2.1.4 Tour: Virtual screening of FDA-approved drugs against HIV protease

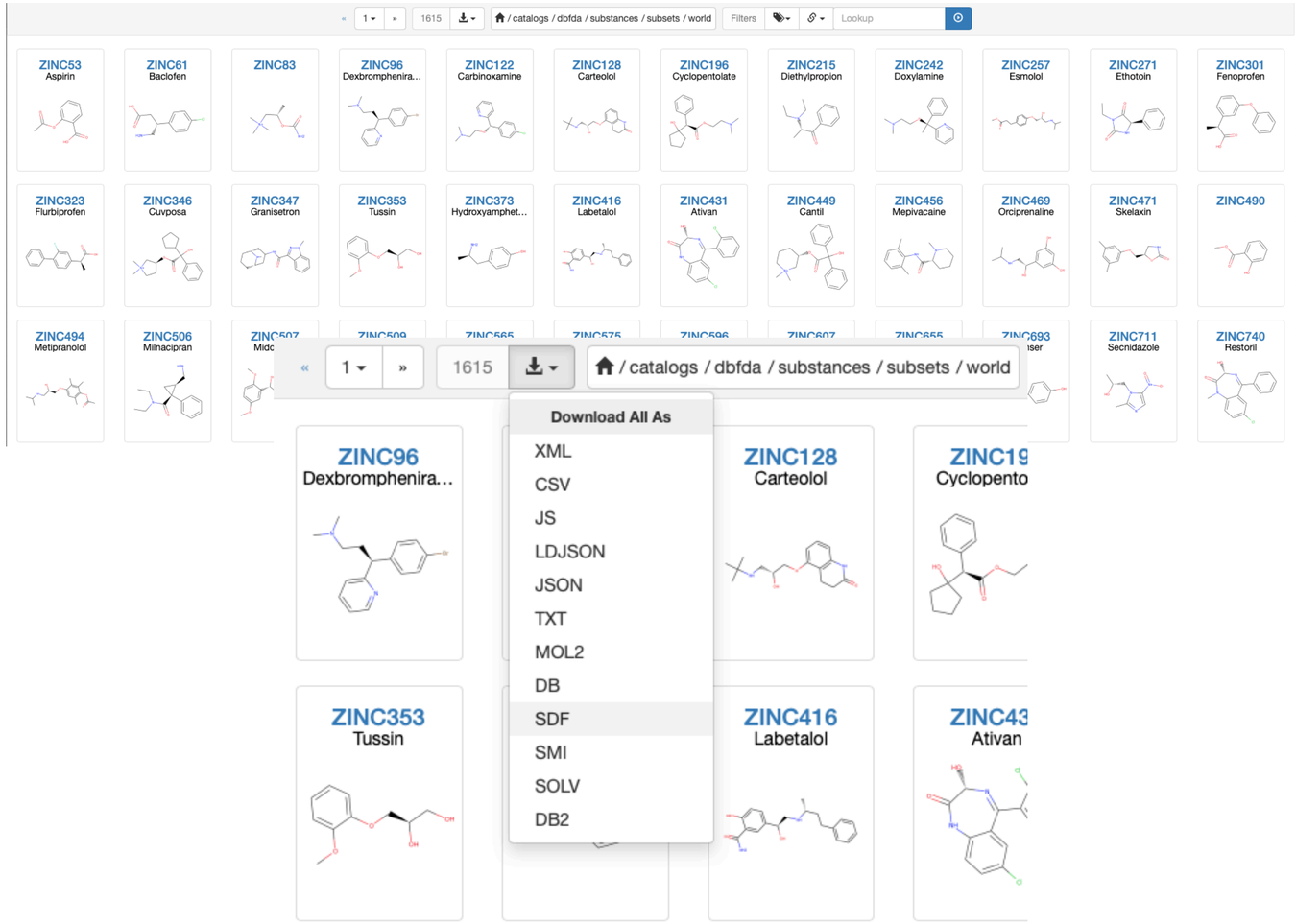
- This module will be a tour of a virtual screening calculation
- The tour should help you perform your own virtual screening

Overview

- In terms of clinical impact, there are many advantages of docking an existing drug against a new target
  - Clinical trials have already been performed
    - Demonstrated safety
    - Understanding of pharmacokinetics
  - Even before a new clinical trial, doctors can prescribe an existing drug for off-label use
- I will describe how I docked the FDA approved drugs against HIV protease
  - Protease is not a new target, but the same procedure can be followed for a new target
  - The steps include downloading the library, converting to AutoDock's ligand format, transferring files to XSEDE Bridges, submitting a job that runs AutoDock Vina, transferring files back to my computer, and performing some analysis
  - Files for this tour are on GitHub

Preparing a chemical library

- First, I went to the ZINC15 web site and downloaded all substances in the “DrugBank FDA only” catalog in SDF format. It was a 3.8 MB file.



http://zinc15.docking.org/catalogs/dbfda/substances/subsets/world/

File formatting

- ZINC provides the library in a file format that AutoDock Tools is unfamiliar with, SDF
- To convert to the format that AutoDock uses, I used Open Babel with the command
  - `obabel dbfda-world.sdf -O dbfda.pdbqt -m’
  - -m means that the molecule is split into multiple files
- This generated 1657 pdbqt files in the same directory

Virtual Screening Scripts

- I wrote a few scripts to manage the virtual screening
  - sync_virtual_screen.sh, to transfer files back and forth between my computer and XSEDE's Bridges
  - create_vina_sh.py, a python script to create a shell script, script0.sh, to run vina on every ligand file in a directory and use a specified number of cores
  - vina_multithread.job, a SLURM batch script to
    - run create_vina_sh.py based on the number of cores that SLURM provides
    - run the resulting shell script0.sh

```
Minh-IIT-MBP2018:[~/Documents/GitHub/Chem456/static_files/tutorials/hivpr-dockin
g]: more sync_virtual_screen.sh
rsync -Cuavz virtual_screen/ dminh@bridges.psc.xsede.org:~/virtual_screen/
rsync -Cuavz dminh@bridges.psc.xsede.org:~/virtual_screen/ virtual_screen/
```

Virtual Screening Procedure

- First, I transferred the files to Bridges using sync_virtual_screen.sh
- Then, I logged onto Bridges and executed the command `sbatch vina_multithread.job'
- Next, I transferred the results to my own computer using sync_data.sh
- Finally, I performed some preliminary analysis on an ipython notebook, AnalyzeVS.ipynb, also exported to html format.