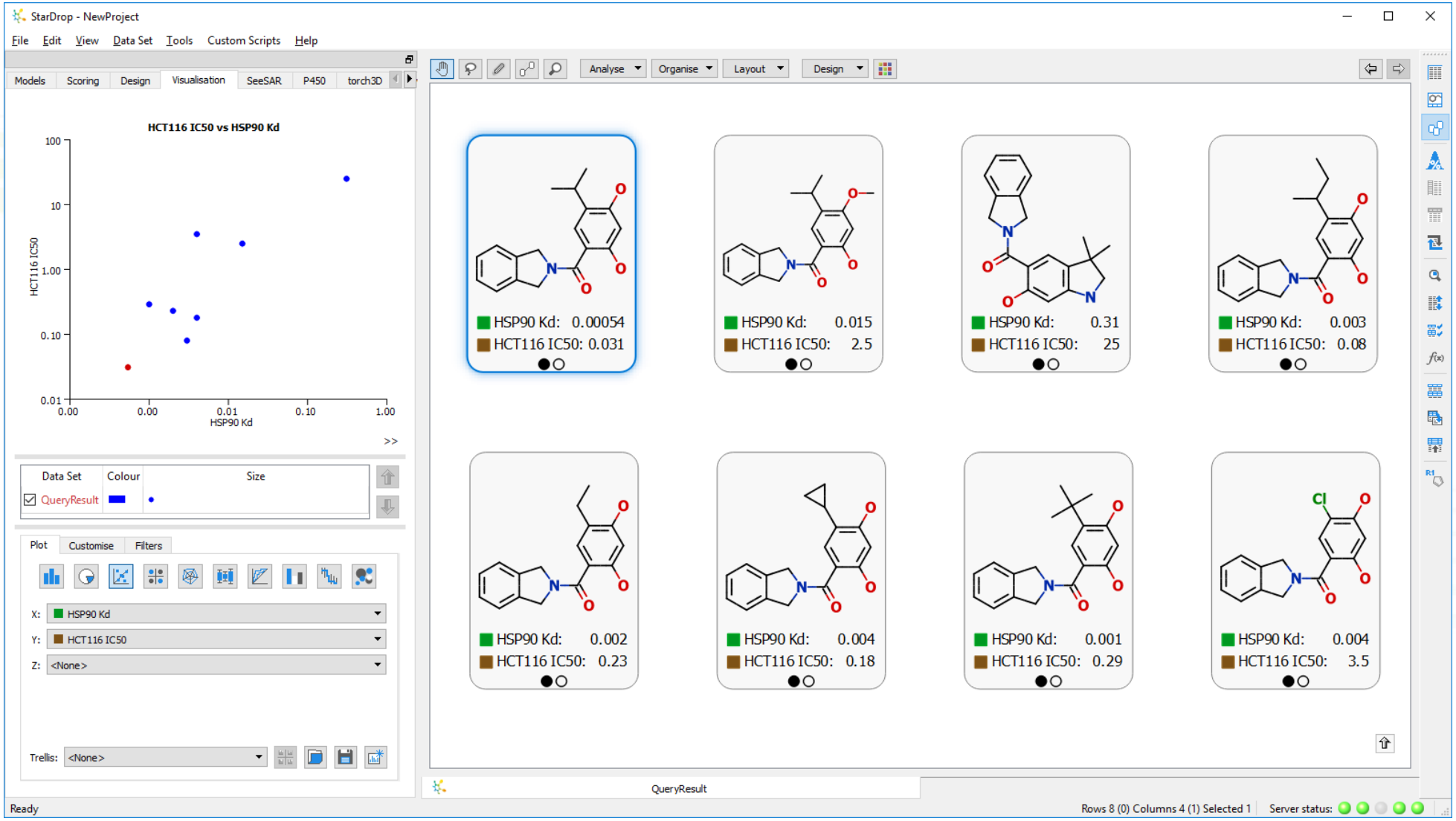
Unsubstituted isoindolines



Initial SAR

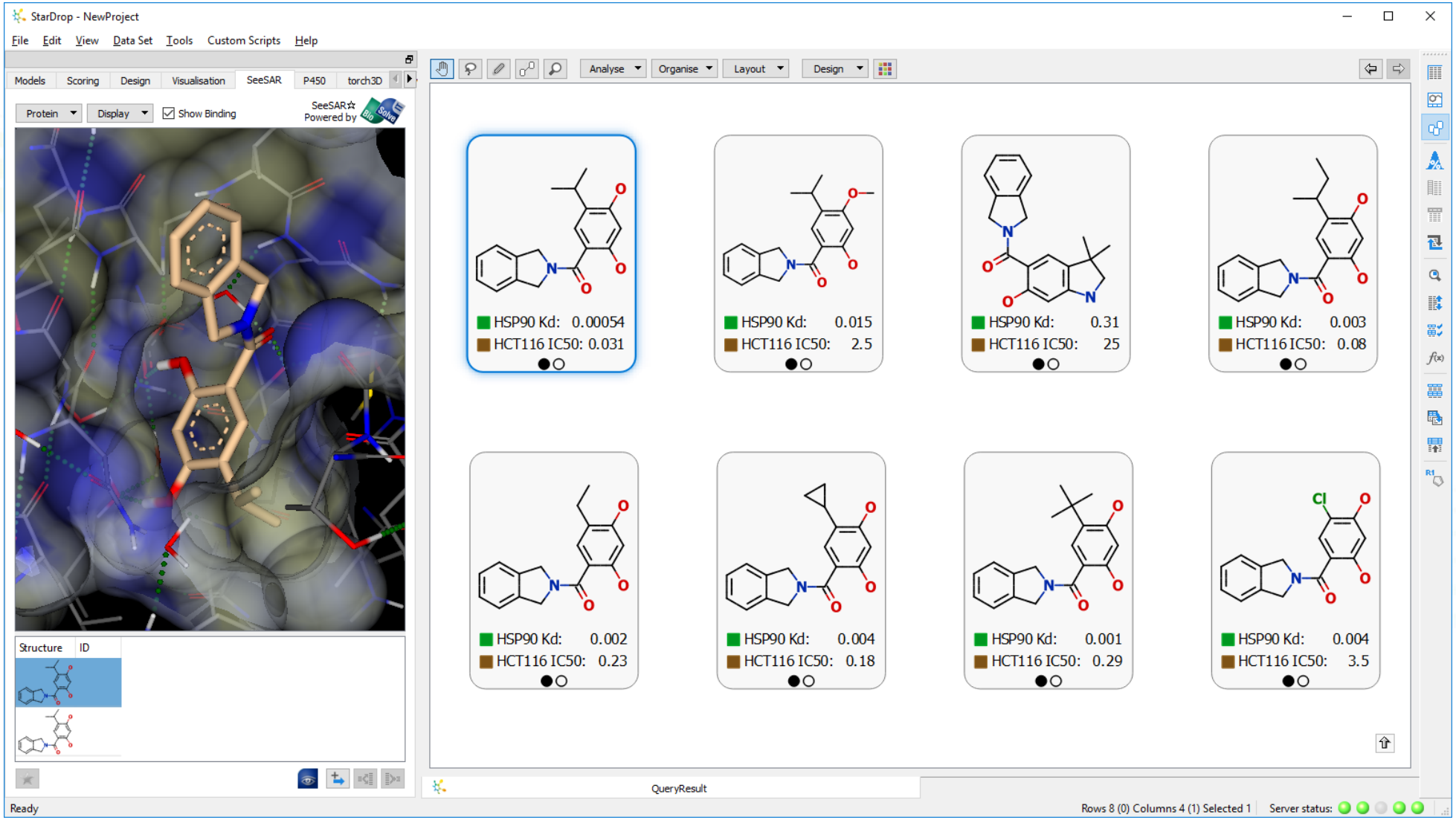


Initial SAR



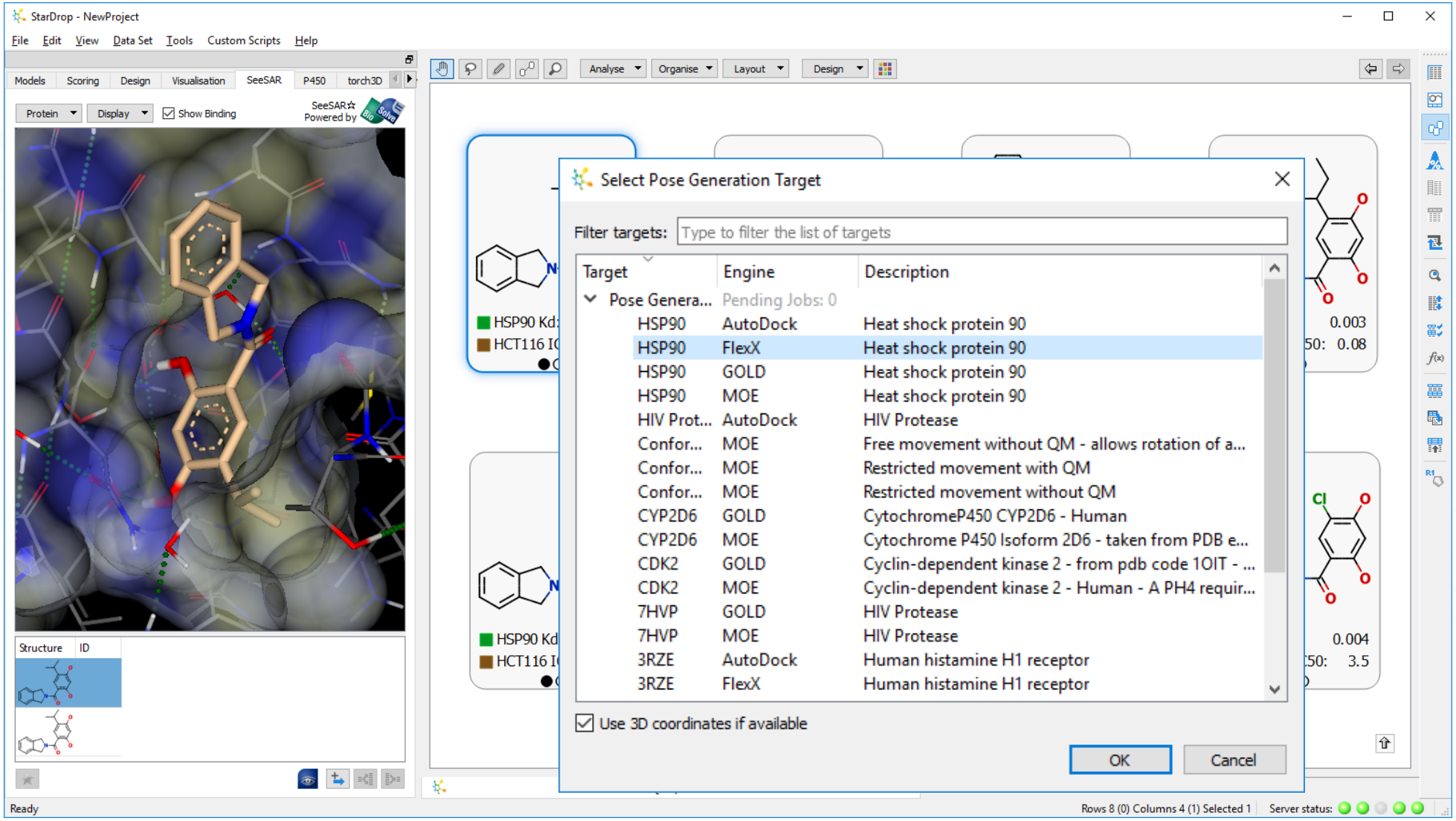
M.D. Segall *et al.* (2015) Drug Discov. Today **20**(9) pp. 1093-1103

