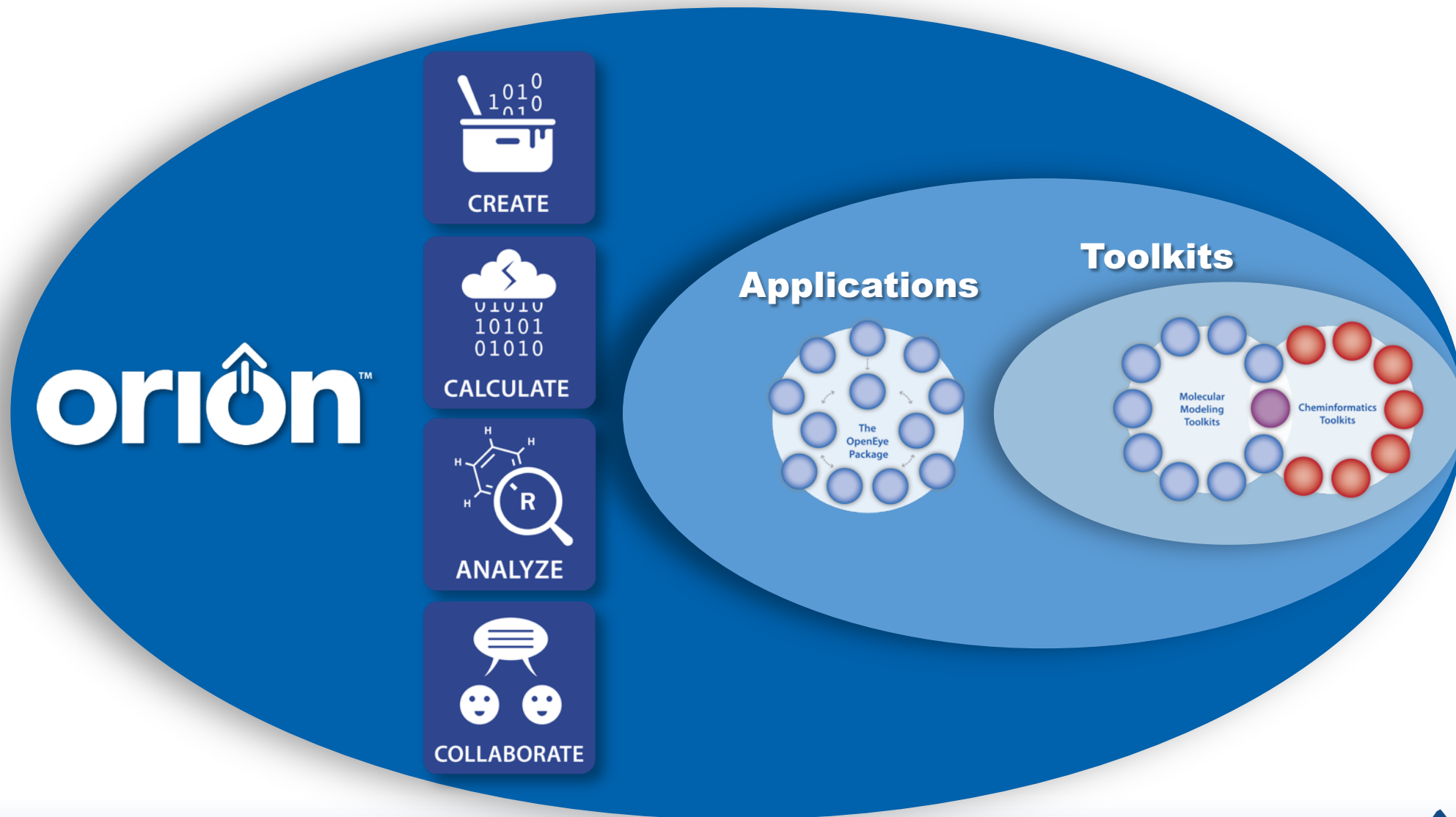


Toolkit & Application 2021.1 Update

Jesper Sørensen

Scientific Tools at OpenEye



Outline

- Platform & Language Support
- Minor updates
- VIDA 5.0
- Docking receptors
- GPU Efforts

Acknowledgments

OpenEye toolkits and applications
development team

Outline

- Platform & Language Support
- Minor updates
- VIDA 5.0
- Docking receptors
- GPU Efforts

Compiler and Language Support

- Python 3.9 support has been added
- We now officially support C++17
 - GCC 7.4 is our minimum supported GCC compiler
 - Support has been added for GCC 9.x
 - VS2017 and VS2019 are supported

Platform & Language Support

Package	Versions	Linux Platforms	Windows	MacOS*
Python	3.7, 3.8, 3.9	RHEL7/8, Ubuntu18/20	Win10	10.14, 10.15, 11
C++		RHEL7/8, Ubuntu18/20	Win10 (VS2017,VS2019)	10.14, 10.15, 11
Java	1.8, 11	RHEL7/8, Ubuntu18/20	Win10	10.14, 10.15, 11
C#			Win10 (VS2017,VS2019)	
Applications		RHEL7/8, Ubuntu18/20	Win10	10.14, 10.15, 11

* Arm/M1 Mac is not supported at this time

Outline

- Platform & Language Support
- Minor updates
- VIDA 5.0
- Docking receptors
- GPU Efforts

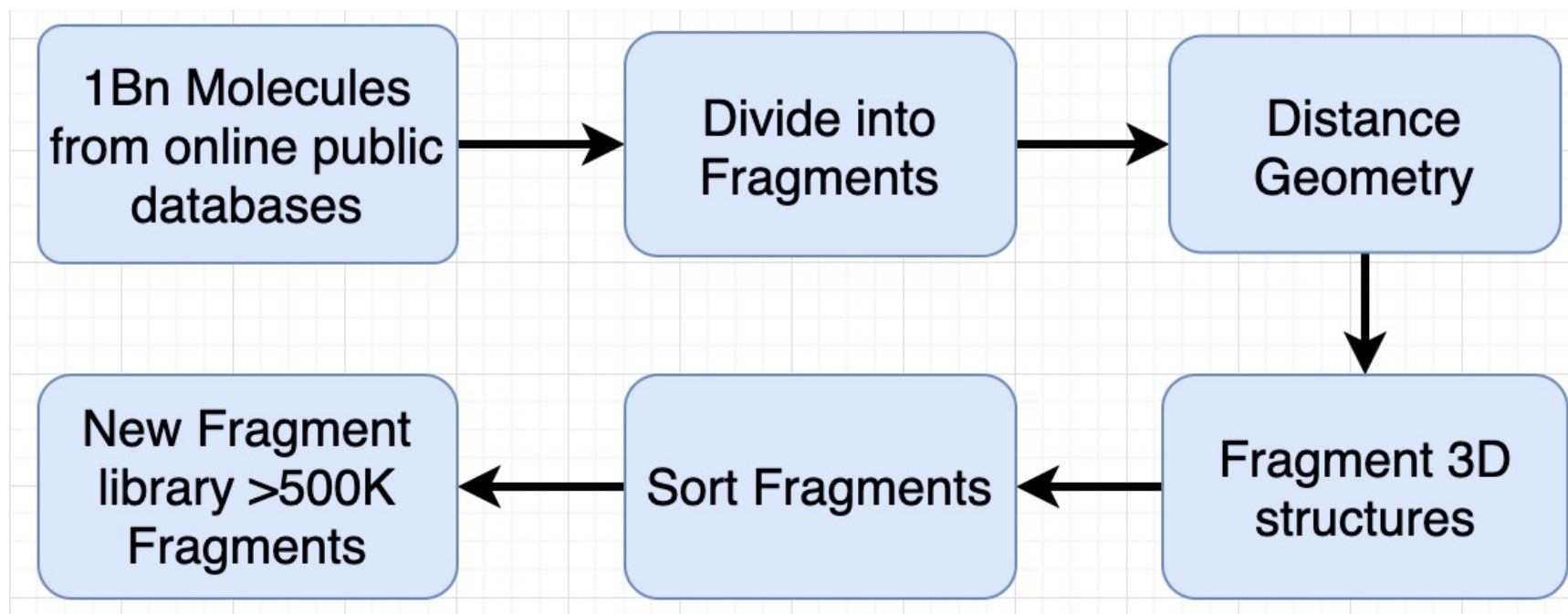
Toolkit Updates – OEChem TK

- ZSTD has been updated to version 1.4.9
- Better performance for OEZ file format
- OERecord API was released in the Spring 2020
- Switching molecule storage on records to use OEZ
- Toolkit change that will positively affect Orion
- Smaller dataset sizes, faster data transfer, faster read/write

Toolkit Updates – OEChem TK

- Protein residue perception was improved with recognition for standard protein terminal capping groups
- InChI has been updated to v1.06
Includes support for non-standard InChI strings with pseudo-atoms present

Toolkit Updates – Omega TK



Existing fragments library: ~80K

Toolkit Updates

- Lexichem TK

Added support for input of many common drug names

- SZYBKI & Szybki TK

AMBER FF14SB support in previous release

Parsley 1.2 support in previous release

Added Parsley 1.3.0 from the Open Force Field Initiative

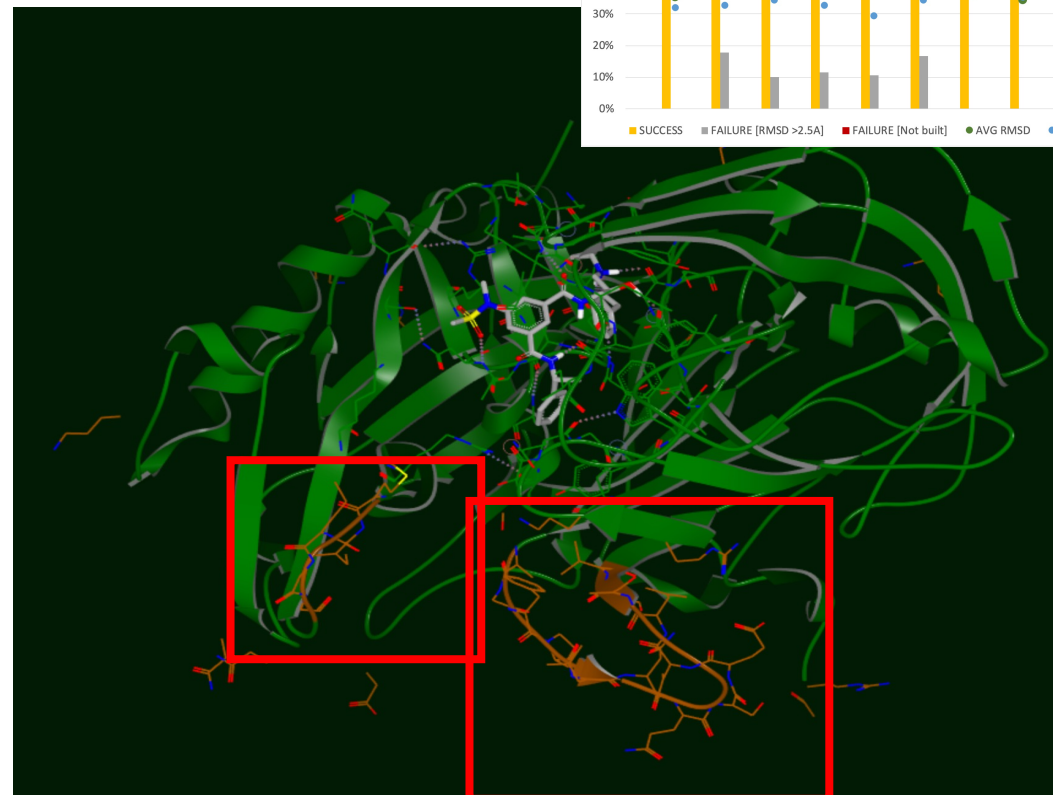
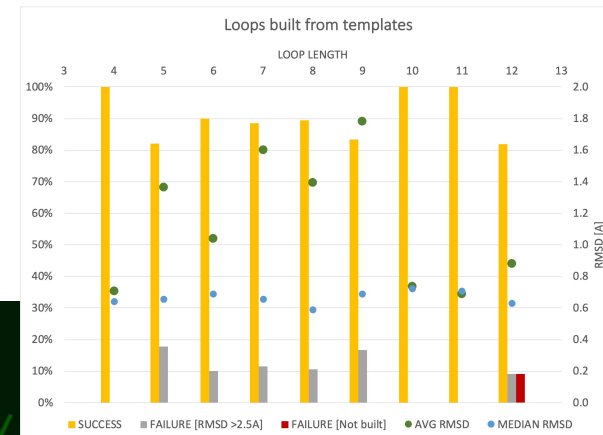
Toolkit Updates – Spruce TK

