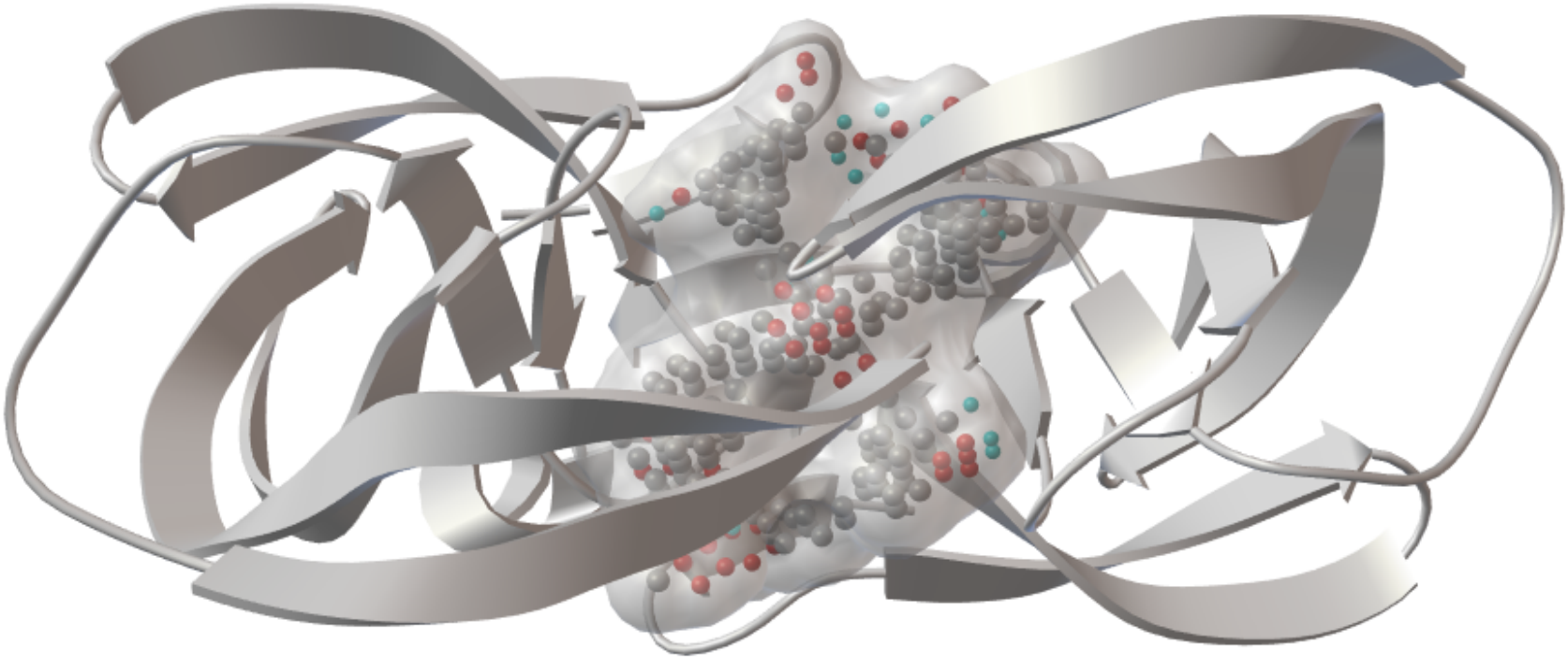
AutoLigand:
A Tool for Finding Ligand Binding Sites

The Croucher Foundation Advanced Studies Institute
December 14-19 2009 - CUHK, Hong Kong