Crystal Structure
PDB 2XAB

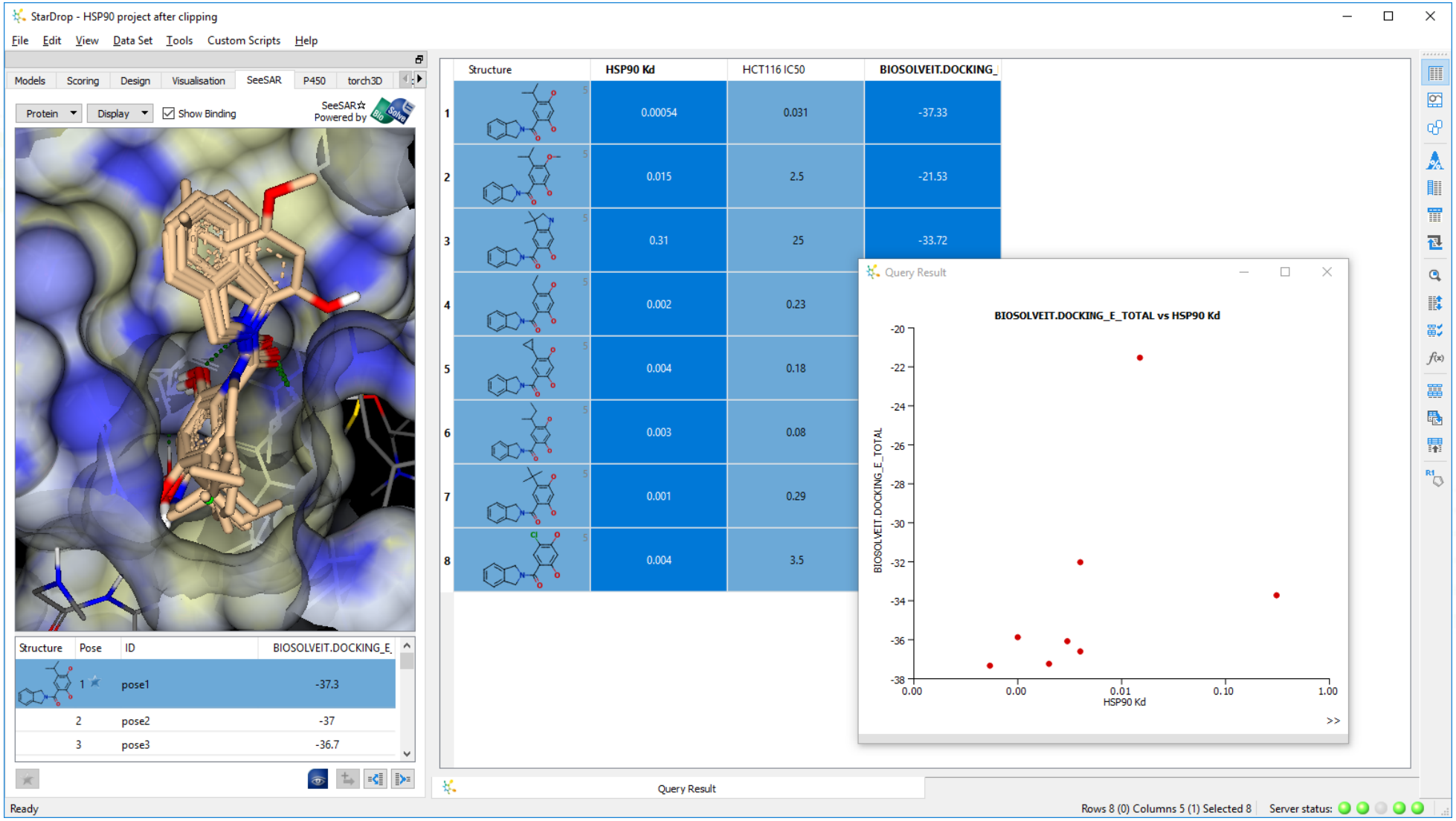


M.D. Segall *et al.* (2015) Drug Discov. Today **20**(9) pp. 1093-1103

Dock with HSP90 Model

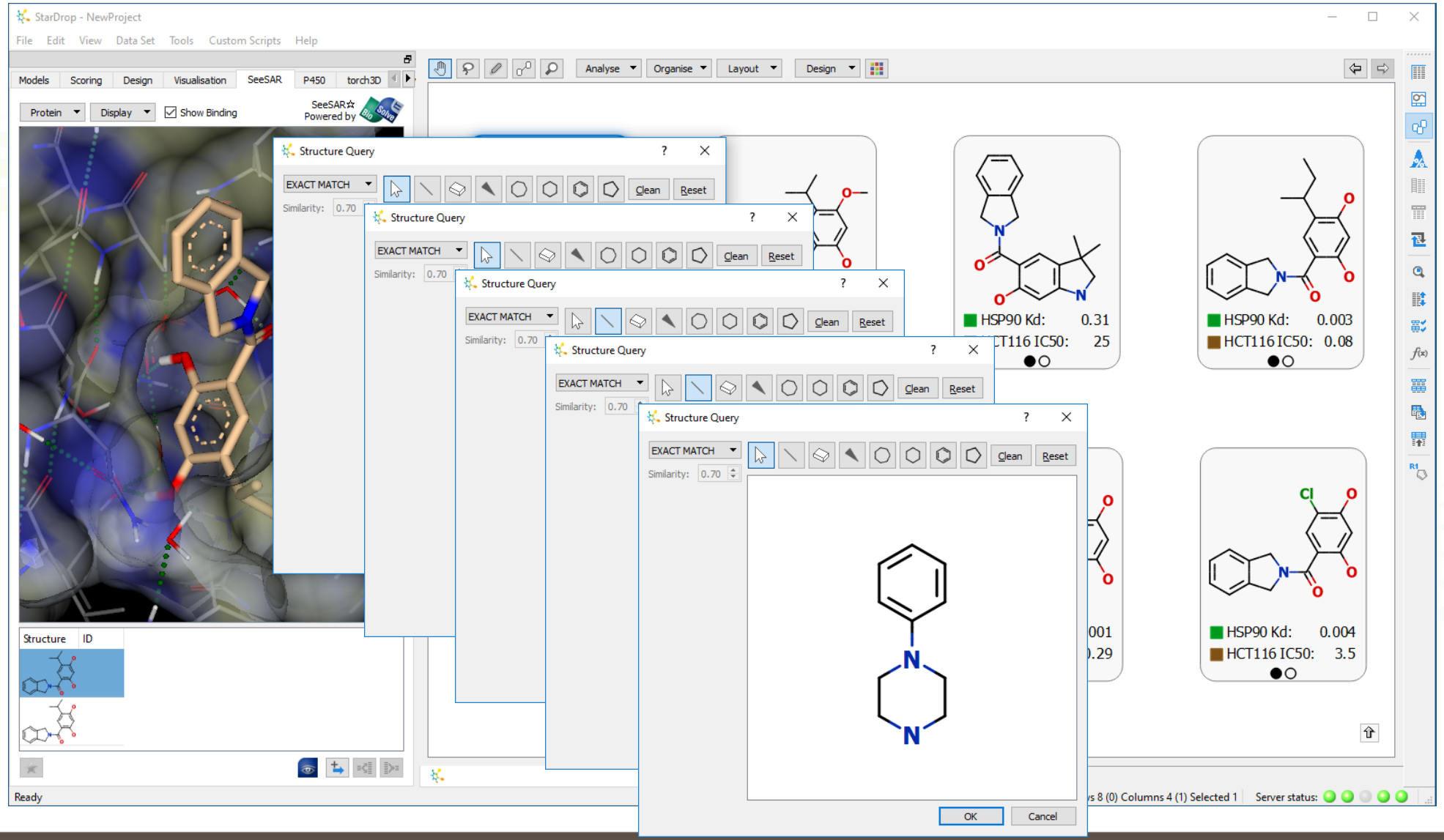