- Spruce TK was released Fall 2019
- Automated structure to model-ready protein preparation
 - Includes Iridium structure assessment
- Loop modeling was added in Fall 2020
 - Loop database available for download
 - Can build a database of corporate structures or e.g. target specific structures

Loop modeling

- Input
 - OEResidue of anchor residues
 - FASTA of expected sequence
- Extensive control parameters
 - Remove extra residues at the anchors (def: 1)
 - Build disulfide bridges (def: yes)
 - Fuzzy seq search, strict proline (def: yes)

Validation set:
Rossi et al.
(*Prot. Sci.* 2007)



DVA between GLU 310 A and THR 314 A
GFPLNQSEVLAS between ALA 157 A and VAL 170 A

Toolkit Updates – Spruce TK

- Improvements to PDB file perception for standard residues
- The superposition API has been refactored from the preliminary API
- OESuperposeMethod namespace
 - Global Sequence, (Binding) Site Sequence, Difference Density Matrix (DDM), Secondary Structure Elements (SSE), ...
- OESuperposeOptions
- OESuperpose
- OESuperposeResults

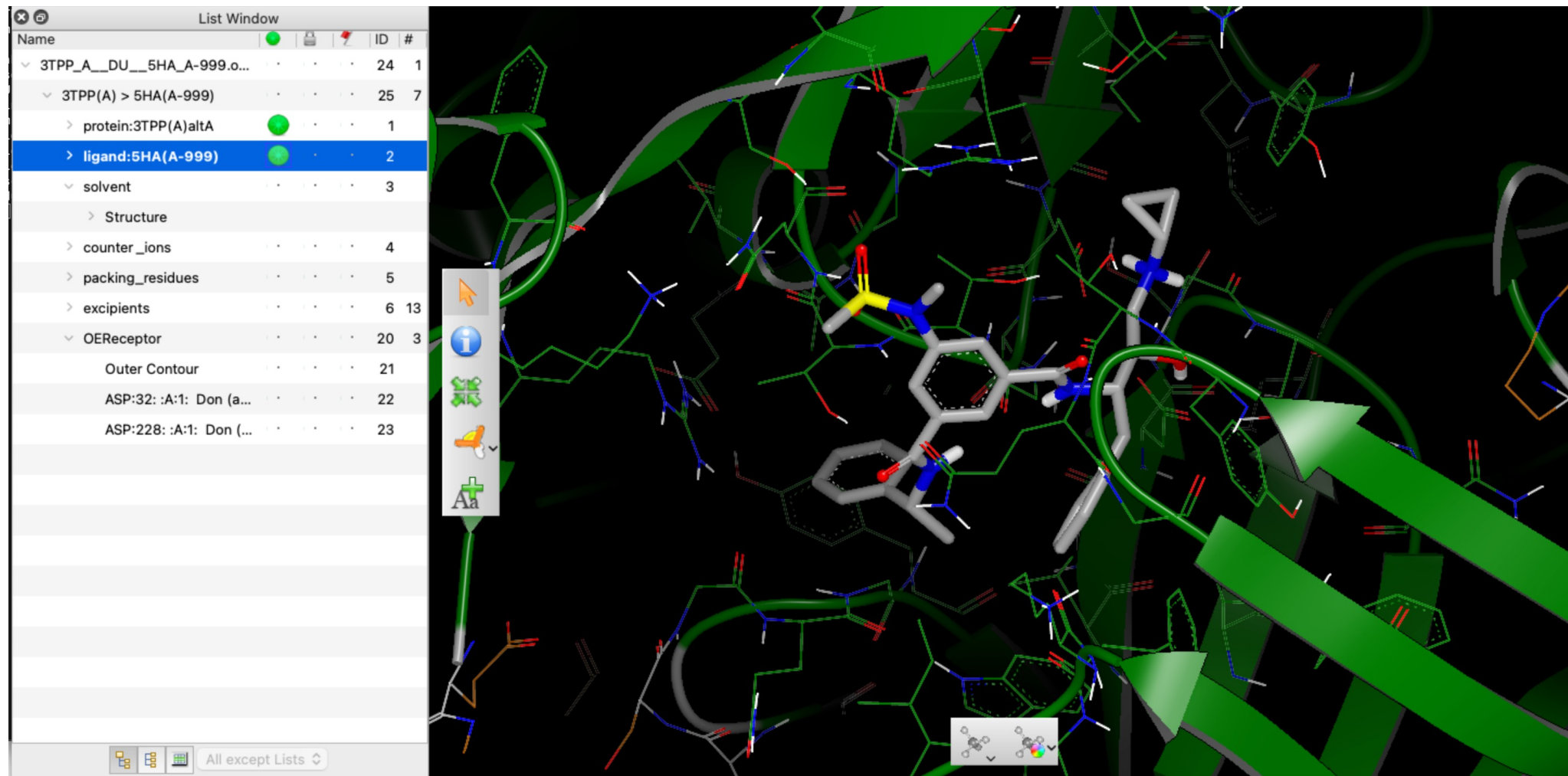
Outline

- Platform & Language Support
- Minor updates
- **VIDA 5.0**
- Docking receptors
- GPU Efforts

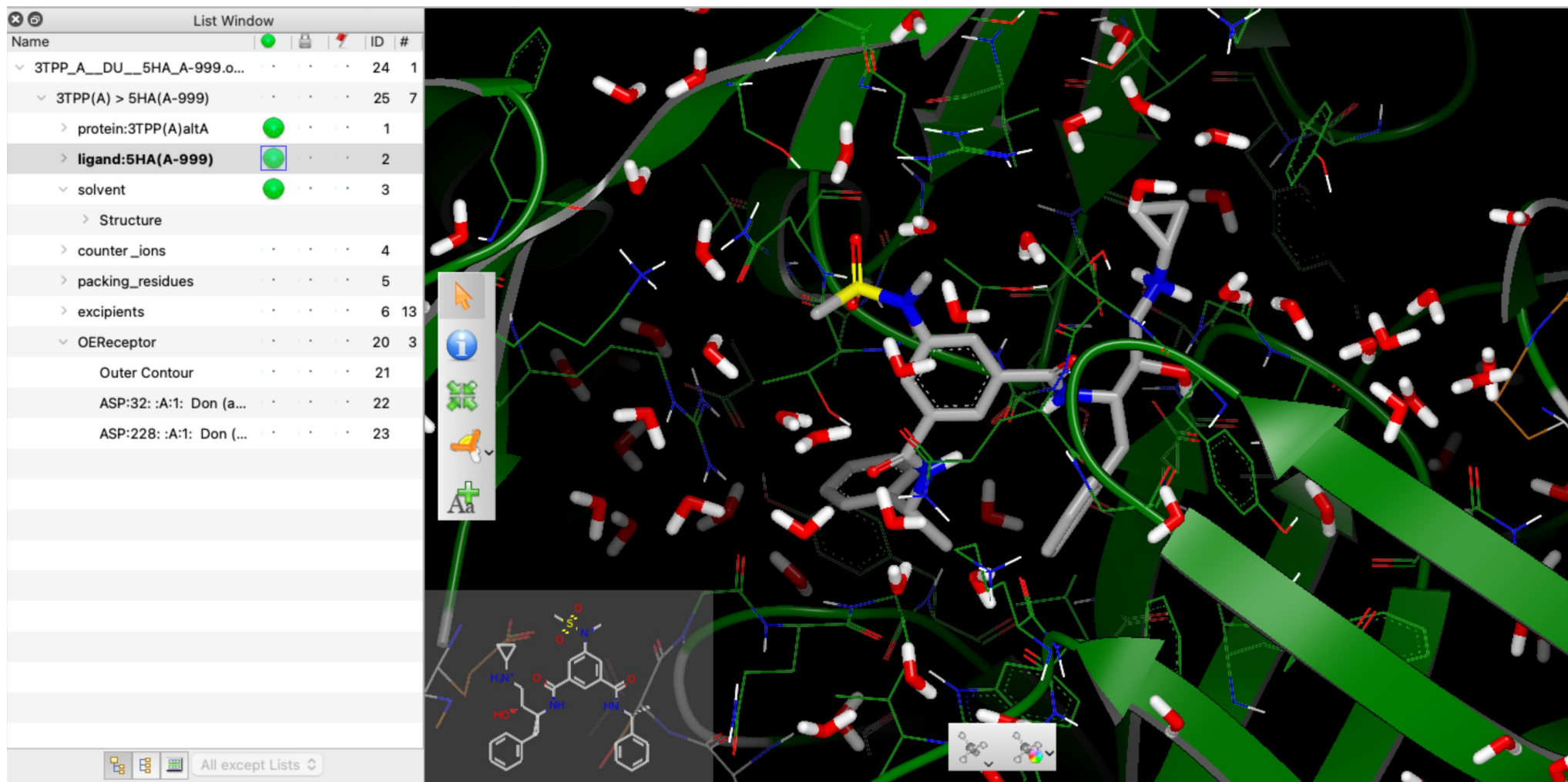
VIDA 5.0

- Python 3
- Qt 5 and PySide2
 - Enables full Qt-library access
- OpenEye Toolkits 2021.1.0
- VIDA will release regularly along side the toolkit and application bundle