Stefano Forli and Rodney M. Harris
The Scripps Research Institute

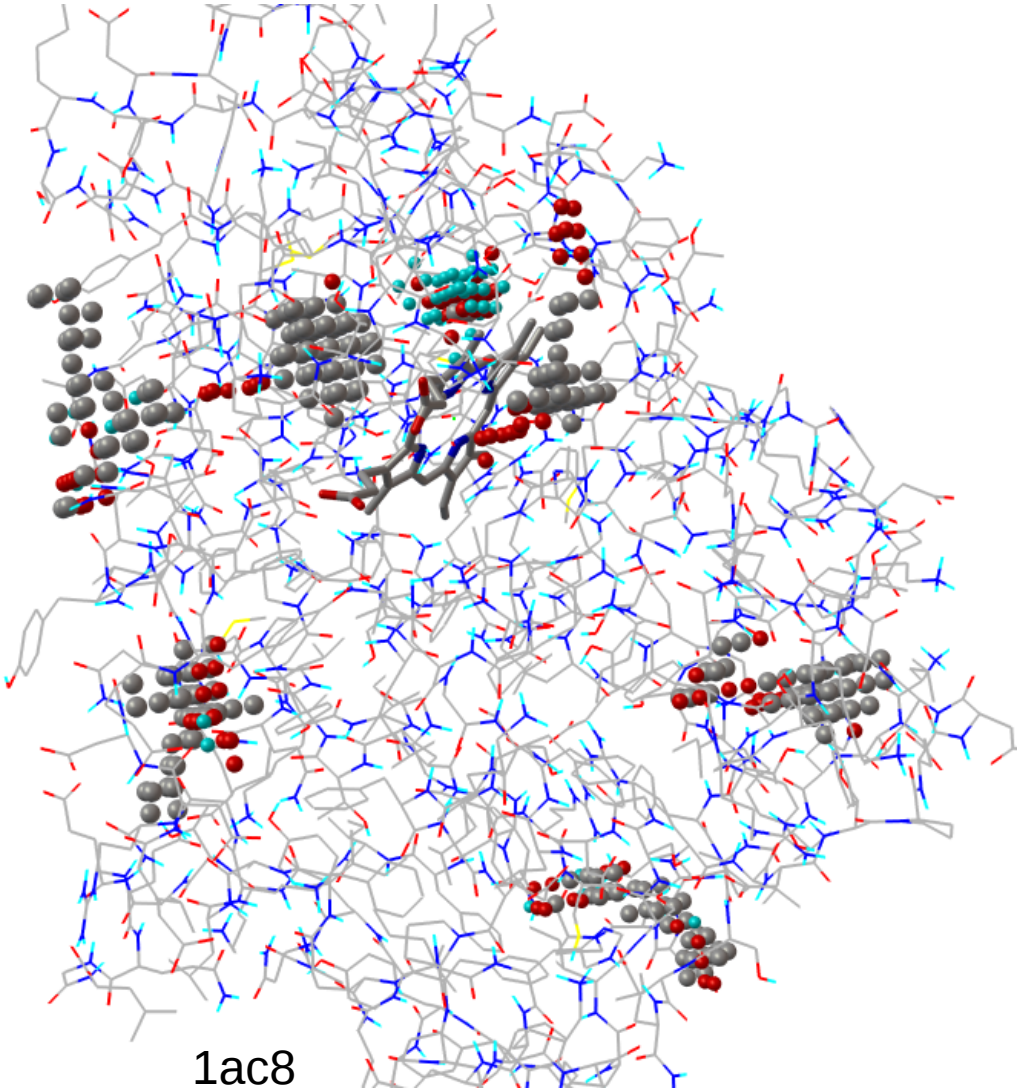


Overview

- Description and uses of AutoLigand
- Algorithms (how it works)
  - Flood fill
  - Optimization
  - Local minimum check
- Statistics and scoring
- Results