HSP90 Docking Results

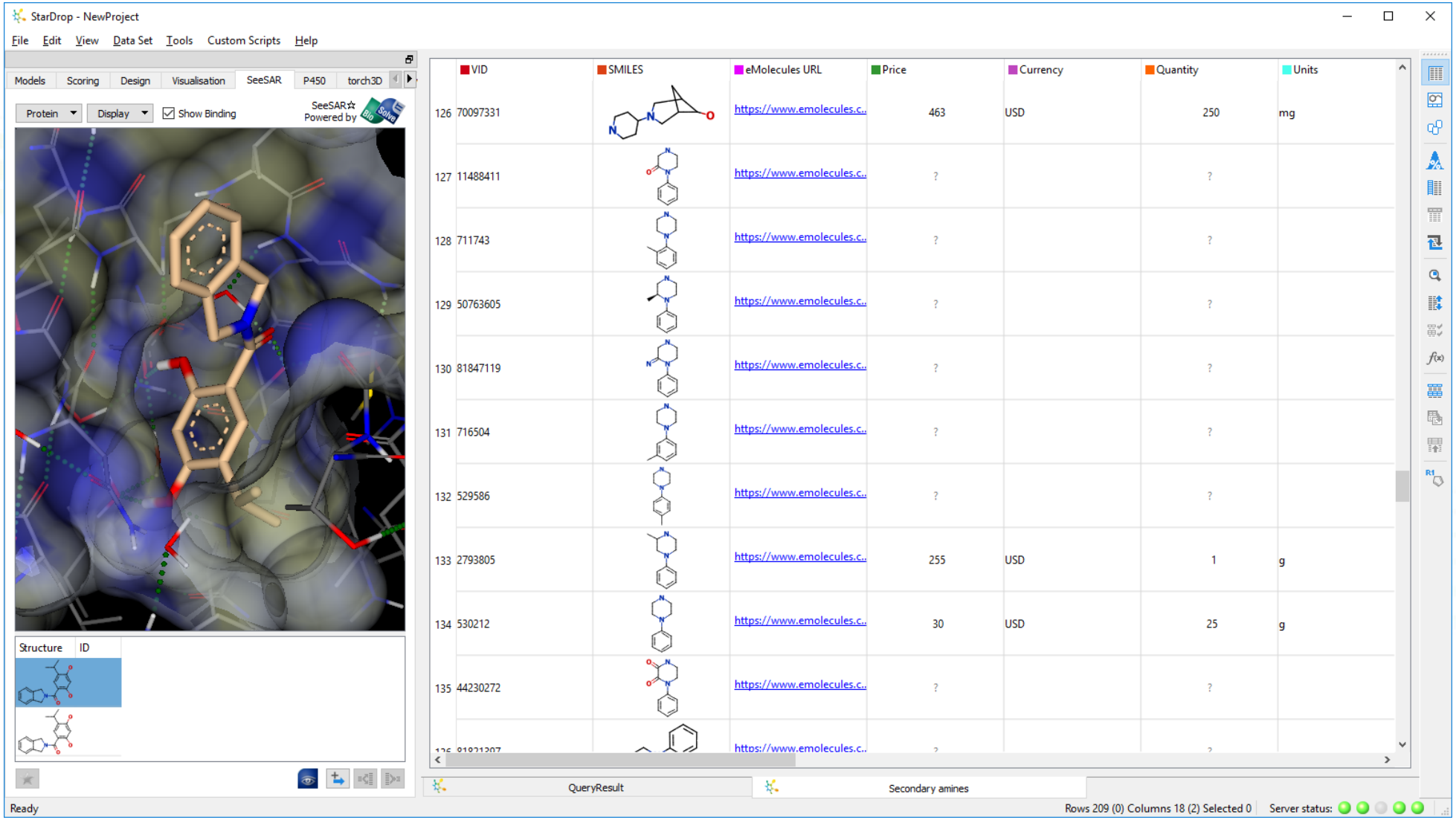


Search eMolecules

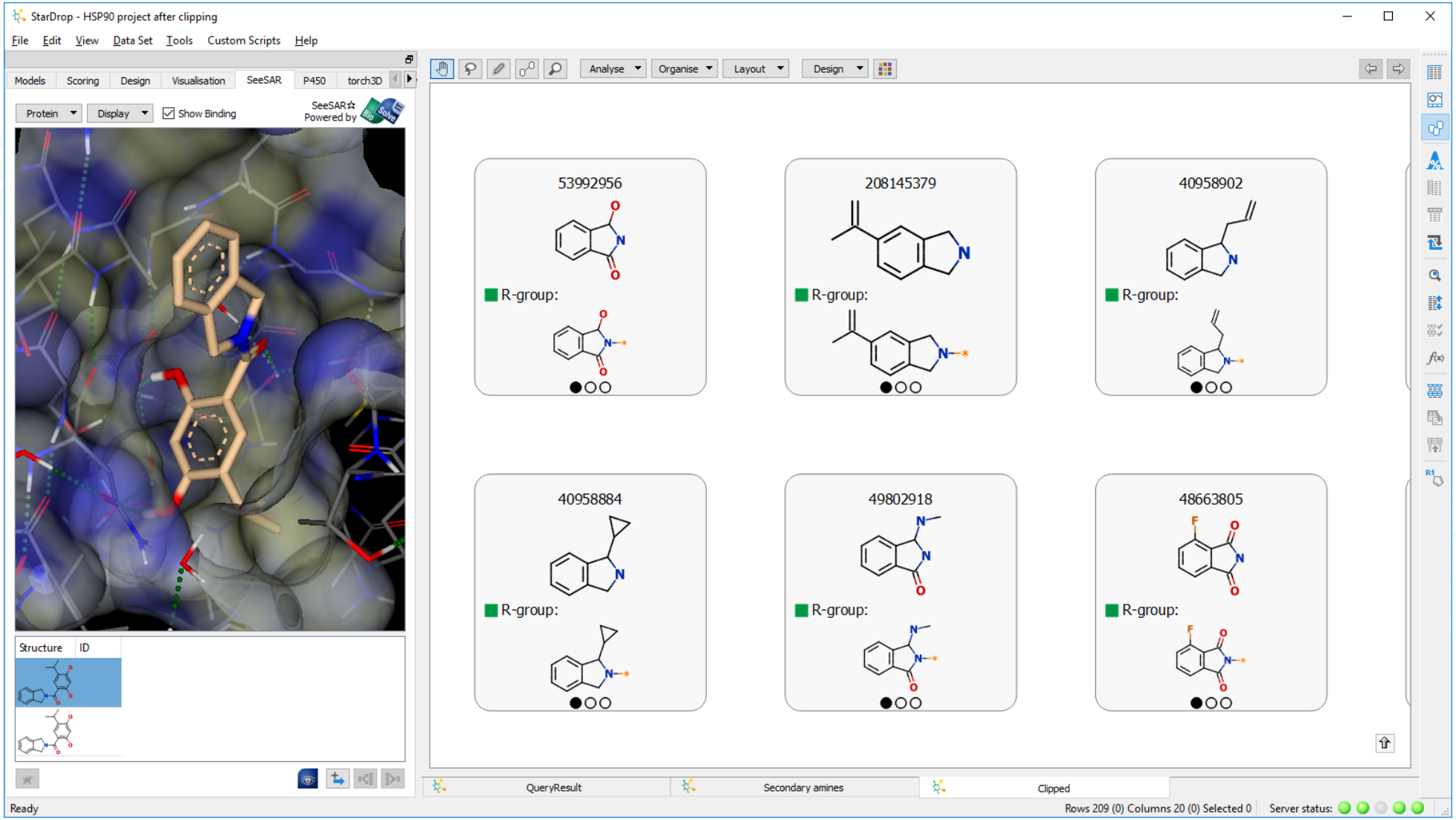
Substituted isoindolines and potential replacements