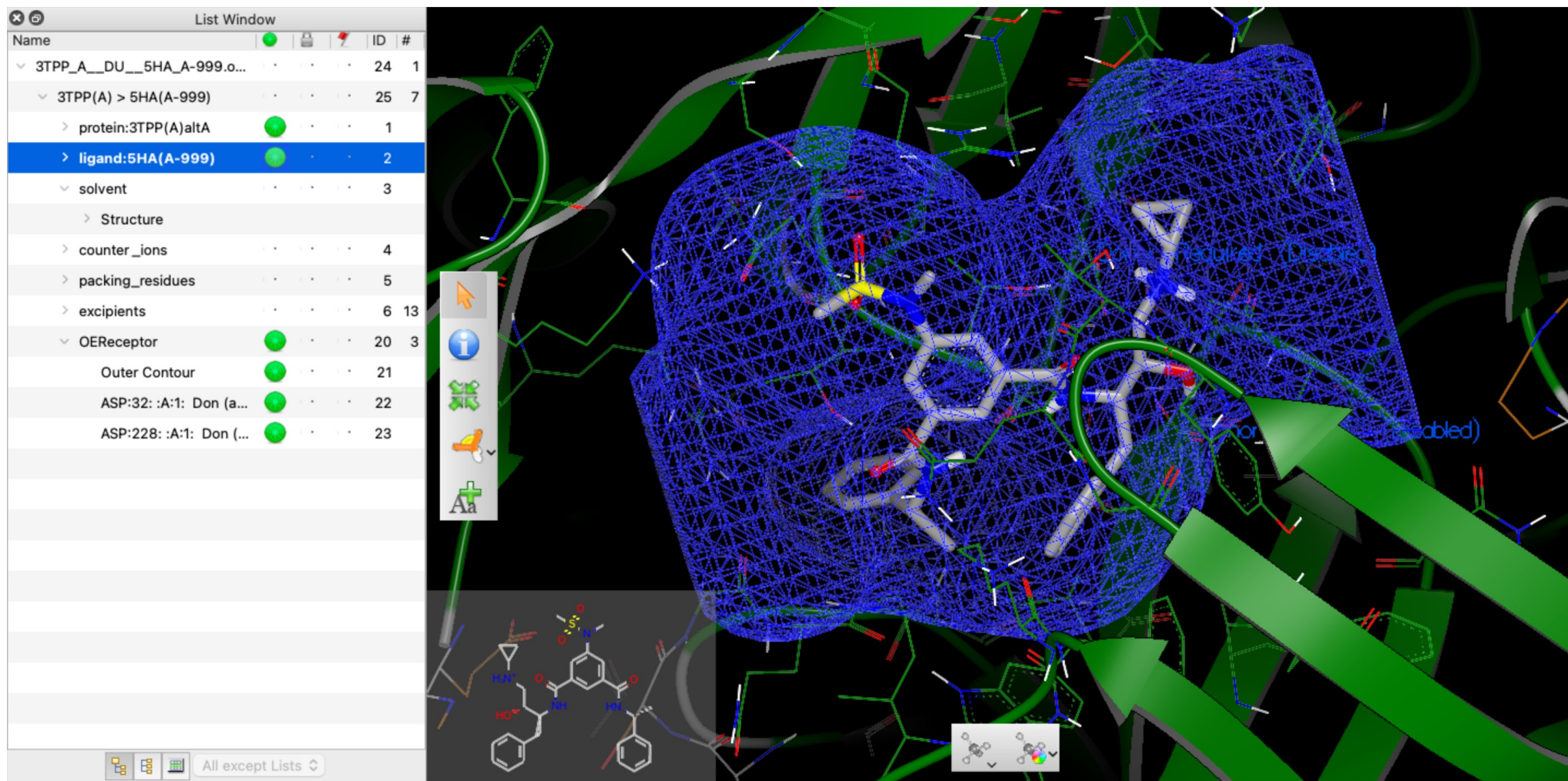
VIDA 5.0 – design unit and receptor support



VIDA 5.0 – design unit and receptor support



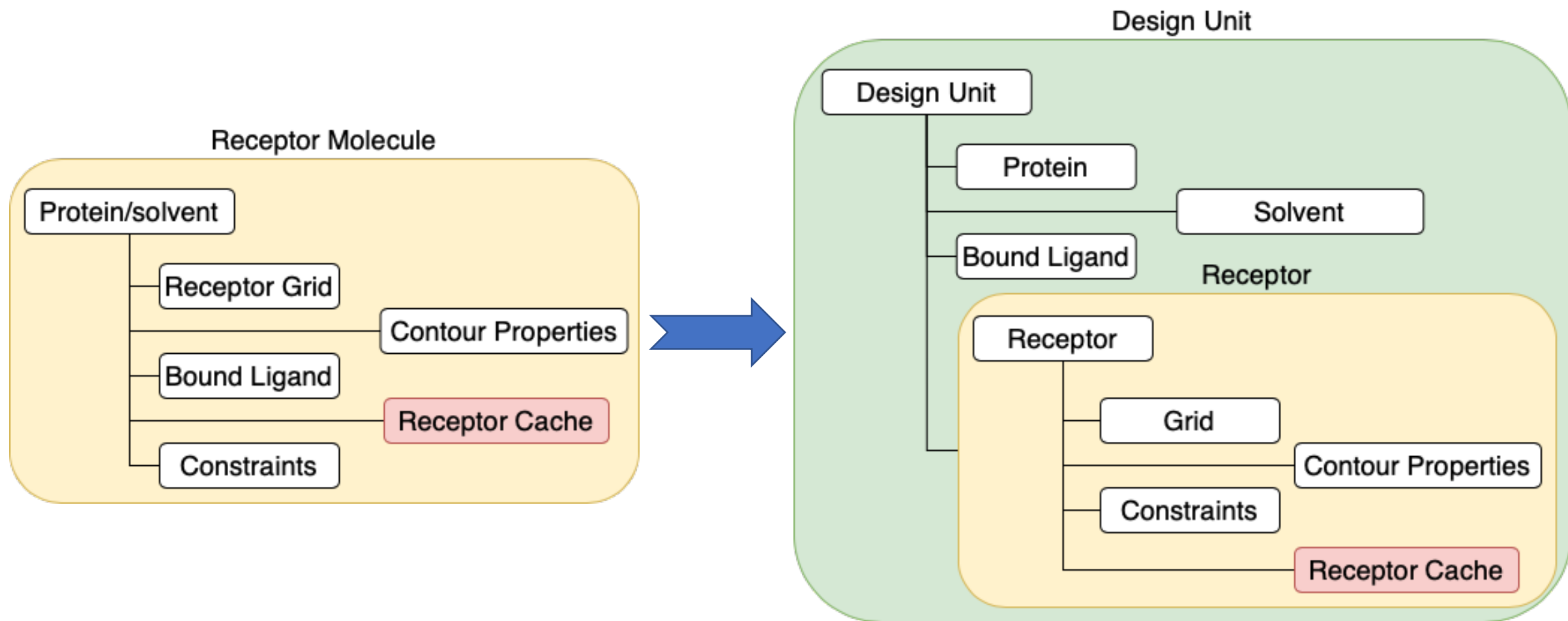
VIDA 5.0 – design unit and receptor support



Outline

- Platform & Language Support
- Minor updates
- VIDA 5.0
- Docking and Receptors
- GPU Efforts

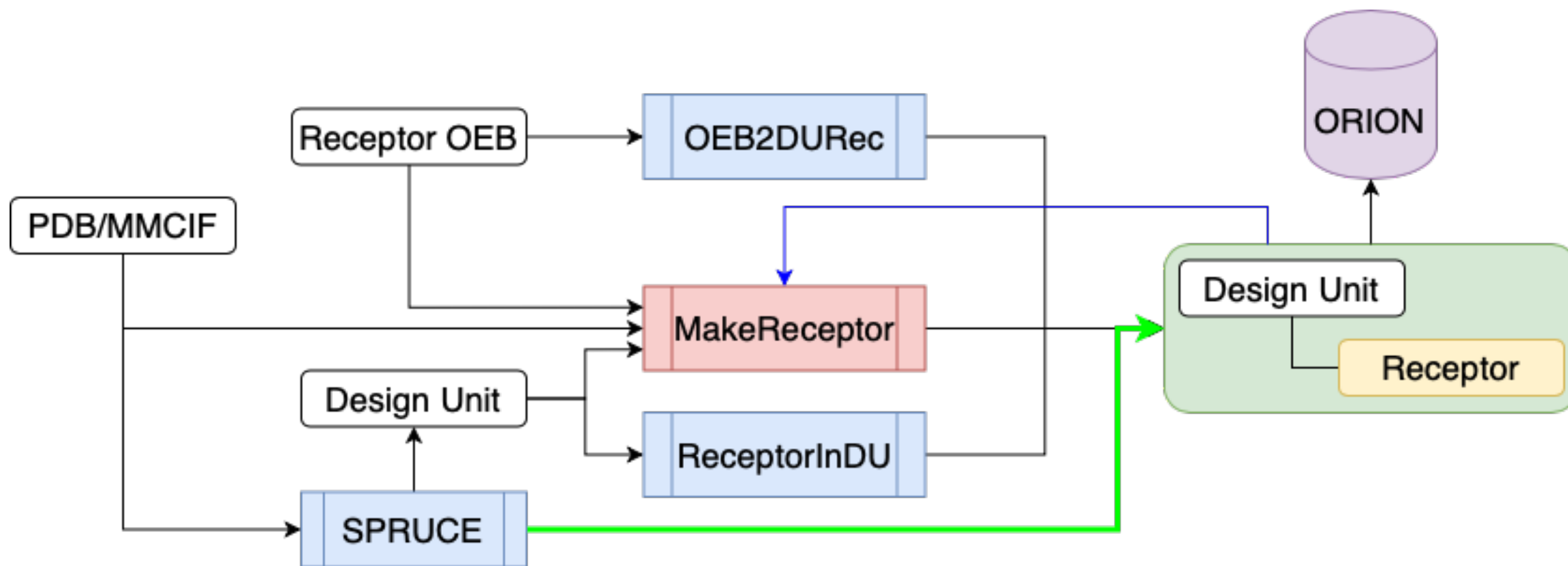
Fall 2020 - Receptor in Design Unit



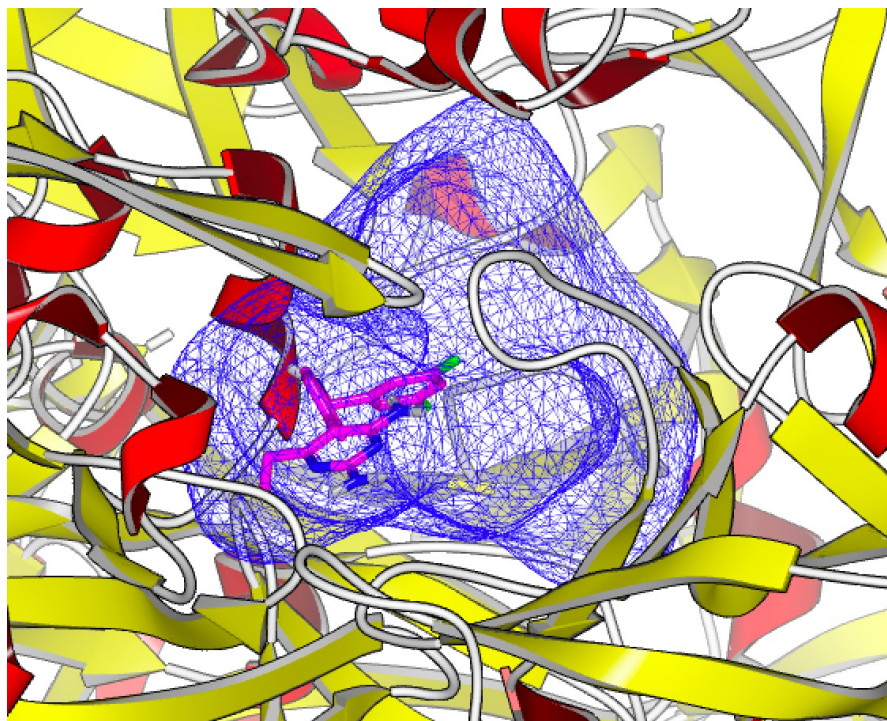
Spring 2021

- SPRUCE now generates docking receptors on design units by default
- MakeReceptor can now read OEDU files (and still OEB files)
- MakeReceptor can now run protein preparation using SPRUCE with a PDB/MMCIF & MTZ input
- Design units with receptors can now be visualized in VIDA