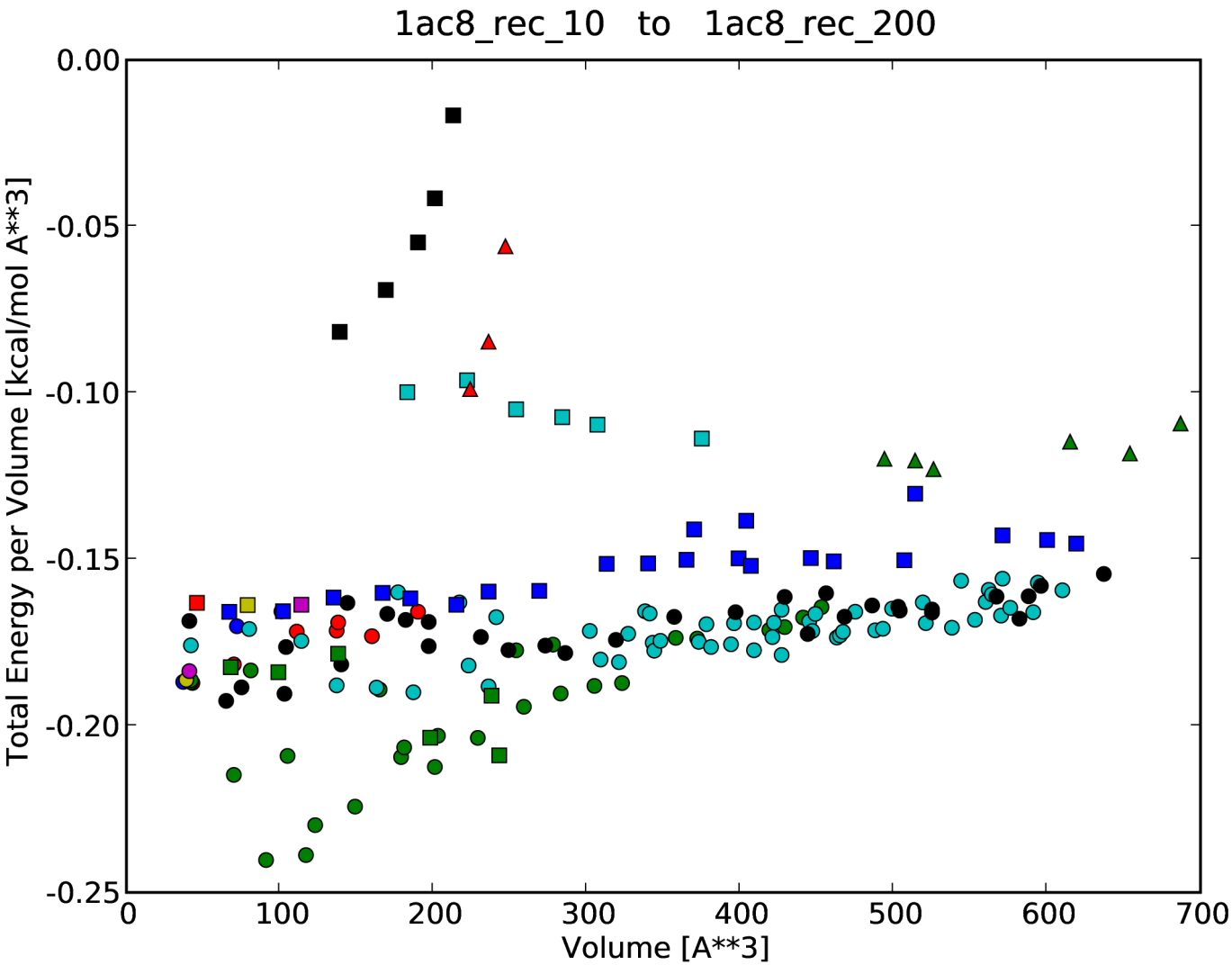
AutoLigand Uses

- Find unknown sites
  - scan complete receptor molecule
  - report top ten hits
- Optimize known site
  - use energy per volume of fill
  - identify atom types