Search eMolecules
Results

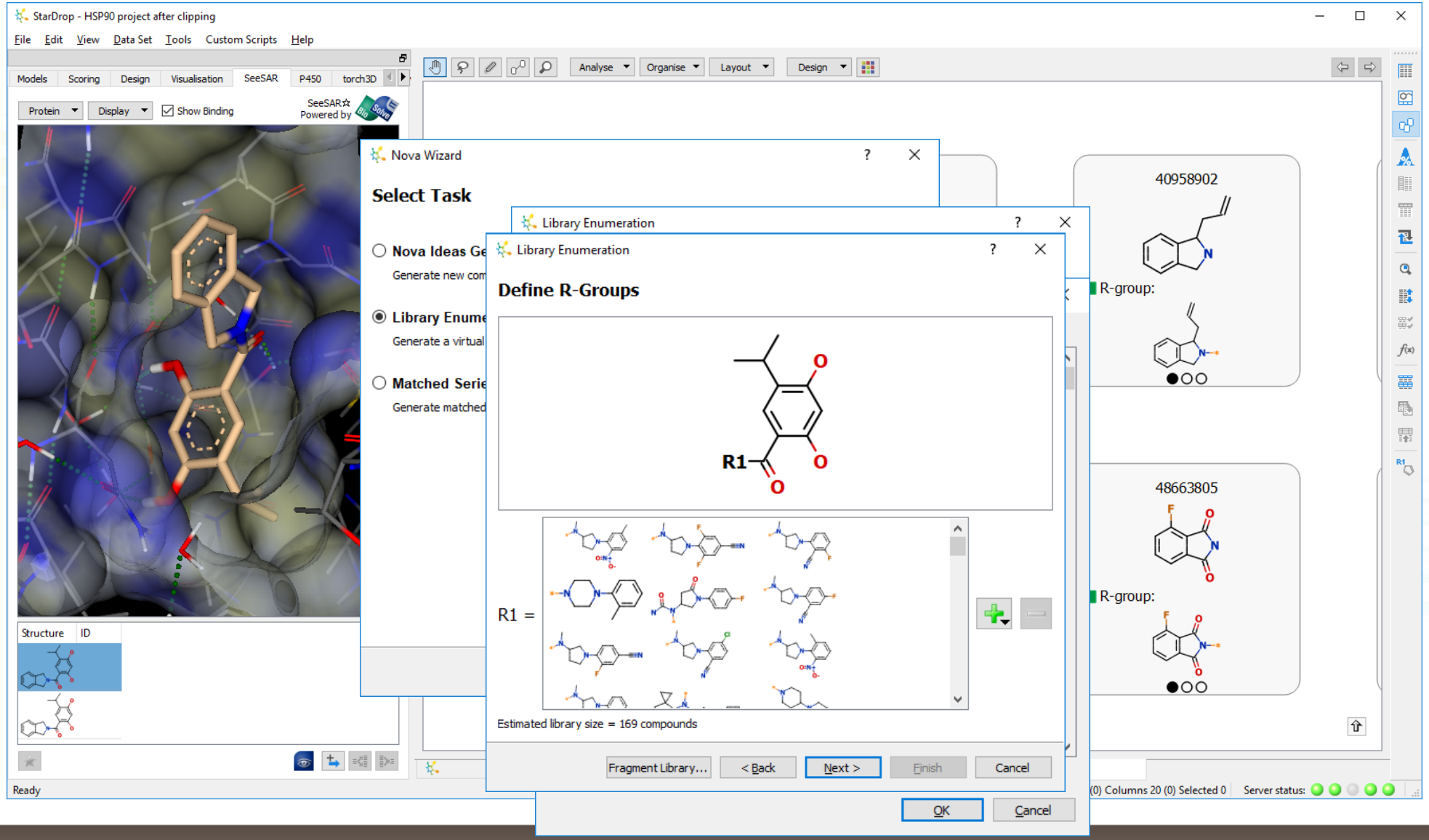


Clip Reagents

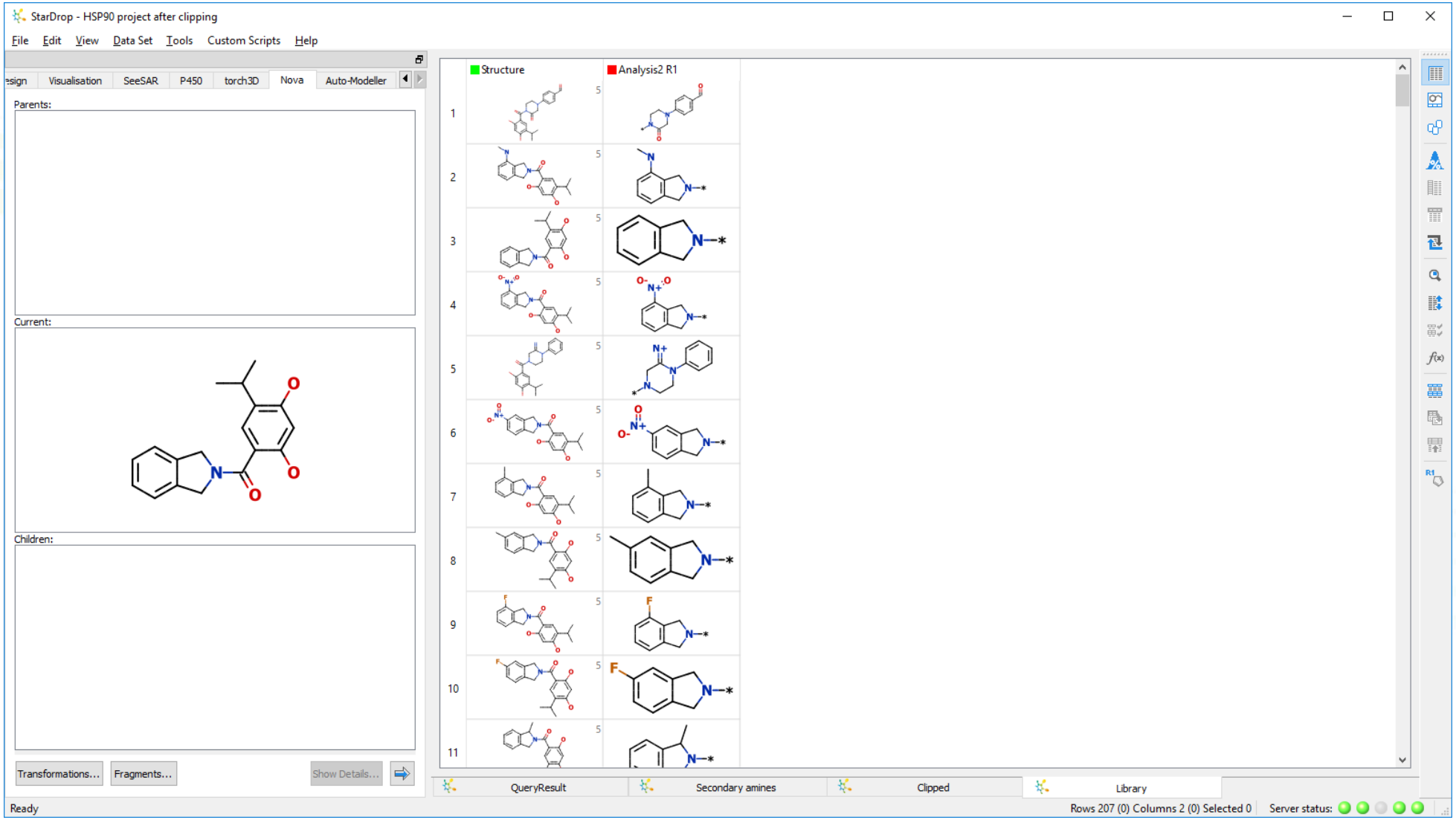


M.D. Segall *et al.* (2015) Drug Discov. Today **20**(9) pp. 1093-1103

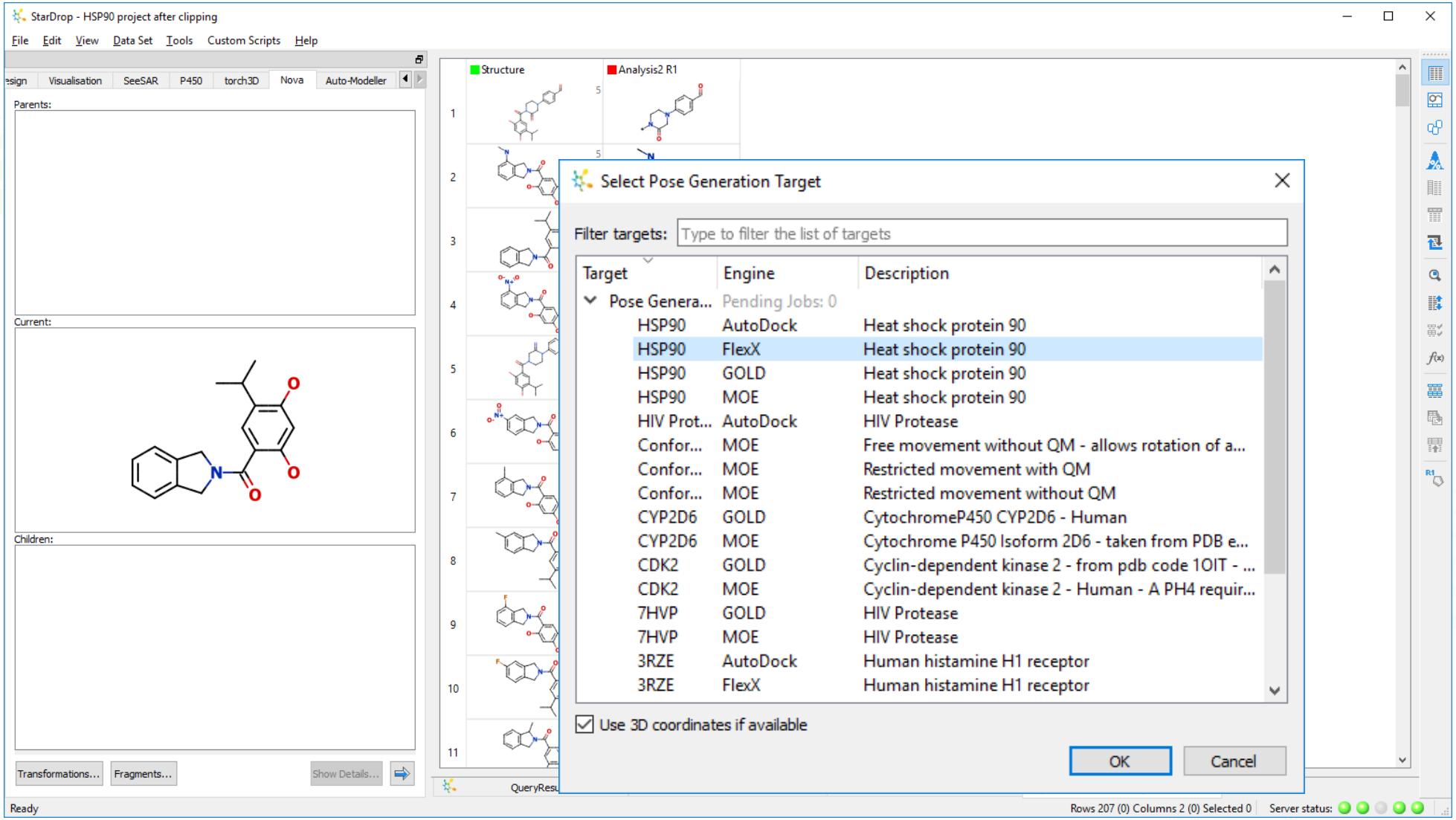
Enumerate Library