Receptor in 2021.Spring



Receptor in Design Unit

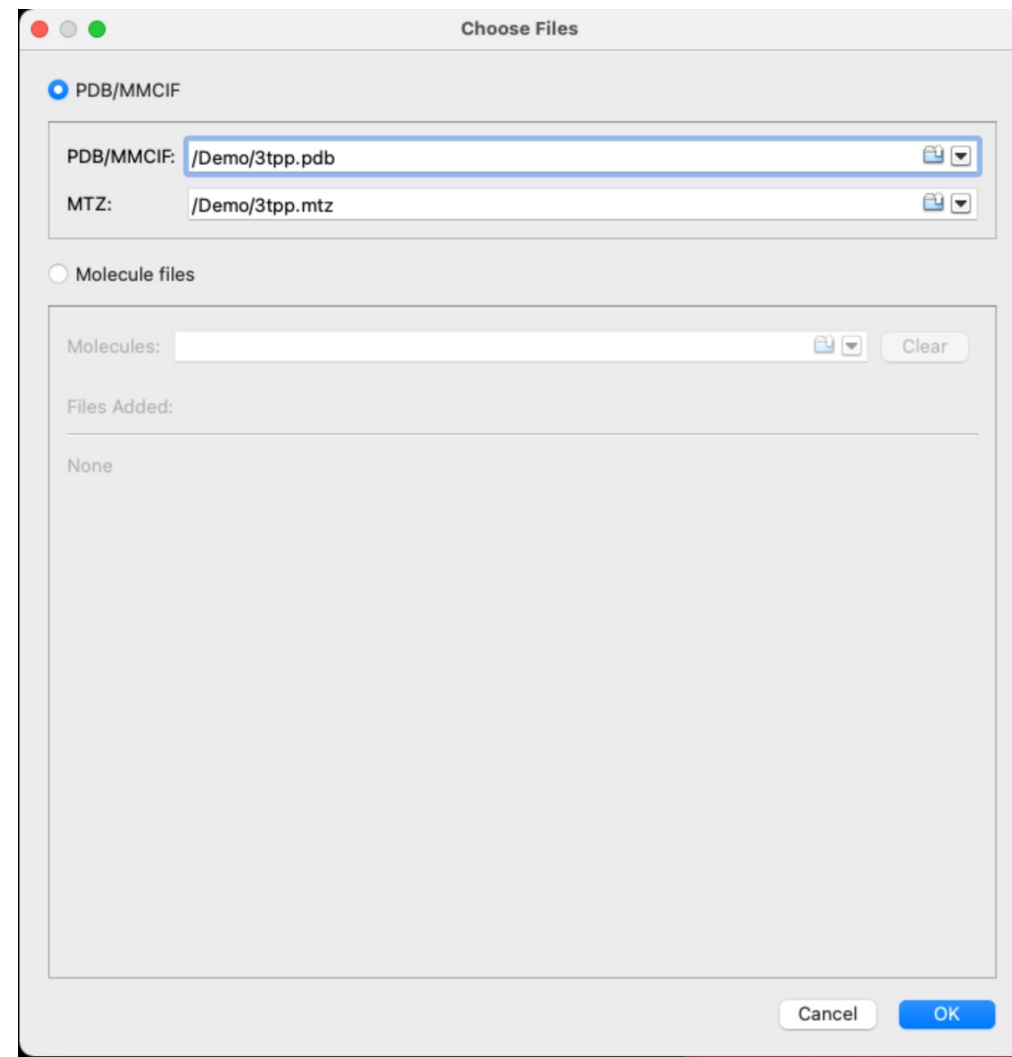
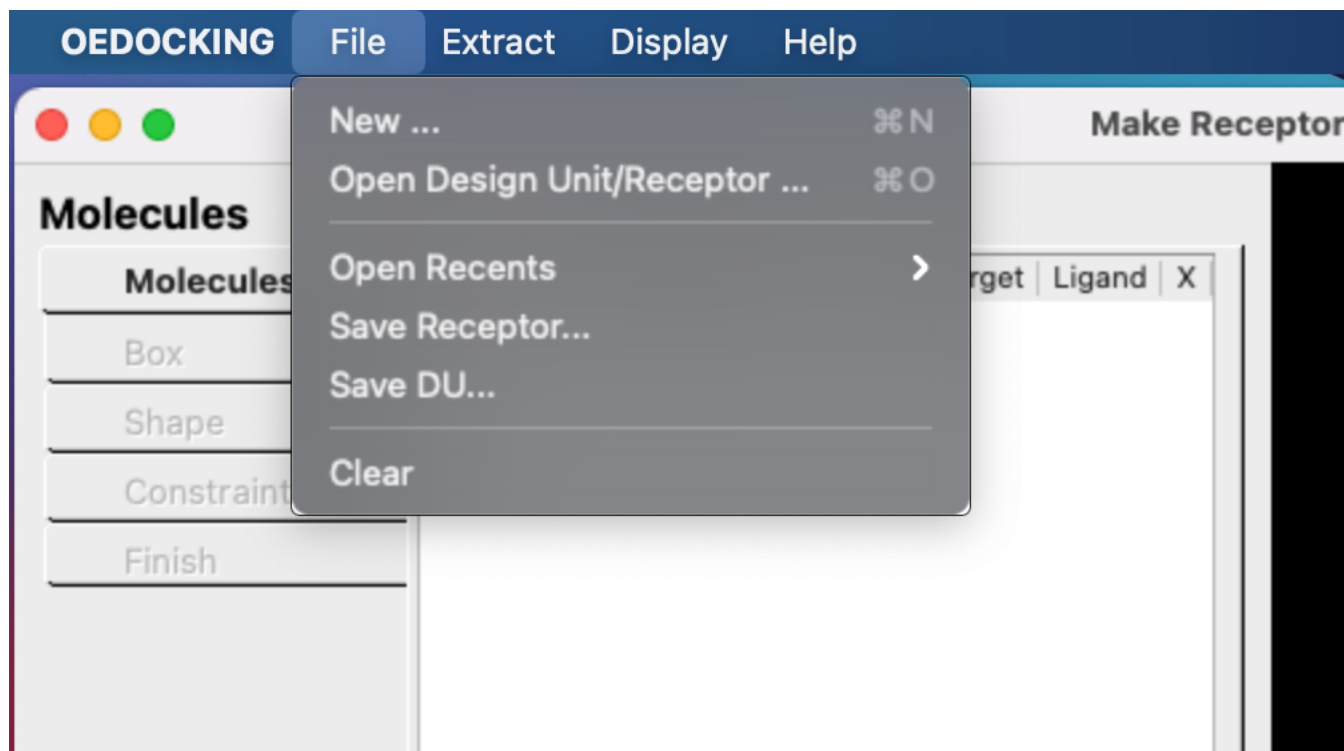


- SPRUCE prepped protein
 - Hydrogens
 - Side-chains and Loops
 - Partial Charges
- Improved pose prediction
 - Flexible-POSIT with AMBER-FF14SB and Parsley
- Better integration to downstream structure based design applications
 - SZYBKI
 - MD

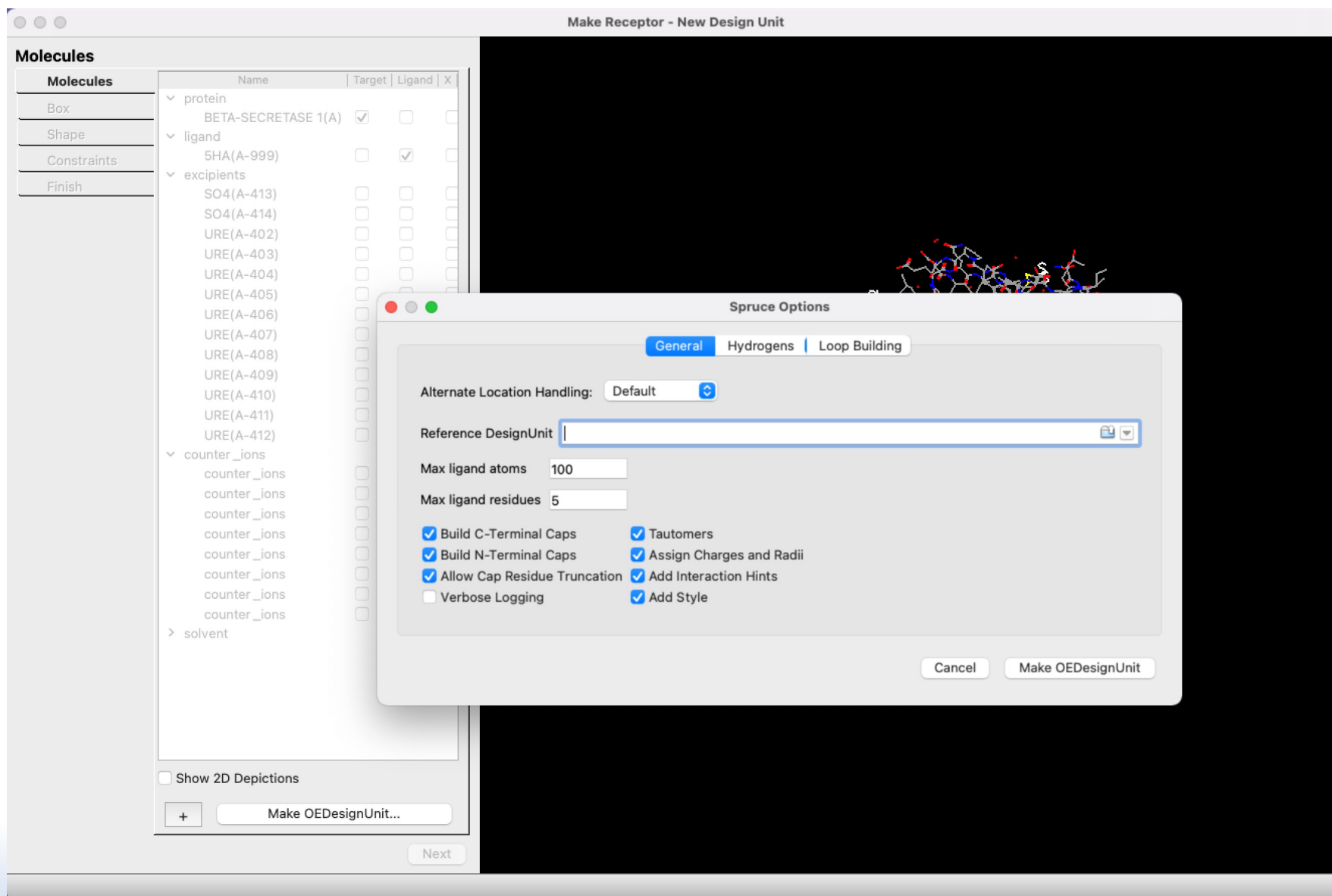
MakeReceptor

- OEDU Native
 - Runs SPRUCE to prepare structure and make design units
 - SPRUCE options available in UI
- Backwards compatibility
 - Works with existing OEB receptor files
 - Ability to choose target (protein) and bound ligand
 - Accept all molecule file formats (PDB, CIF, SDF, MOL, etc.)
 - Ability to combine molecules from multiple files
- Same look and feel, with a widget-based approach