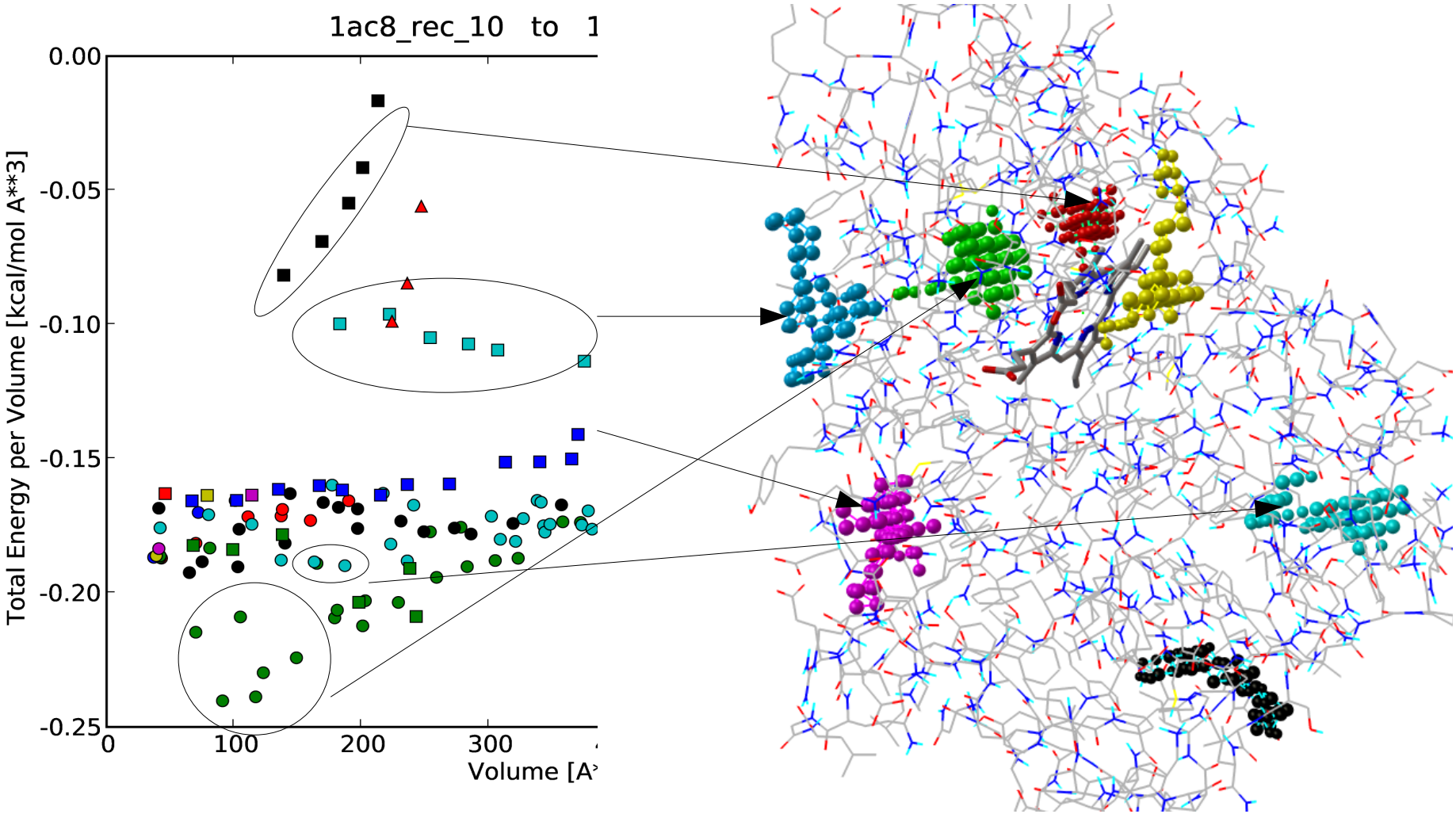


CYTOCHROME C PEROXIDASE

Plot of energy/volume vs volume



Finding the best binding site



Results – HIV Protease

Amprenavir (black)
HIV Protease (CPK colors)