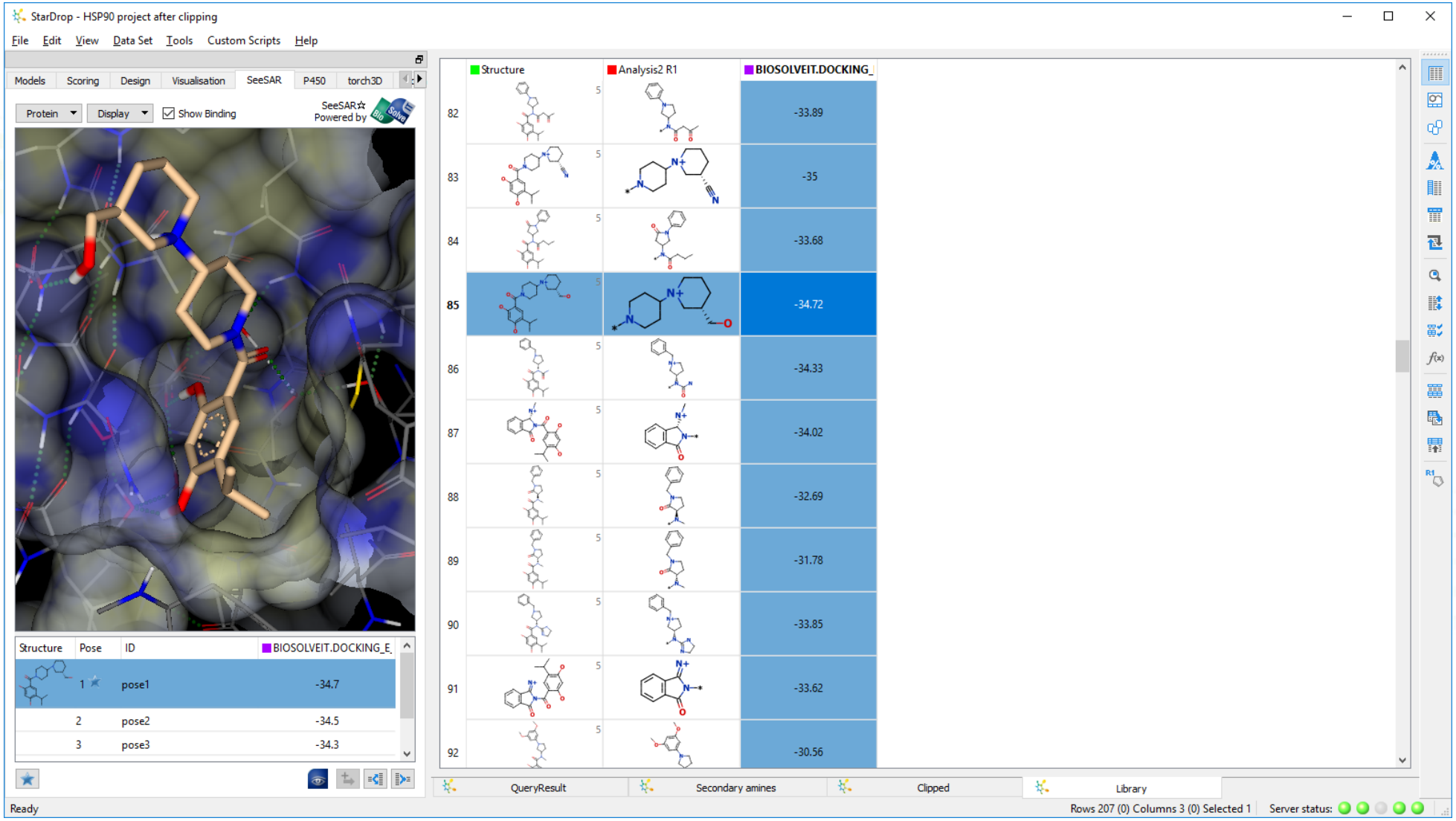
Enumerated Library



Dock Library



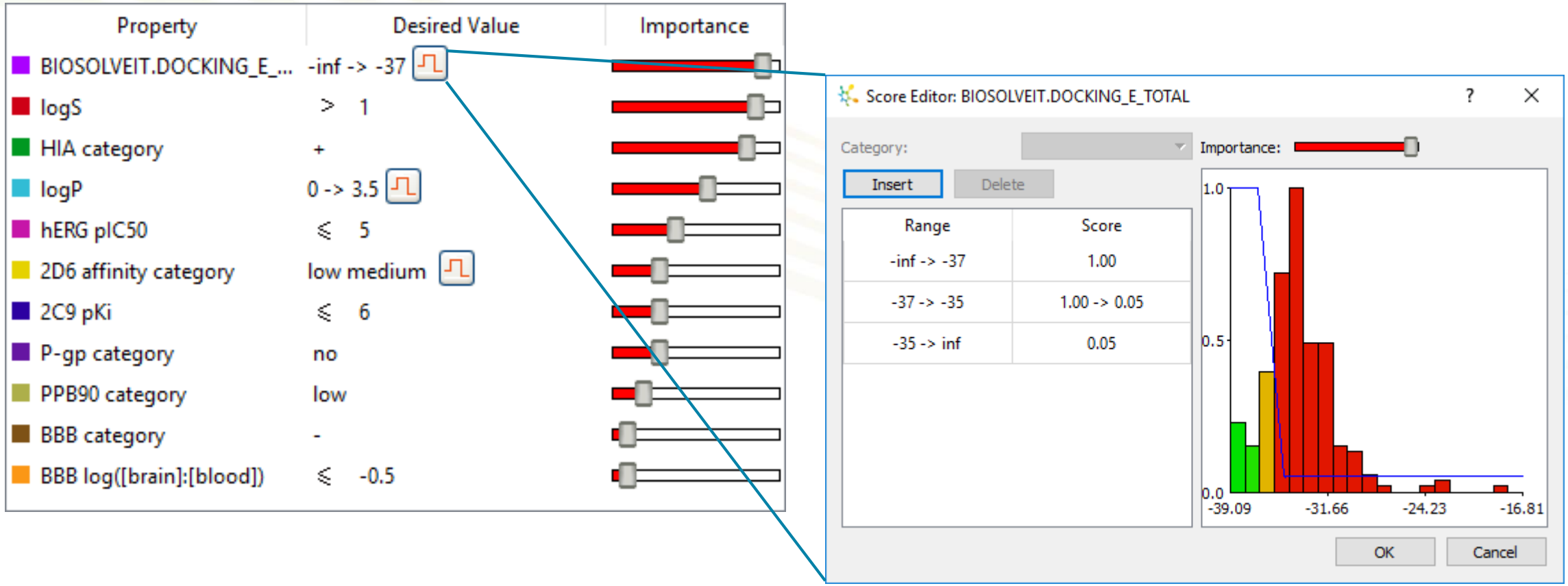
Dock Library



Good Docking Score is Not Enough

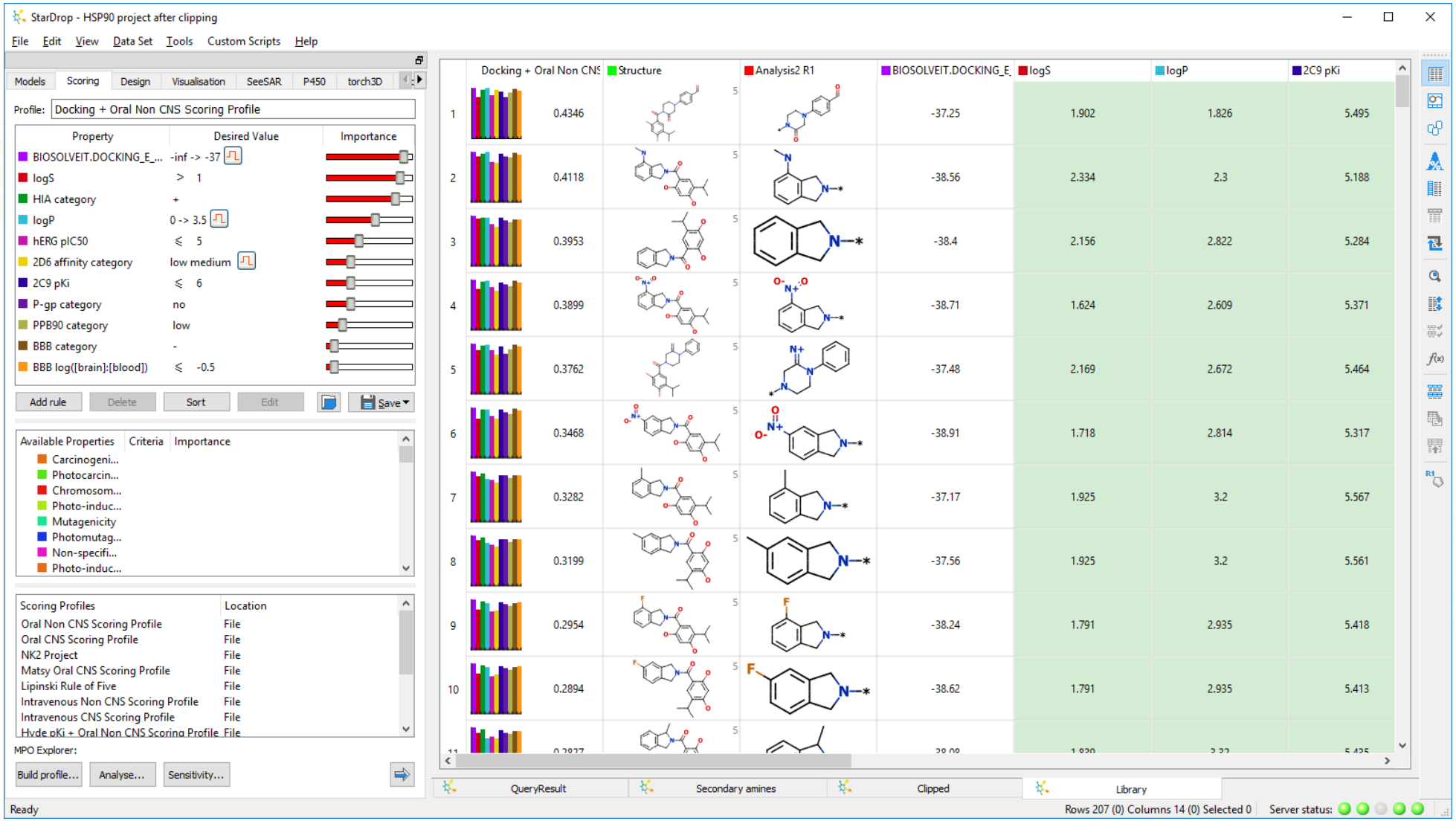
Multi-parameter optimisation