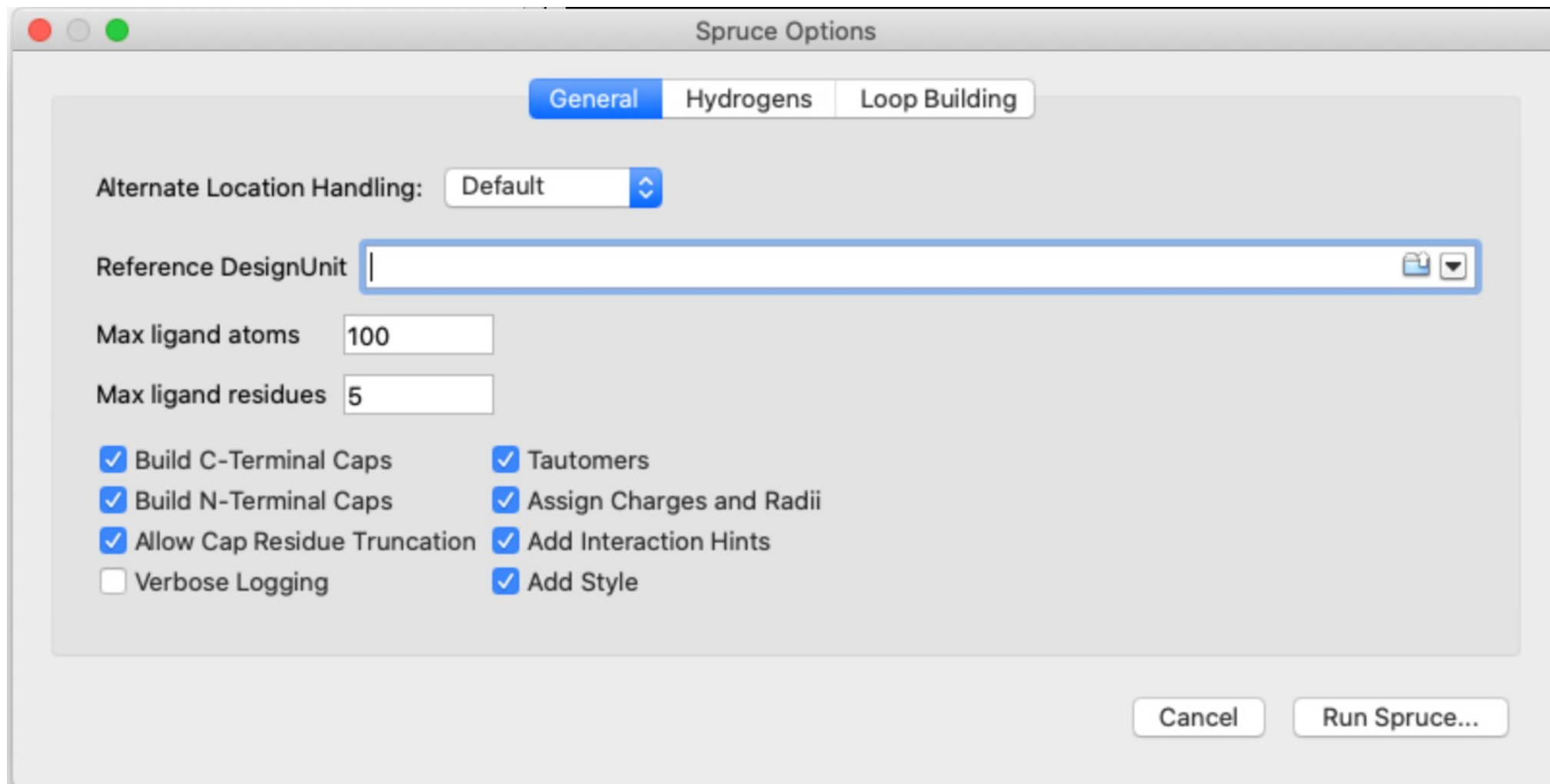
MakeReceptor – New or Open



MakeReceptor – SPRUCE GUI



MakeReceptor - SPRUCE GUI



The screenshot shows the 'Spruce Options' window with three tabs: 'General' (selected), 'Hydrogens', and 'Loop Building'. The 'General' tab contains the following settings:

- Alternate Location Handling: Default (dropdown menu)
- Reference DesignUnit: (empty text field with file selection icon)
- Max ligand atoms: 100 (text field)
- Max ligand residues: 5 (text field)
- Build C-Terminal Caps: ☒
- Build N-Terminal Caps: ☒
- Allow Cap Residue Truncation: ☒
- Verbose Logging: ☐
- Tautomers: ☒
- Assign Charges and Radii: ☒
- Add Interaction Hints: ☒
- Add Style: ☒

At the bottom right are 'Cancel' and 'Run Spruce...' buttons.

Toolkit Updates – OEDocking TK

- OEMakeReceptorOptions class controls what is included in grid generation and docking

```
mask = (OEDesignUnitComponent::Protein |  
        OEDesignUnitComponent::Solvent)
```

New OEMakeReceptorOptions

- New target predicate to select specific e.g. water molecules in solvent component of a design unit

```
pred = "HOH:365: :A|HOH:135: :A"
```

- Enabled auto-constraint perception in toolkit and ability to edit these -> Enables future constraint perception in Orion

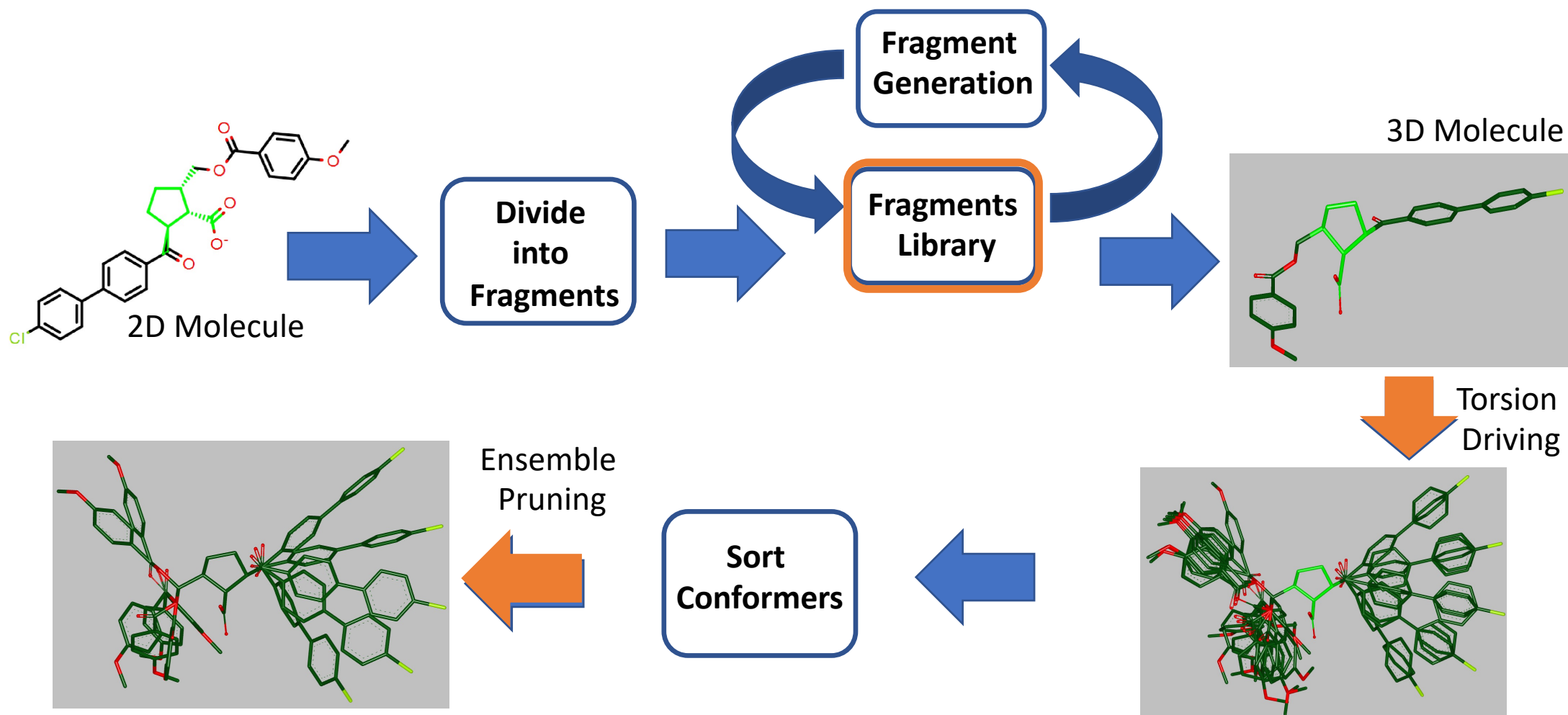
Toolkit Updates – OEBio TK

- OEReceptor class in stores the selection predicate
- OESubsetDesignUnit and OEUpdateDesignUnit will now clear the OEReceptor if the mask or predicate are matched in something that updated or removed
- OEDesignUnit class method OEGetComponents now also accepts a predicate in addition to a mask
- OEDesignUnit class has gotten two new molecule based constructors intended for prepared structures.

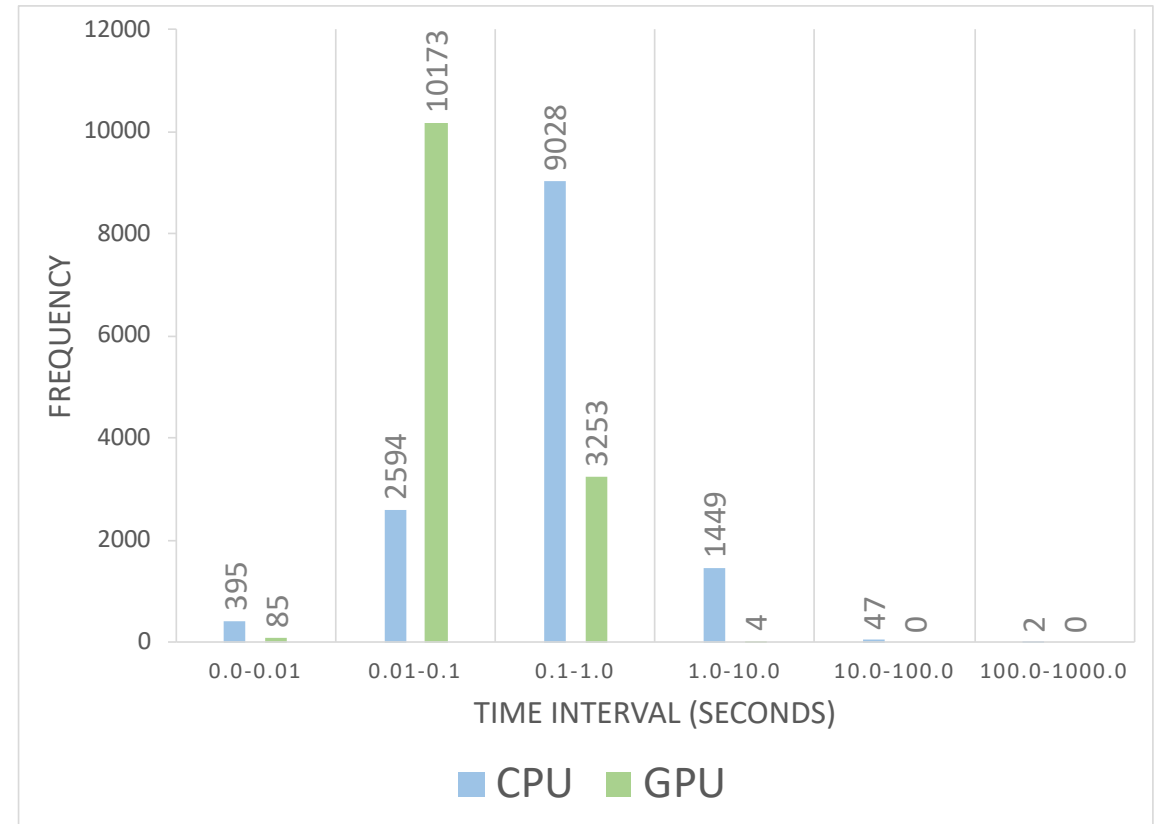
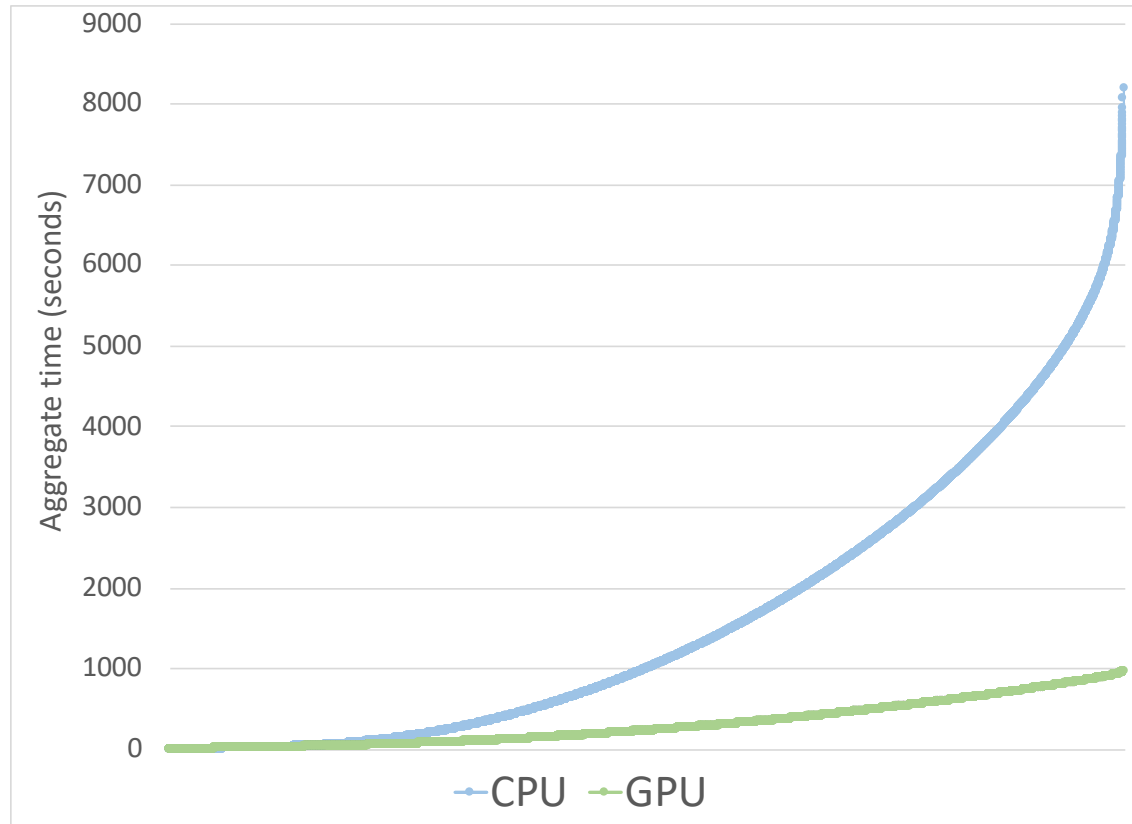
Outline