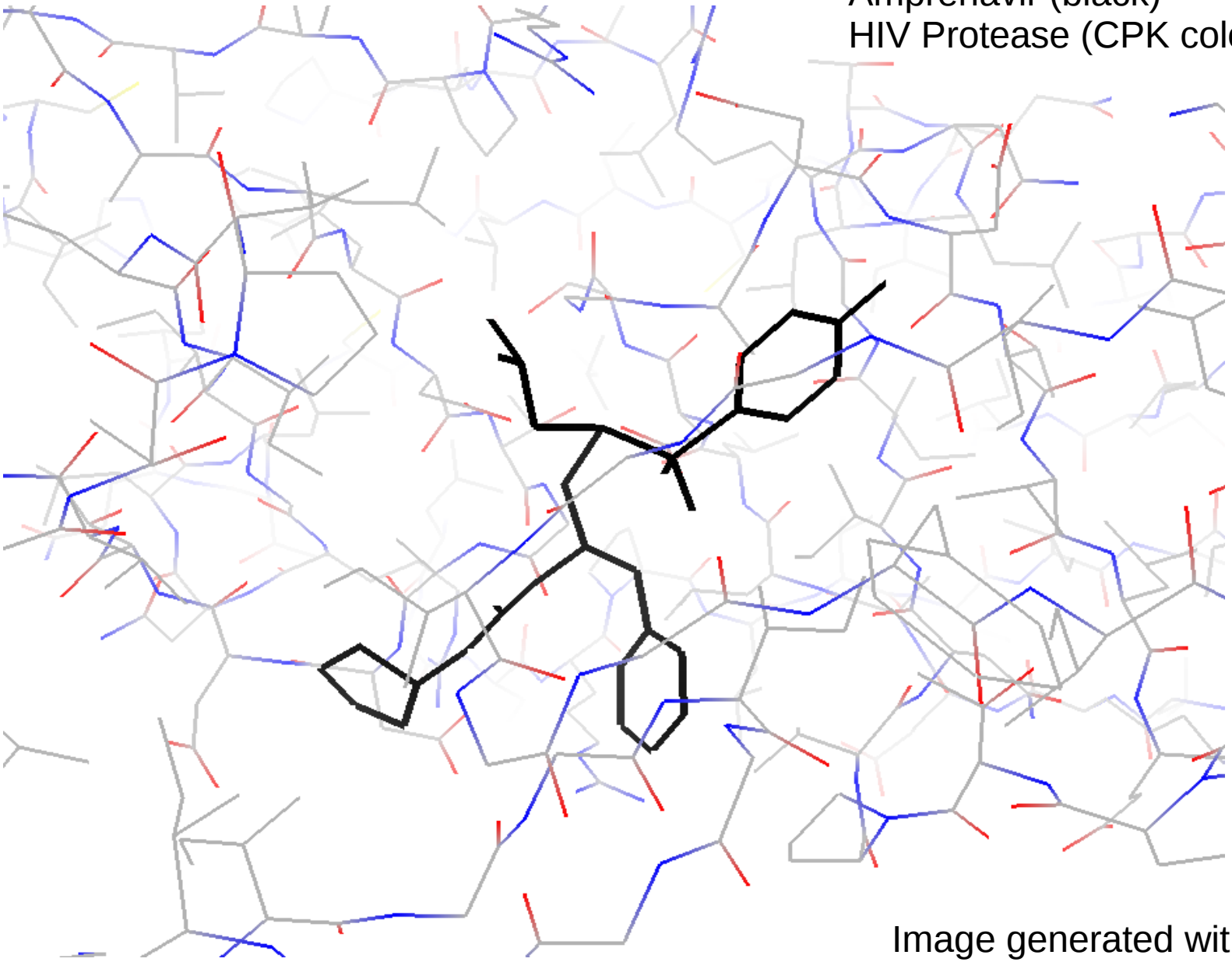


Image generated with PMV 1.4

Fill Example – HIV Protease

50 points from AutoLigand