- Also require good Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) properties



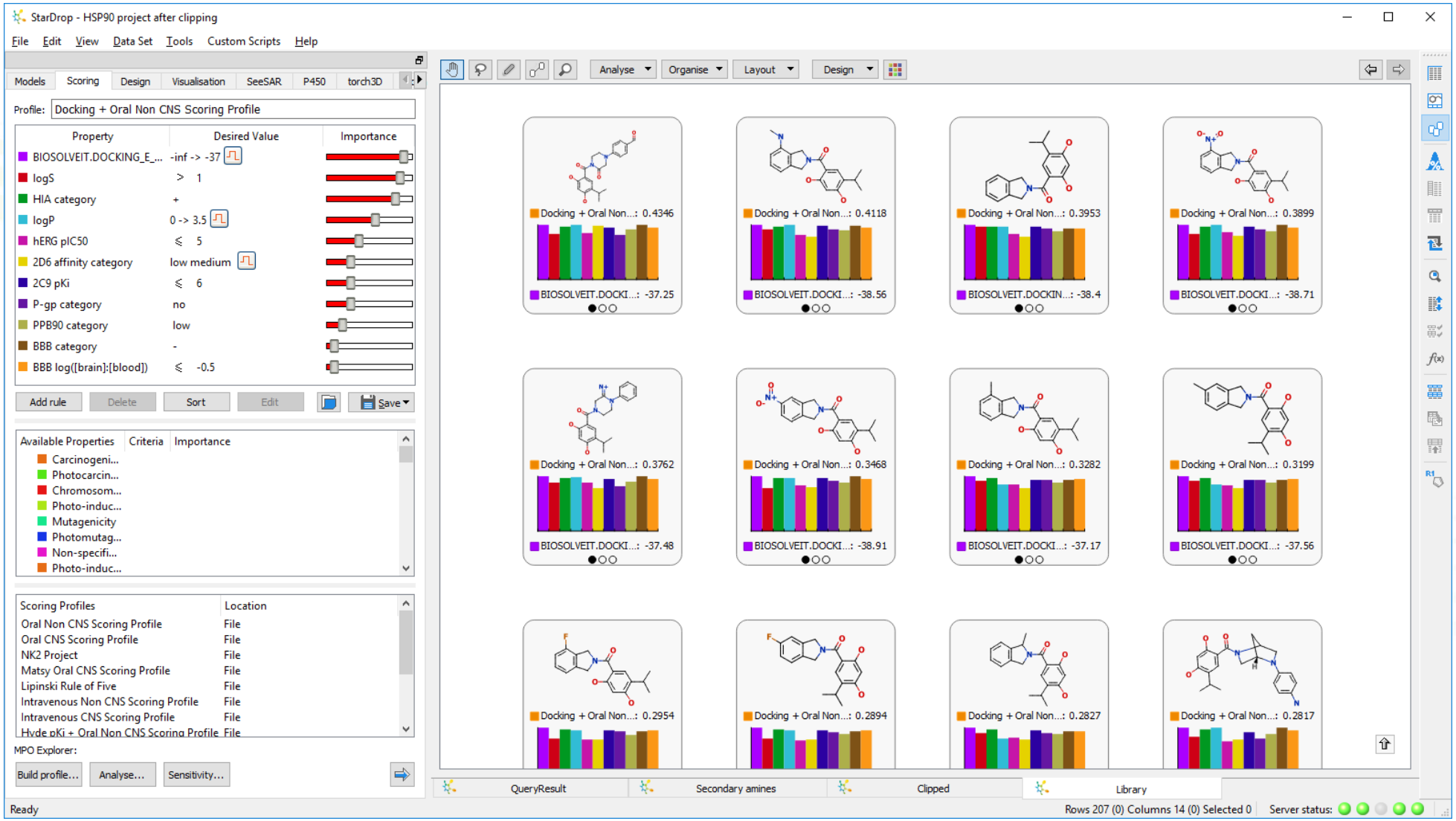
M.D. Segall (2012) Curr. Pharm. Des. **18**(9) pp. 1292-1310

MPO Scores



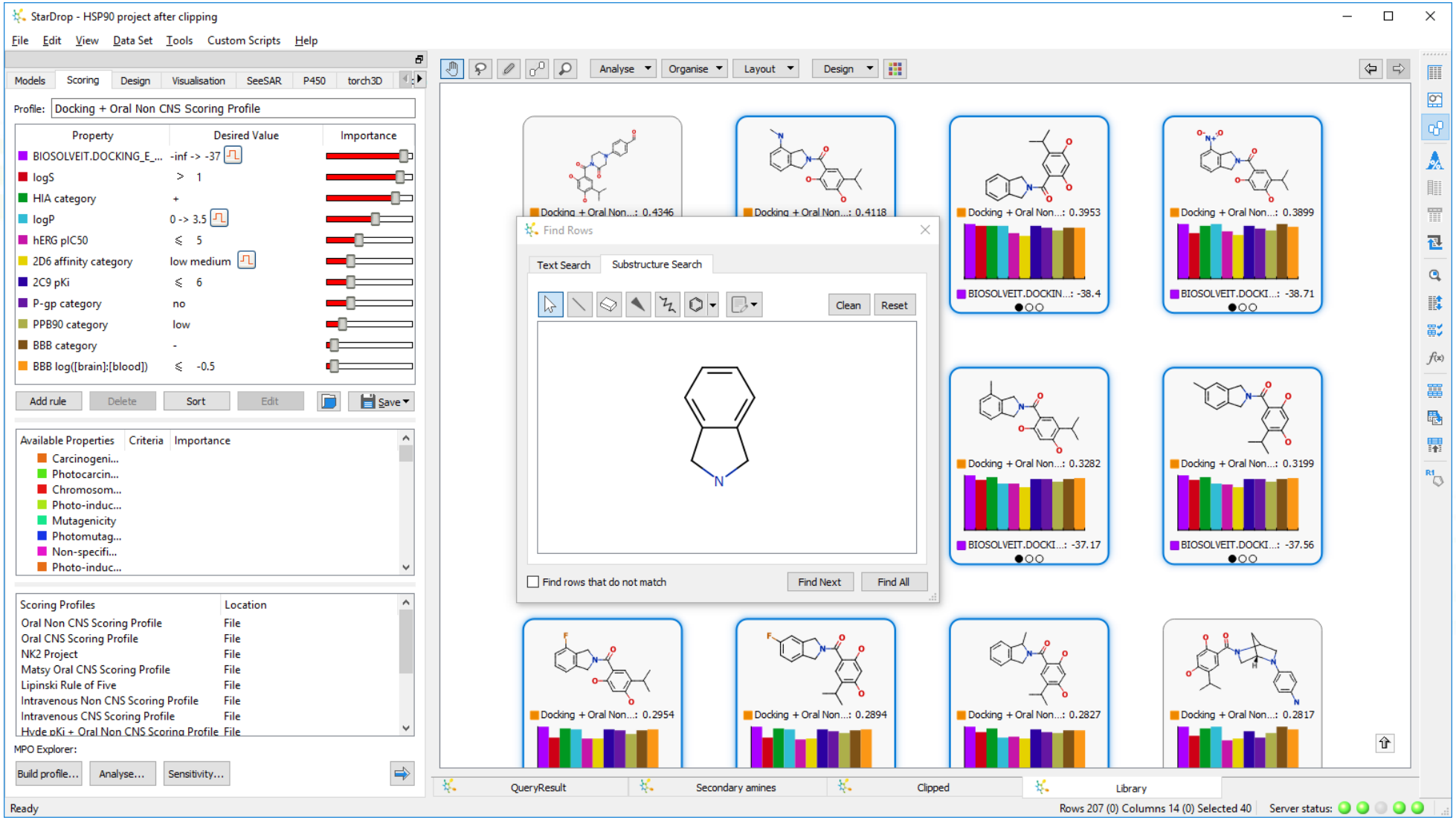
M.D. Segall (2012) Curr. Pharm. Des. **18**(9) pp. 1292-1310