- Platform & Language Support
- Minor updates
- VIDA 5.0
- Docking receptors
- GPU Efforts

Conformer Generation With OMEGA

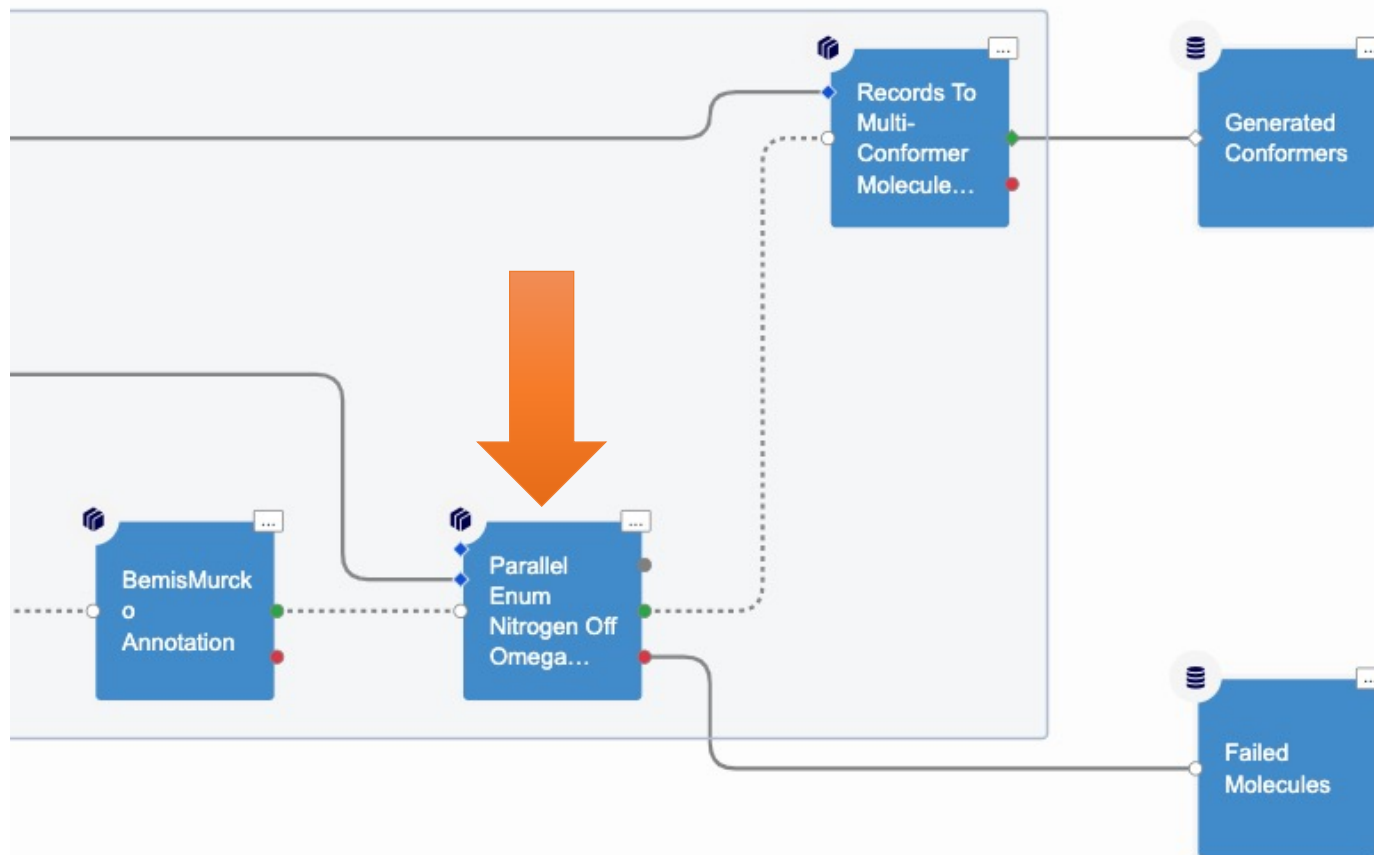


OMEGA Performance with GPU



Dataset: Almela MJ, Lozano S, Lelièvre J, Colmenarejo G, Coterón JM, Rodrigues J, *et al.* *PLoS ONE* **10**, 8 (2015)