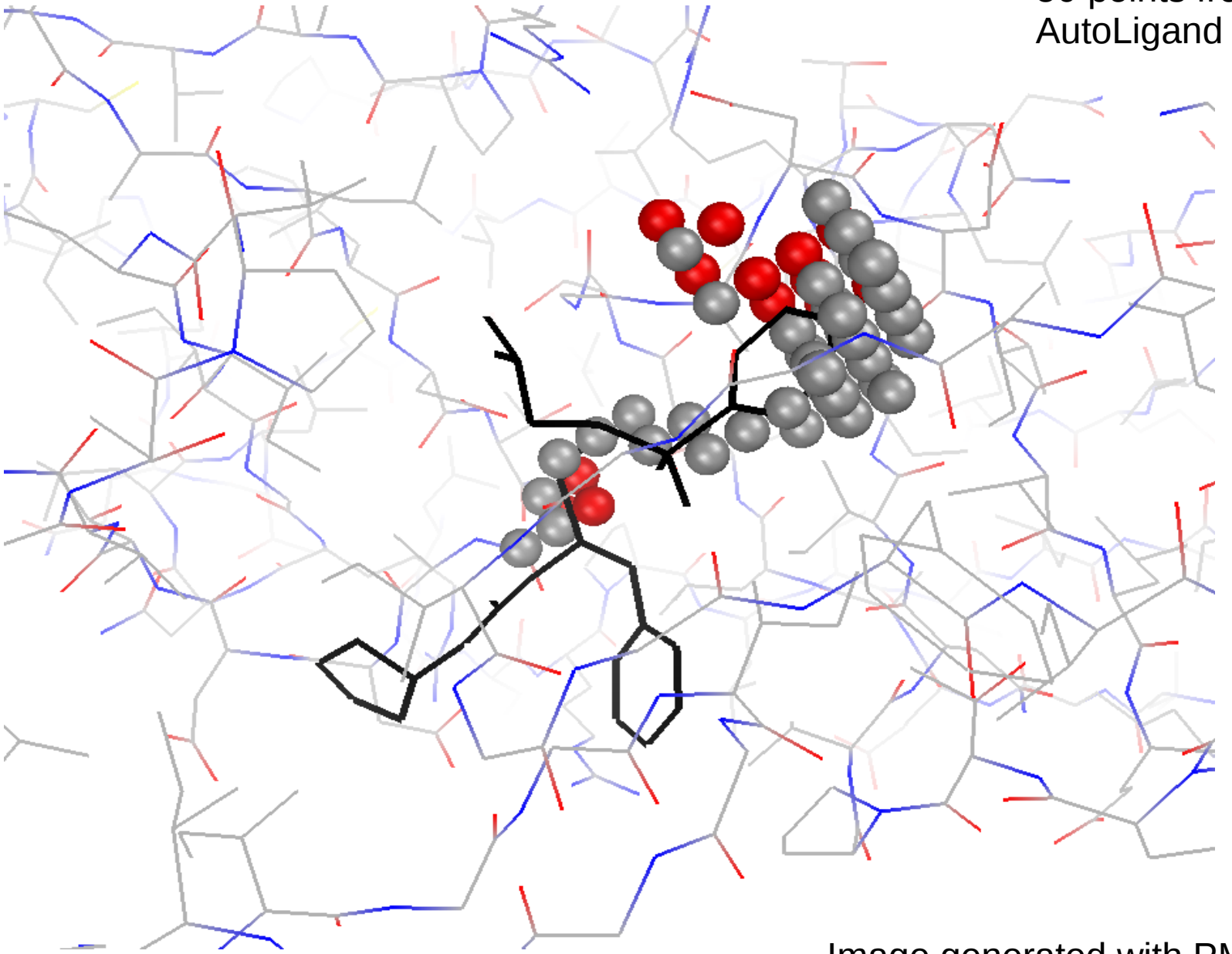


Image generated with PMV 1.4

Fill Example – HIV Protease

100 points from AutoLigand