Compound Overview



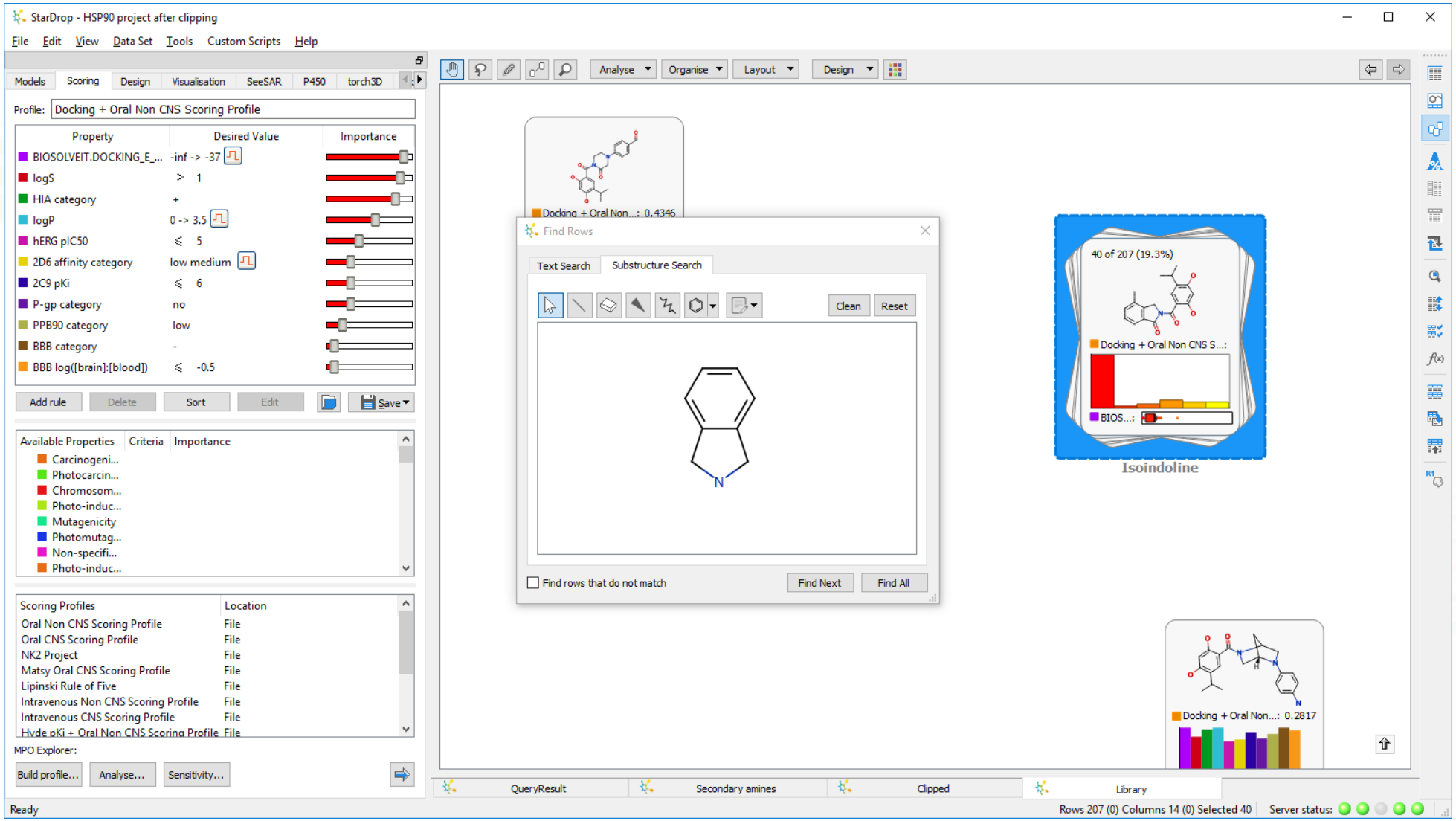
M.D. Segall (2012) Curr. Pharm. Des. **18**(9) pp. 1292-1310