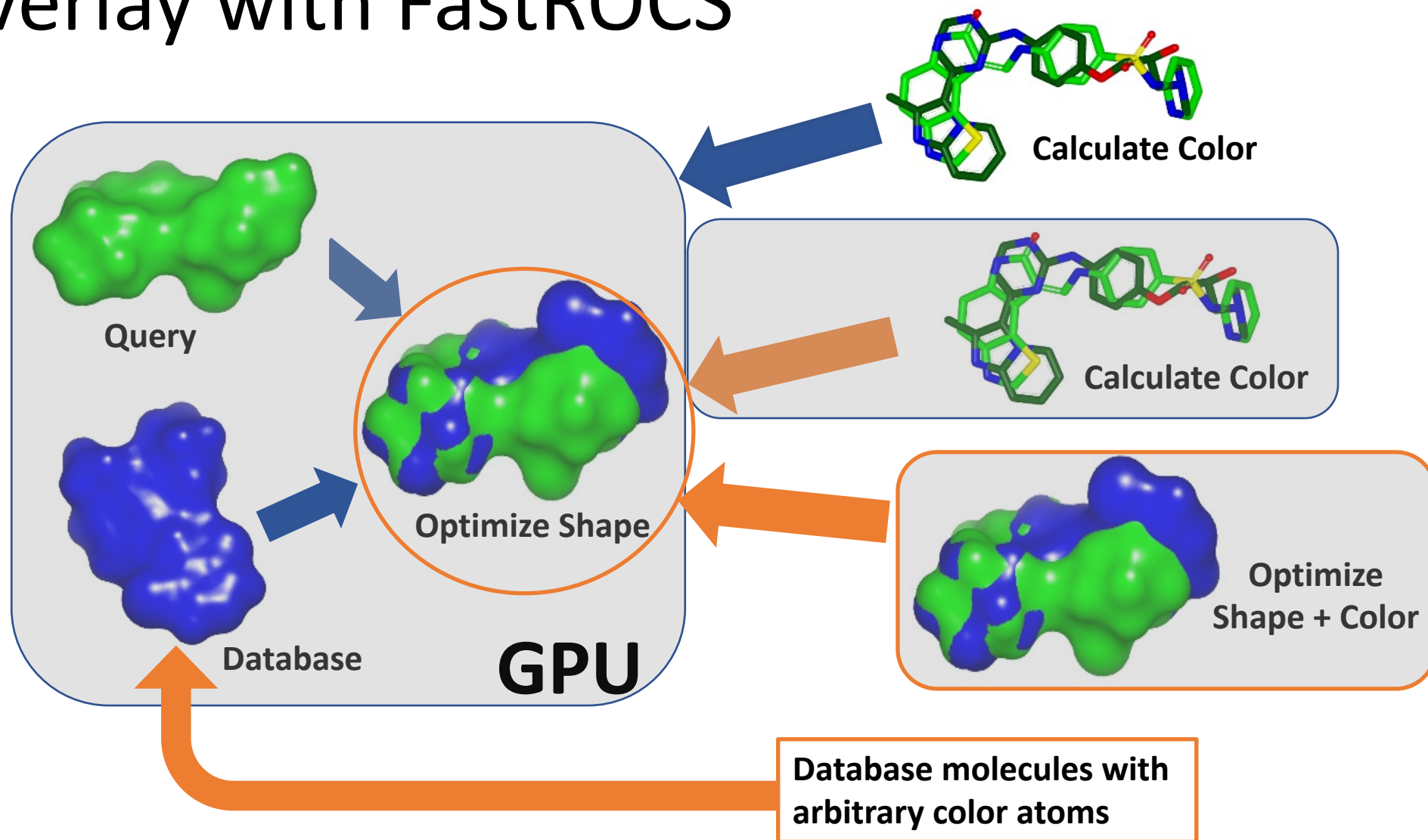
Database Prep in Orion



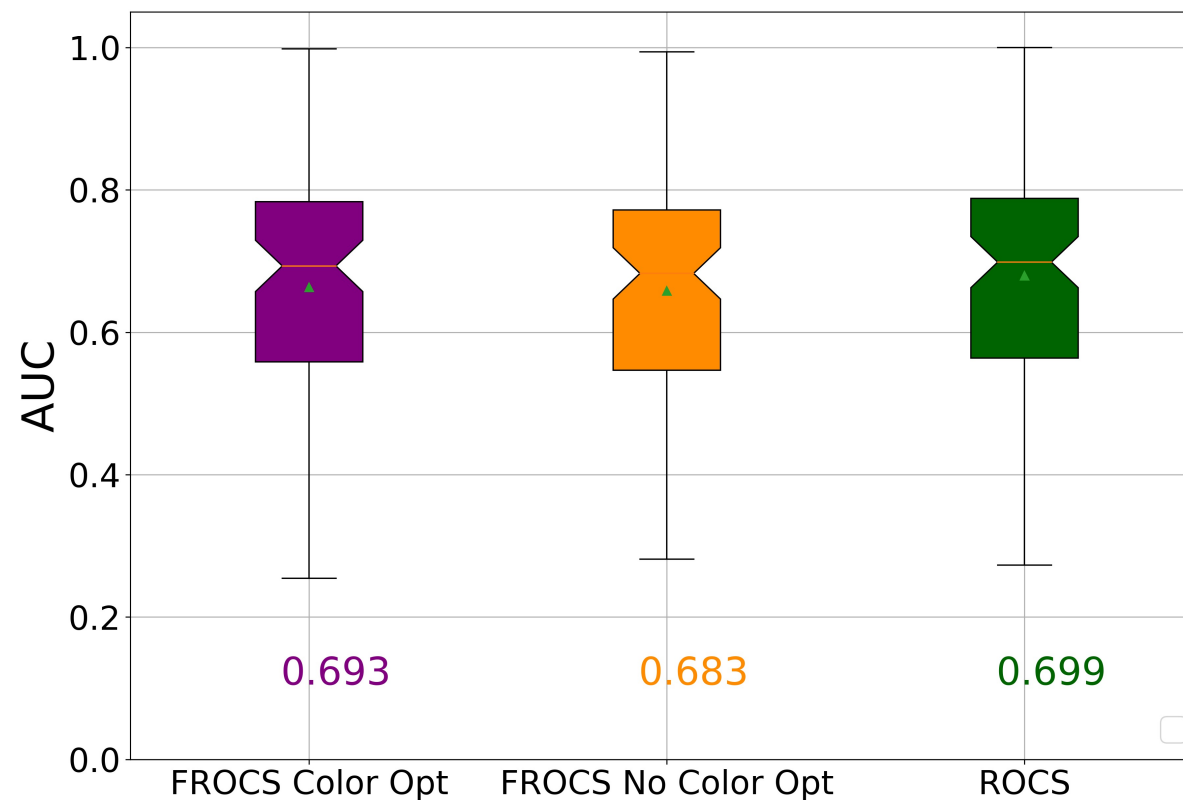
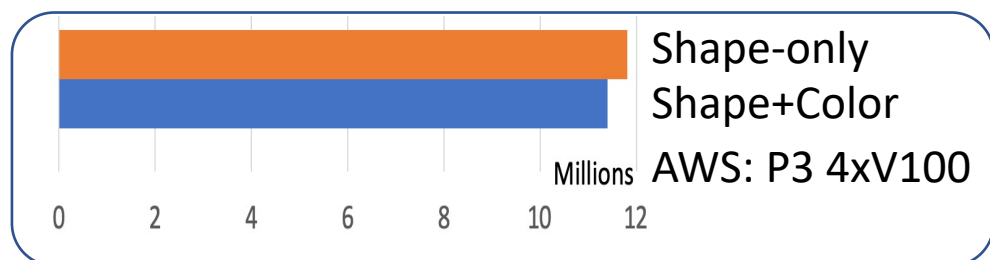
Database Prep Floe

- Multi-threaded OMEGA Cube
 - Cost of GPU Instances
 - Utilize all CPU cores
 - Utilize all GPU cores
- Database prep cost reduction

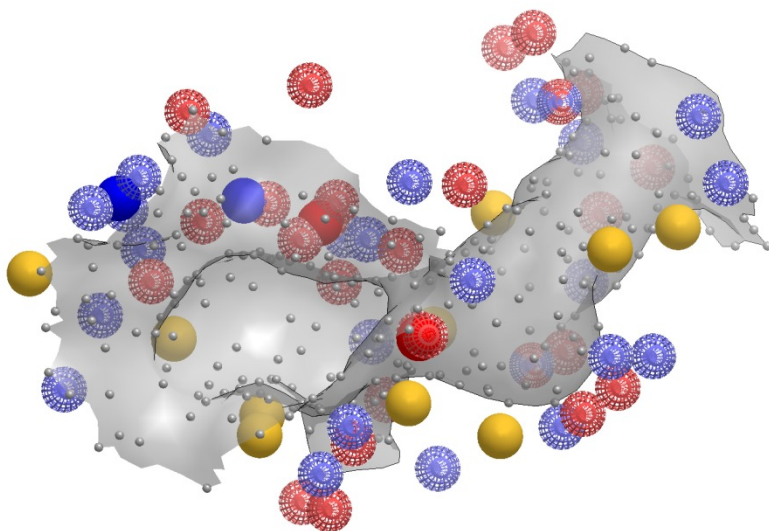
Overlay with FastROCS



FastROCS Performance and Accuracy



SiteHopper: Comparing Binding Sites



Same Protein Different Ligand

- Protein Motion
- Induced Fit
- SAR Analysis/Alignment

Closely Related Proteins

- Homology Modeling
- Selectivity

Same Ligand, Different Protein

- Drug Repurposing
- Off-target effects
- Target Fishing

Protein Pocket with Shape and Color

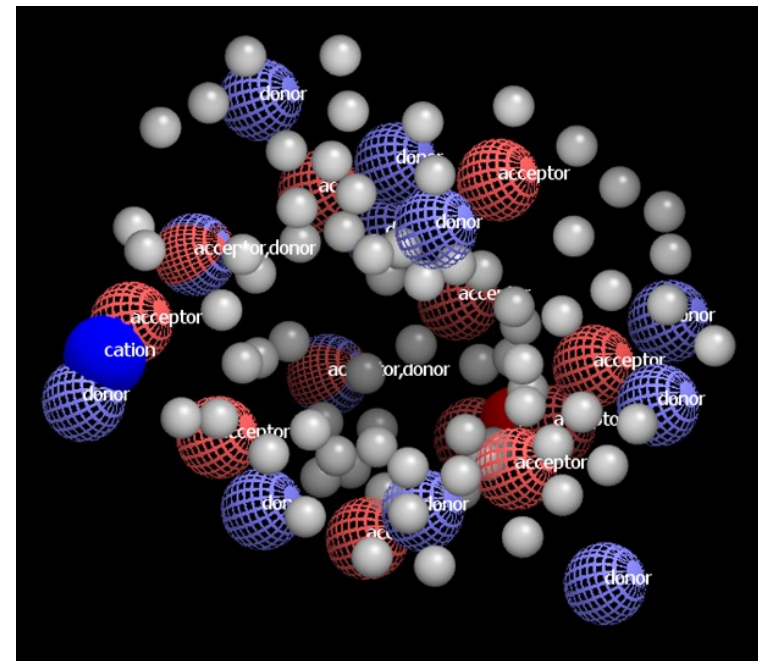
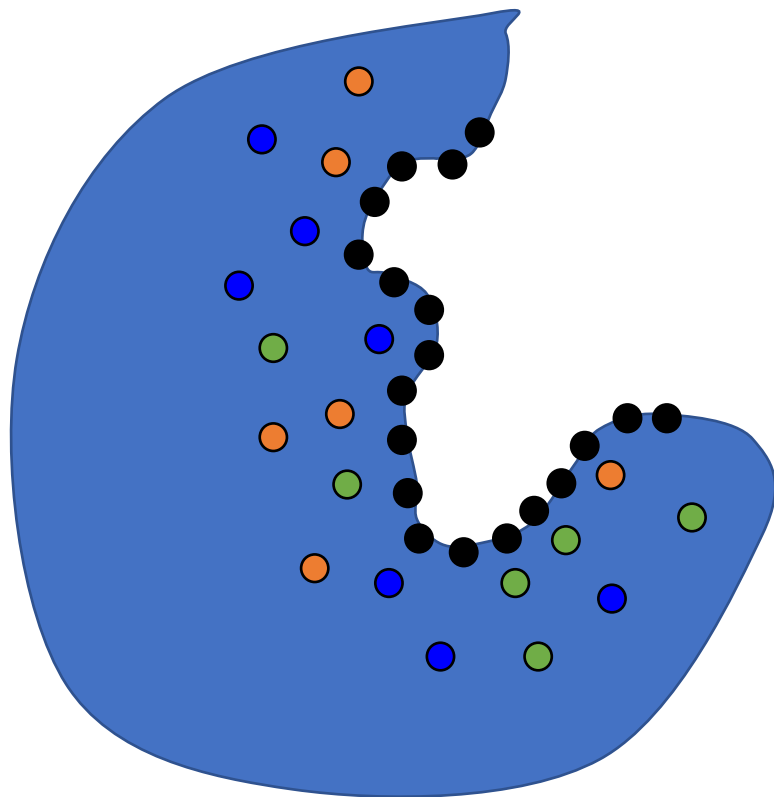
Create Pocket
Surface