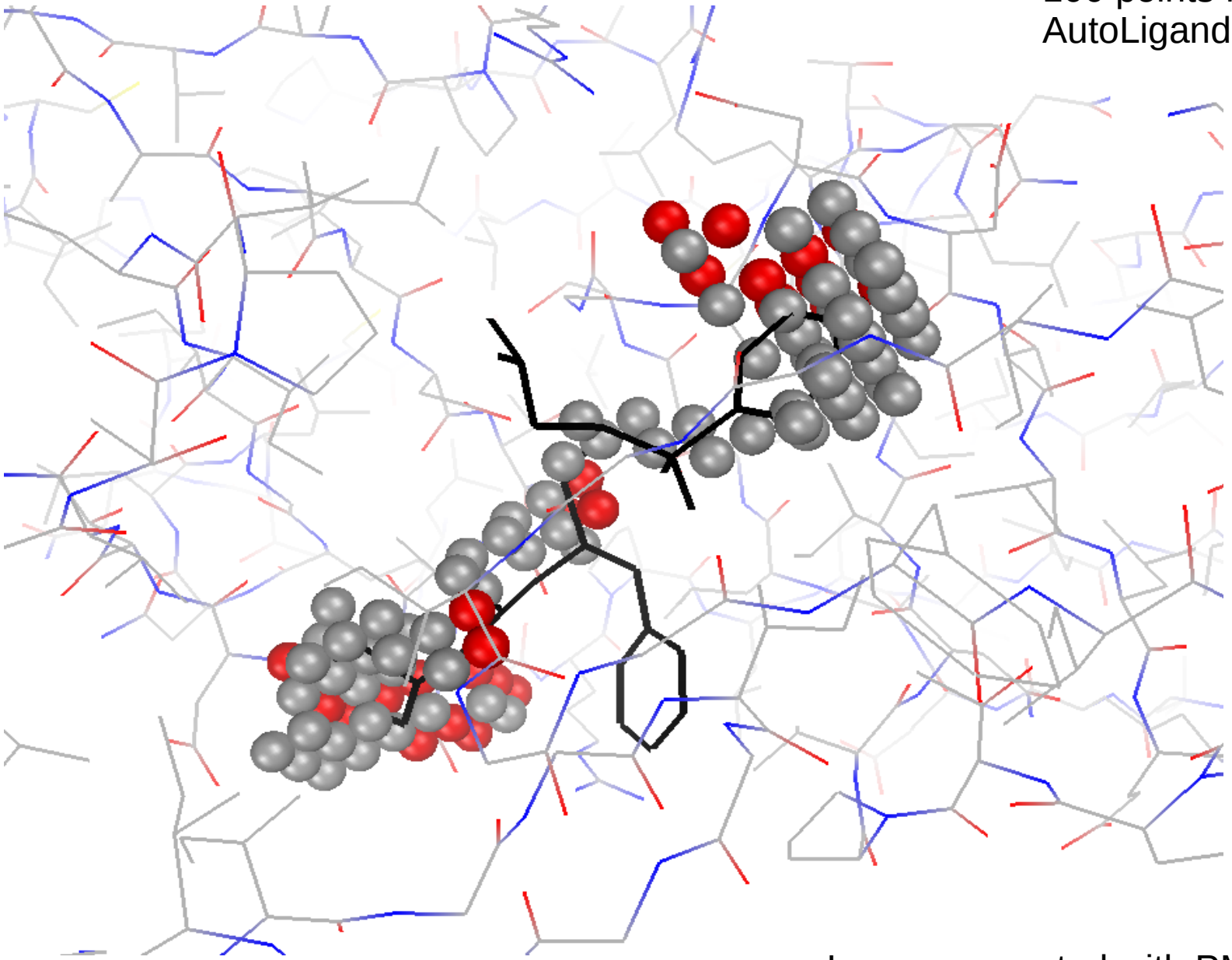


Image generated with PMV 1.4

Fill Example – HIV Protease

150 points from AutoLigand