Identify Chemotypes

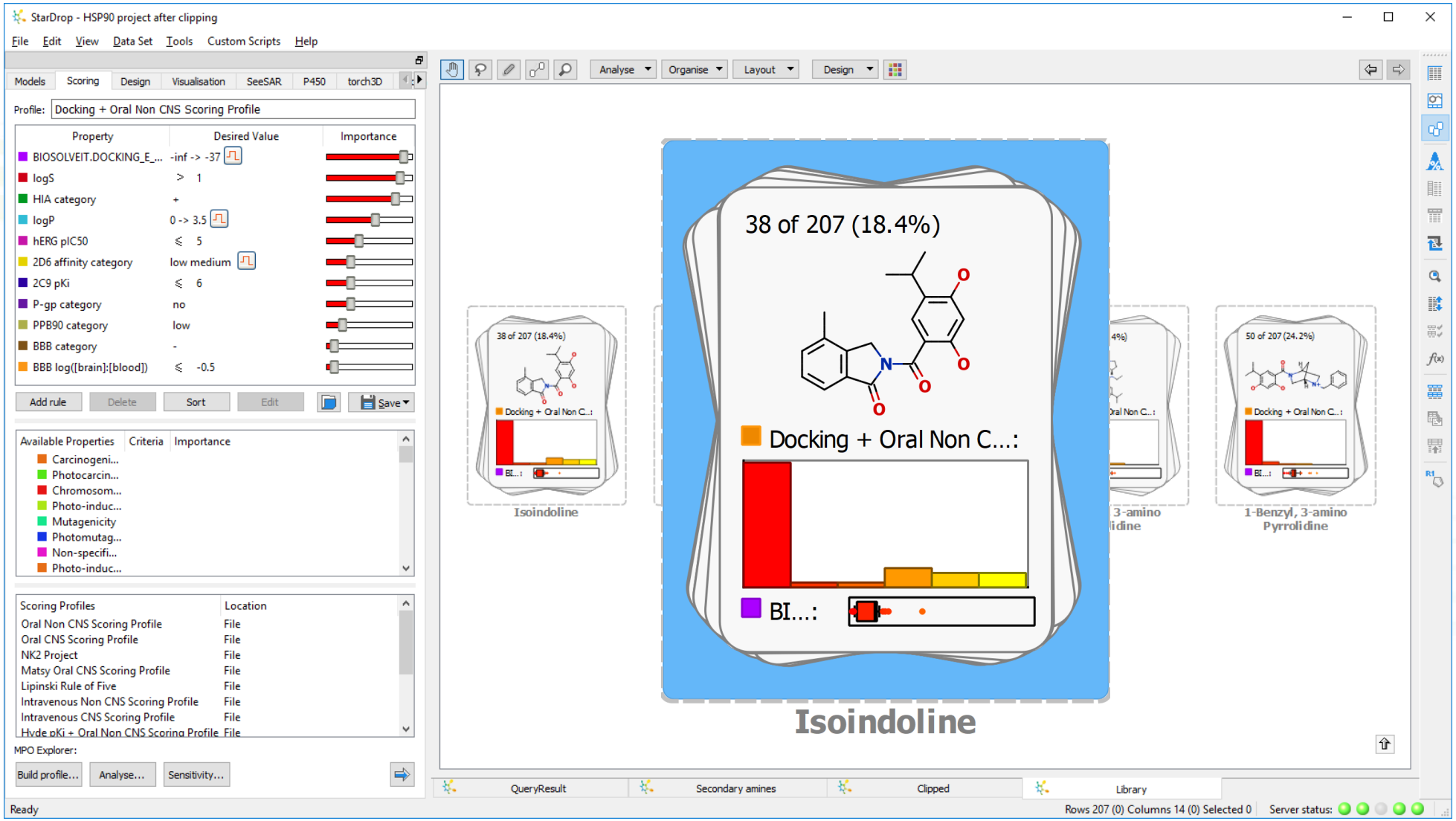


M.D. Segall (2012) Curr. Pharm. Des. **18**(9) pp. 1292-1310

Stack Chemotypes

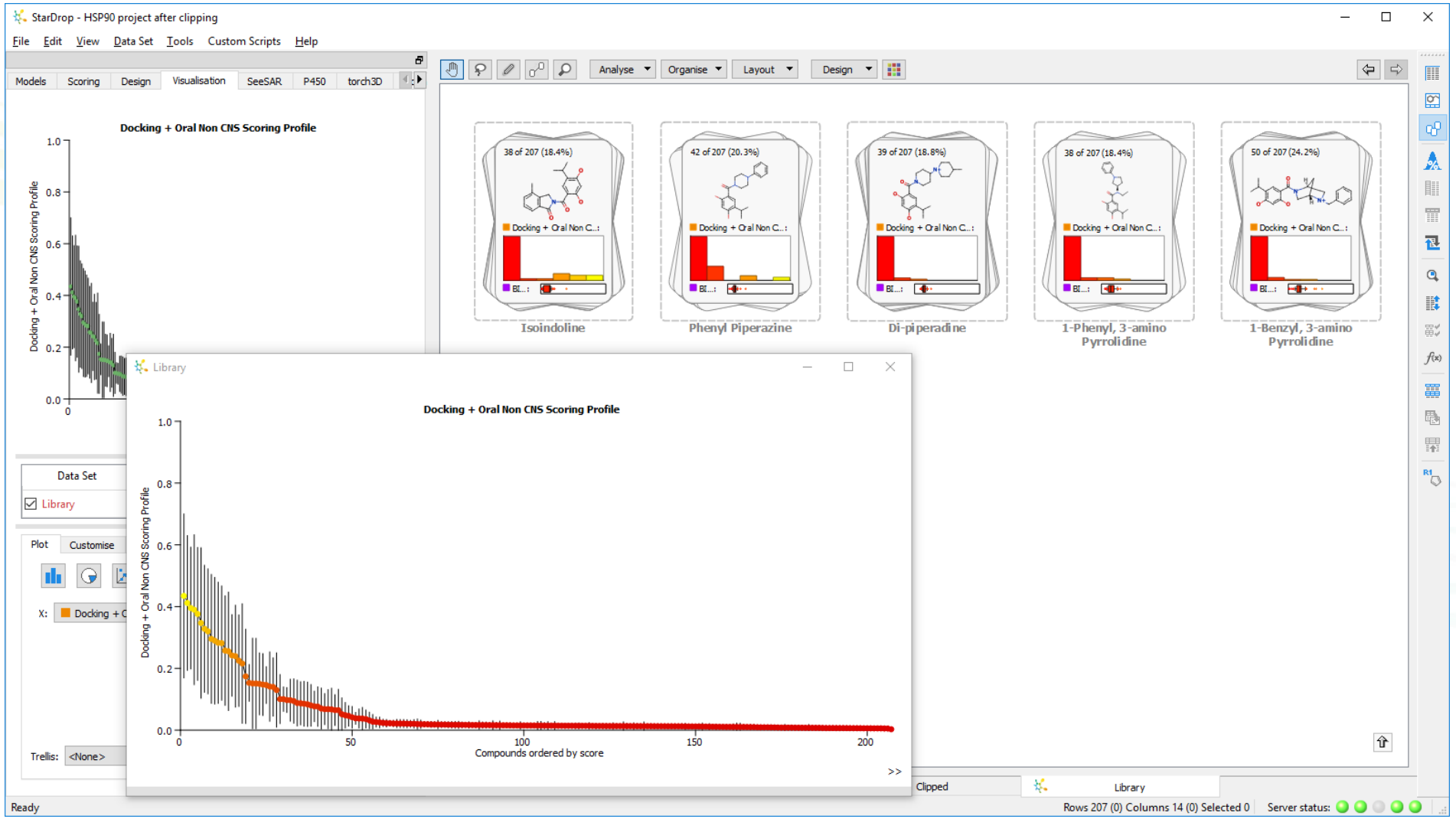


M.D. Segall *et al.* (2015) Drug Discov. Today **20**(9) pp. 1093-1103



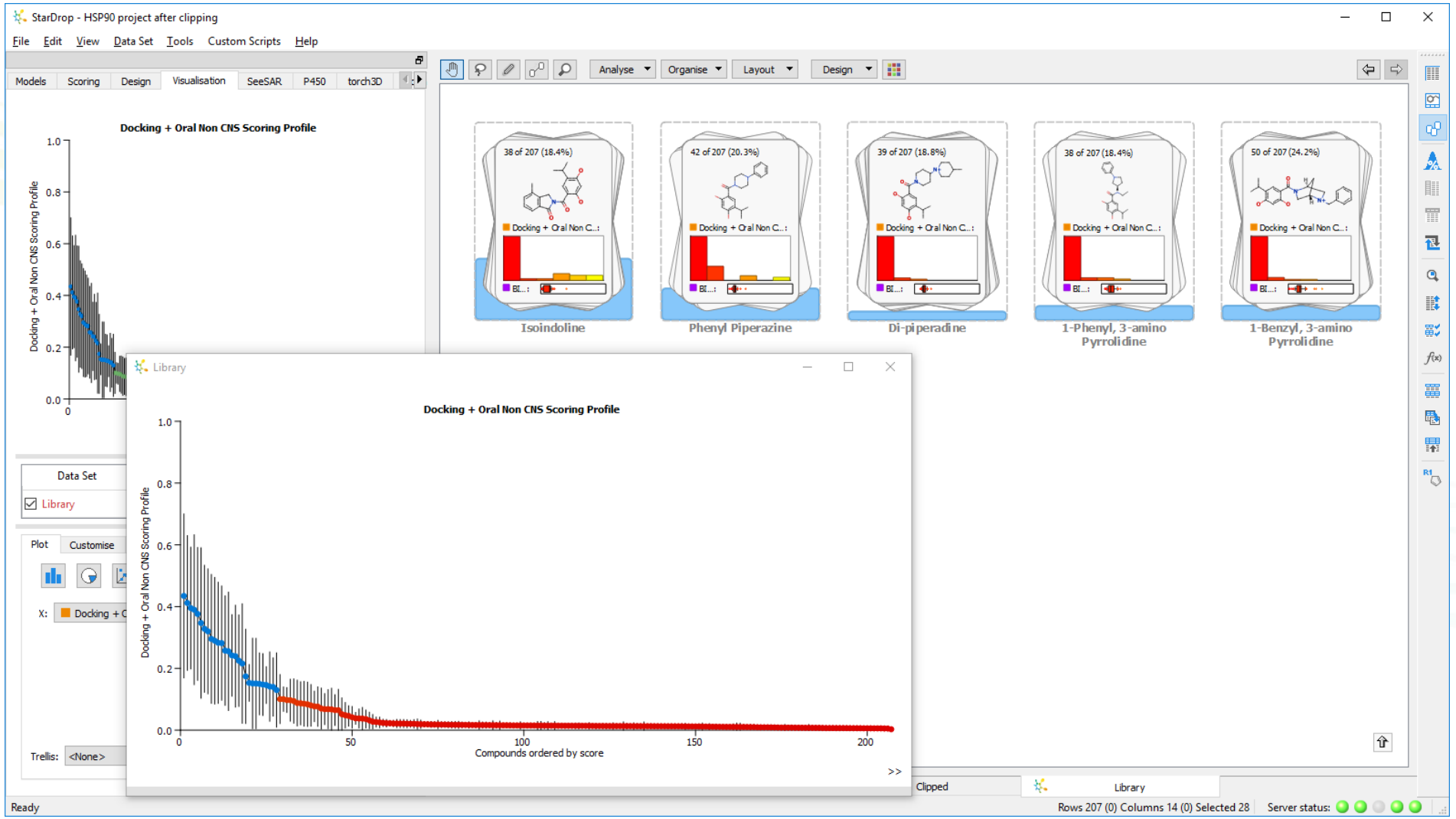
M.D. Segall (2012) Curr. Pharm. Des. **18**(9) pp. 1292-1310

Score Distribution



M.D. Segall (2012) Curr. Pharm. Des. **18**(9) pp. 1292-1310

Compound Selection



M.D. Segall (2012) Curr. Pharm. Des. **18**(9) pp. 1292-1310

Conclusion

- Integration of cheminformatics and computational chemistry tools is essential for efficient drug discovery
  - Quick and good decisions on compound selection and design
- Ease-of-use and interactivity critical
  - All users should be able to intuitively access data and predictions
  - Bring together all data to target high-quality compounds
- Big challenges
  - Compatibility with wide range of vendors and in-house platforms
  - Support for variety of architectures
- For more information: www.optibrium.com

Acknowledgements

- Optibrium team
  - Too many to list!
- BioSolveIT – Access to FlexX and HSP90 model
  - Christian Lemmen
  - Marcus Gastreicht
  - Carsten Detering
- Collaborators in development of query and pose generation interfaces
  - Zoetis, The Edge, ChemAxon, IDBS…
  - BioSolveIT, CCDC, CCG, OpenEye…