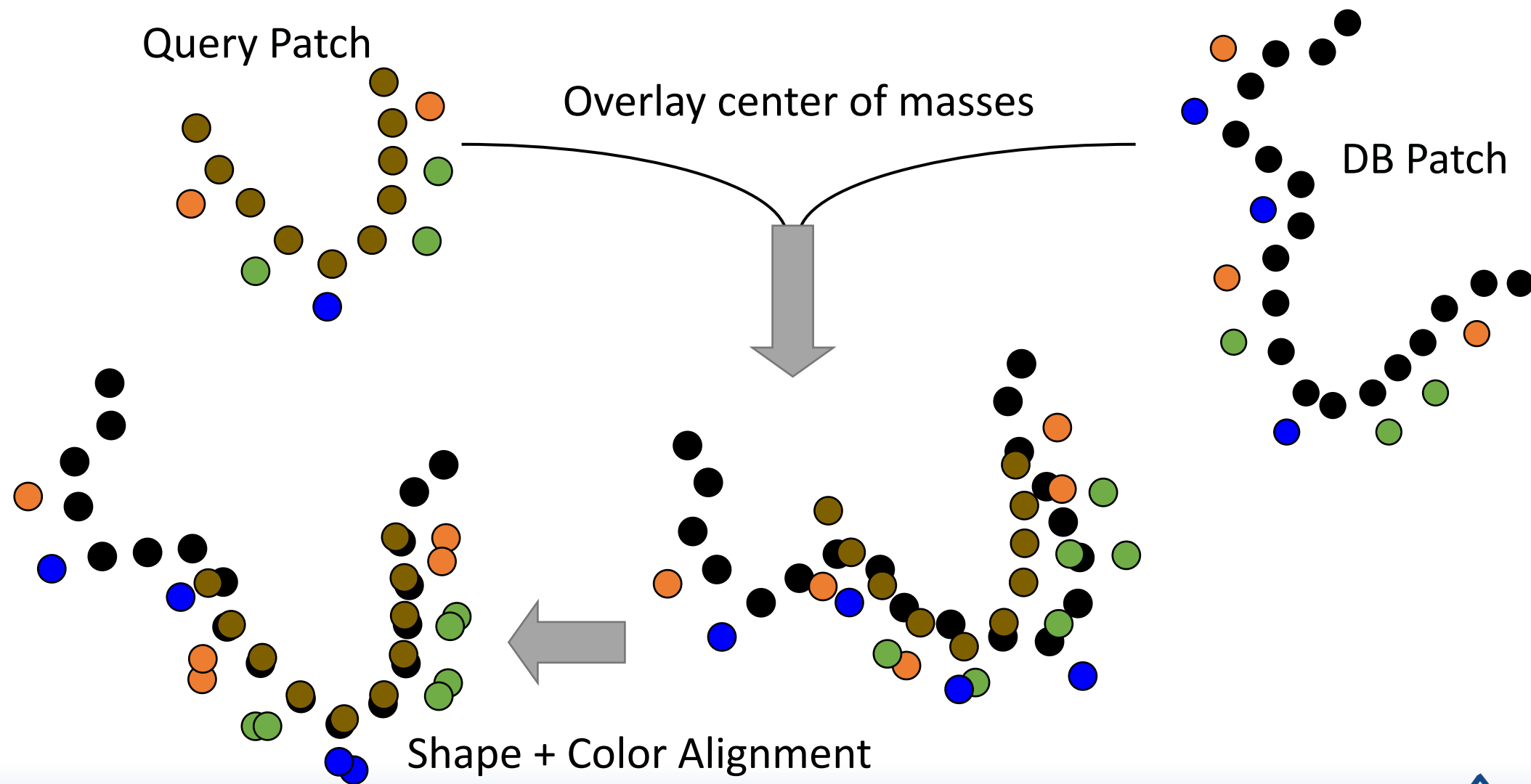
Create
Surface Shape
Molecule



Create
Protein Color
Features



Protein Pocket Overlay

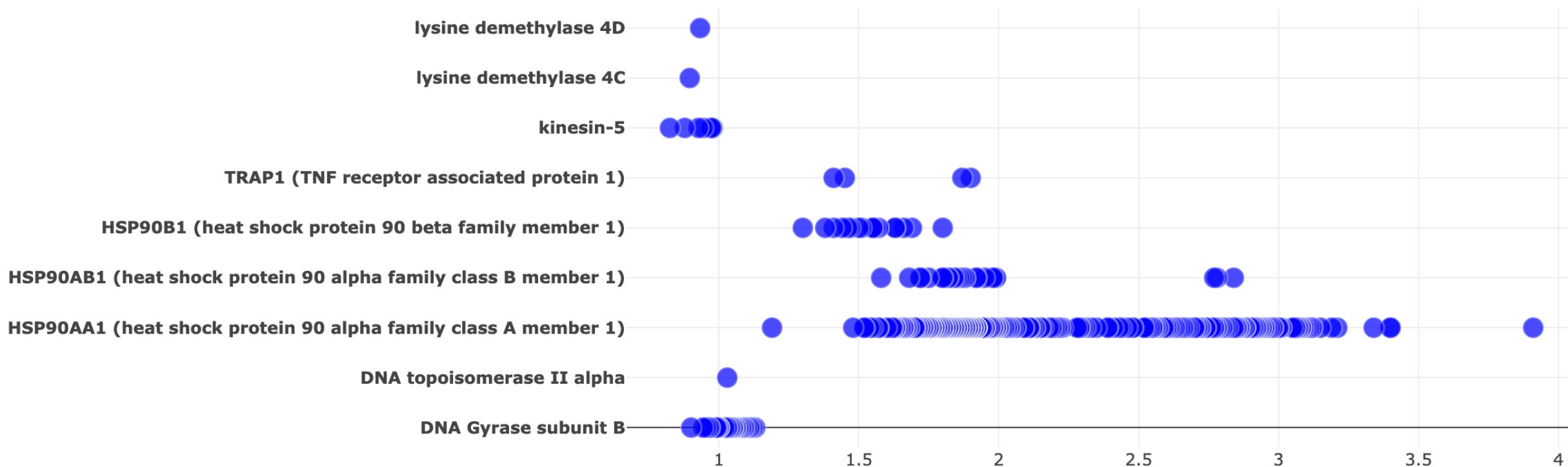


SiteHopper 1.0

- This is the first official release of SiteHopper
 - Linux only and requires an NVIDIA GPU card
- Search is done using the GPU enabling searching 250k+ protein binding sites in just a few minutes
- A database can be built from a directory of OEDesignUnit files. We are making a database available for searching based on liganded sites in the PDB (~280k binding sites).

Orion Search example

- Query HSP90, 1UYG(A) > PU2(A-1224)
- Database: Guide to Pharmacology in Orion ~40,000 DUs



App Search example

- Query HSP90, 1UYG(A) > PU2(A-1224)
- 280,000 liganded DUs from the PDB
- Top 200 hits

App Search example – best hit (self)

VIDA 5.0.0 for OpenEye (Worldwide)

Focused Visible Marked All

List Window

Name	ID	#
1UYG_A__DU__PU2_A-1224.o...	8	1
1UYG(A) > PU2(A-1224)	9	5
protein:1UYG(A)altA	1	
ligand:PU2(A-1224)	2	
solvent	3	
packing_residues	4	
OEReceptor	5	2
sitehopper_hits.oedu	1026	200
1UYG(A) > PU2(A-1224)	1027	4
protein:1UYG(A)altA	10	
ligand:PU2(A-1224)	11	
solvent	12	
binding_site	13	
1UYH(A) > PU0(A-1224)	1028	4
5FNC(AB) > IEE(B-1225)	1029	6
5LR1(A) > 72Y(A-4000)	1030	4
2H55(A) > DZ8(A-1001)	1031	4
protein:2H55(A)altA	34	
ligand:DZ8(A-1001)	35	
solvent	36	
binding_site	37	
1UYK(A) > PUX(A-1224)	1032	4
1UYF(A) > PU1(A-1224)	1033	4

Spreadsheet

Molecule	VIDA ID	Patch Sco...	Patch Shape Sc...	Patch Color Sc...
protein:1UYG(A)altA	10	3.9500	0.9500	3.0000

4

Molecules Proteins Atoms Residues

App Search example – top hits

VIDA 5.0.0 for OpenEye (Worldwide)

Focused Visible Marked All

List Window

Name	ID	#
1UYG_A_DU__PU2_A-1224.o...	8	1
1UYG(A) > PU2(A-1224)	9	5
> protein:1UYG(A)altA	1	
> ligand:PU2(A-1224)	2	
> solvent	3	
> packing_residues	4	
> OEReceptor	5	2
sithopper_hits.oedu	1026	200
1UYG(A) > PU2(A-1224)	1027	4
> protein:1UYG(A)altA	10	
> ligand:PU2(A-1224)	11	
> solvent	12	
binding_site	13	
> 1UYH(A) > PU0(A-1224)	1028	4
> 5FNC(AB) > IEE(B-1225)	1029	6
> 5LR1(A) > 72Y(A-4000)	1030	4
2H55(A) > DZ8(A-1001)	1031	4
> protein:2H55(A)altA	34	
> ligand:DZ8(A-1001)	35	
> solvent	36	
binding_site	37	
> 1UYK(A) > PUX(A-1224)	1032	4
> 1UYF(A) > PU1(A-1224)	1033	4

Spreadsheet

Molecule	VIDA ID	Patch Sco...	Patch Shape Sc...	Patch Color Sc...
protein:2H55(A)altA	34	3.0900	0.6600	2.4300

Molecules Proteins Atoms Residues

App Search example – hit #200

VIDA 5.0.0 for OpenEye (Worldwide)

Focused Visible Marked All

List Window

Name	ID	#
> 2QF6(A) > A56(A-256)	1209	4
> 3B27(A) > B2T(A-1)	1210	4
> 4LWH(A) > FJ5(A-301)	1211	4
> 3HZ1(AB) > 42C(B-237)	1212	6
> 2YE8(AB)altA > 2D3(B-1225)	1213	6
> 6LT8(A) > ET3(A-301)	1214	4
> 3WHA(A)altA > WHA(A-302)	1215	5
> 4LWF(A) > FJ3(A-301)	1216	4
> 2YE8(AB)altB > 2D3(A-1225)	1217	6
> 3HZ1(AB) > 42C(A-237)	1218	6
> 3HHU(B) > 819(B-501)	1219	4
> 4EEH(A) > HH6(A-301)	1220	5
> 5J6N(AB) > 6FF(B-301)	1221	5
> 2WI5(A) > ZZ5(A-1225)	1222	4
> 2QF6(D) > A56(D-256)	1223	4
> 3OWD(A) > MEY(A-1)	1224	4
> 2XDU(A) > MT0(A-1228)	1225	6
> 4YKR(A) > 4EP(A-301)	1226	5
> protein:4YKR(A)altA	1018	
> ligand:4EP(A-301)	1019	
> solvent	1020	
> excipients	1021	3
> binding_site	1025	

protein:4YKR(A)altA

ligand:4EP(A-301)

solvent

excipients

binding_site

Spreadsheet

Molecule	VIDA ID	Patch Sco...	Patch Shape Sc...	Patch Color Sc...
protein:4YKR(A)altA	1018	1.8800	0.4000	1.4800

304

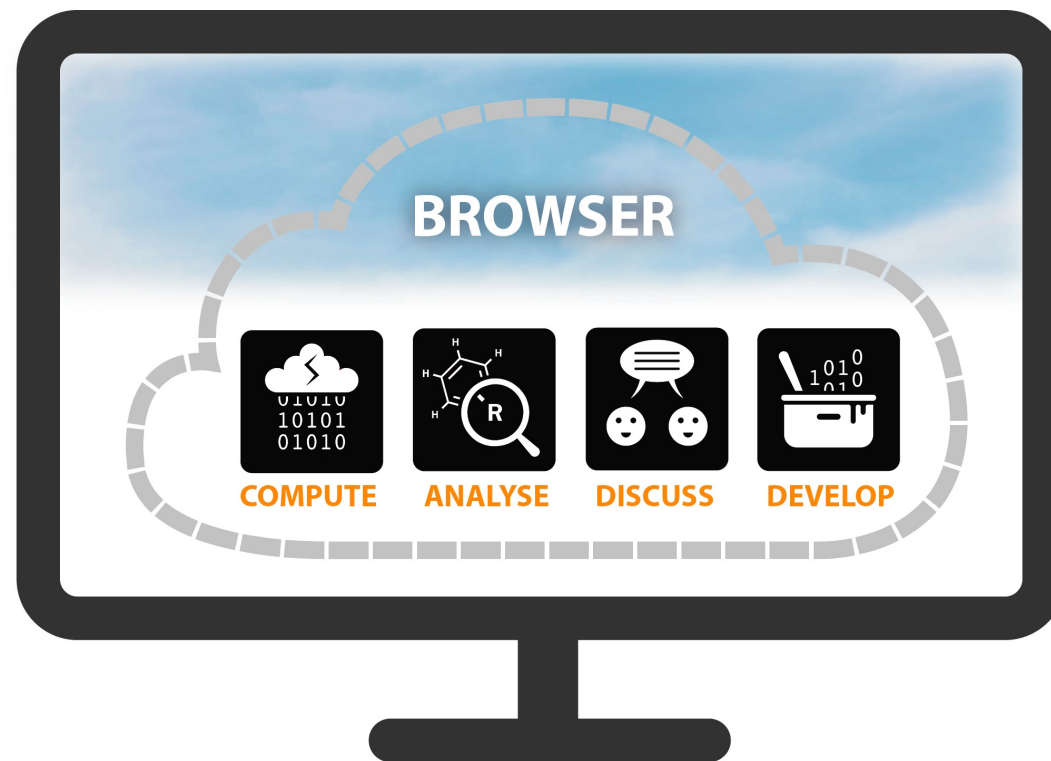
Molecules Proteins Atoms Residues

Summary

- Platform & Language Support
- Minor updates
- VIDA 5.0
- Docking receptors
- GPU Efforts



OpenEye Toolkits 2021.1.0
will be available in the
2nd Orion release of 2021



For more information, please contact:

sales@eyesopen.com | info@eyesopen.com

www.eyesopen.com

+1-505-473-7385

A full-page background image of a starry night sky. The Milky Way galaxy is visible as a dense band of stars stretching across the frame. In the bottom right corner, the dark silhouette of a person is shown standing and looking up at the stars.

Thank You

The End