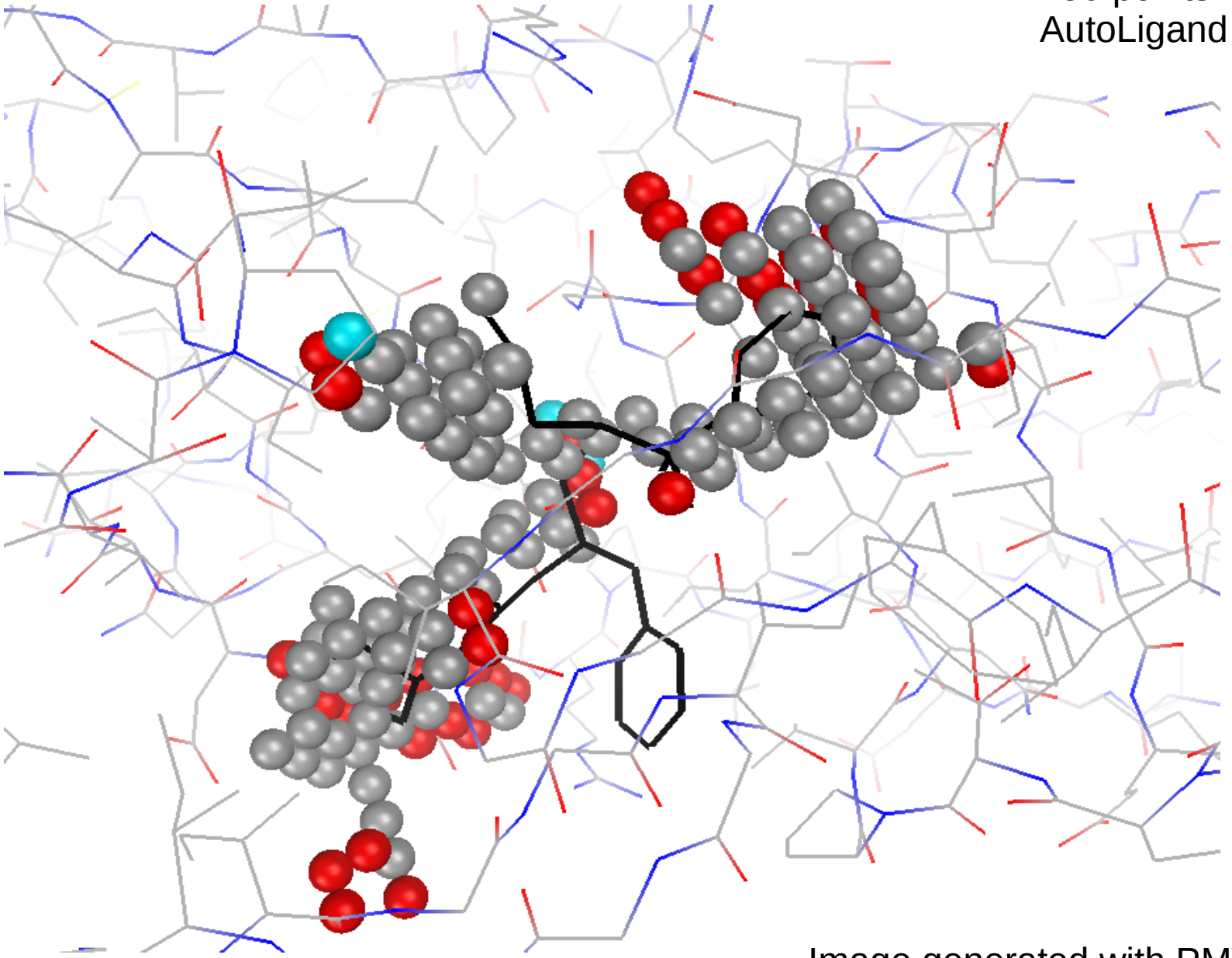


Image generated with PMV 1.4

Fill Example – HIV Protease

200 points from AutoLigand

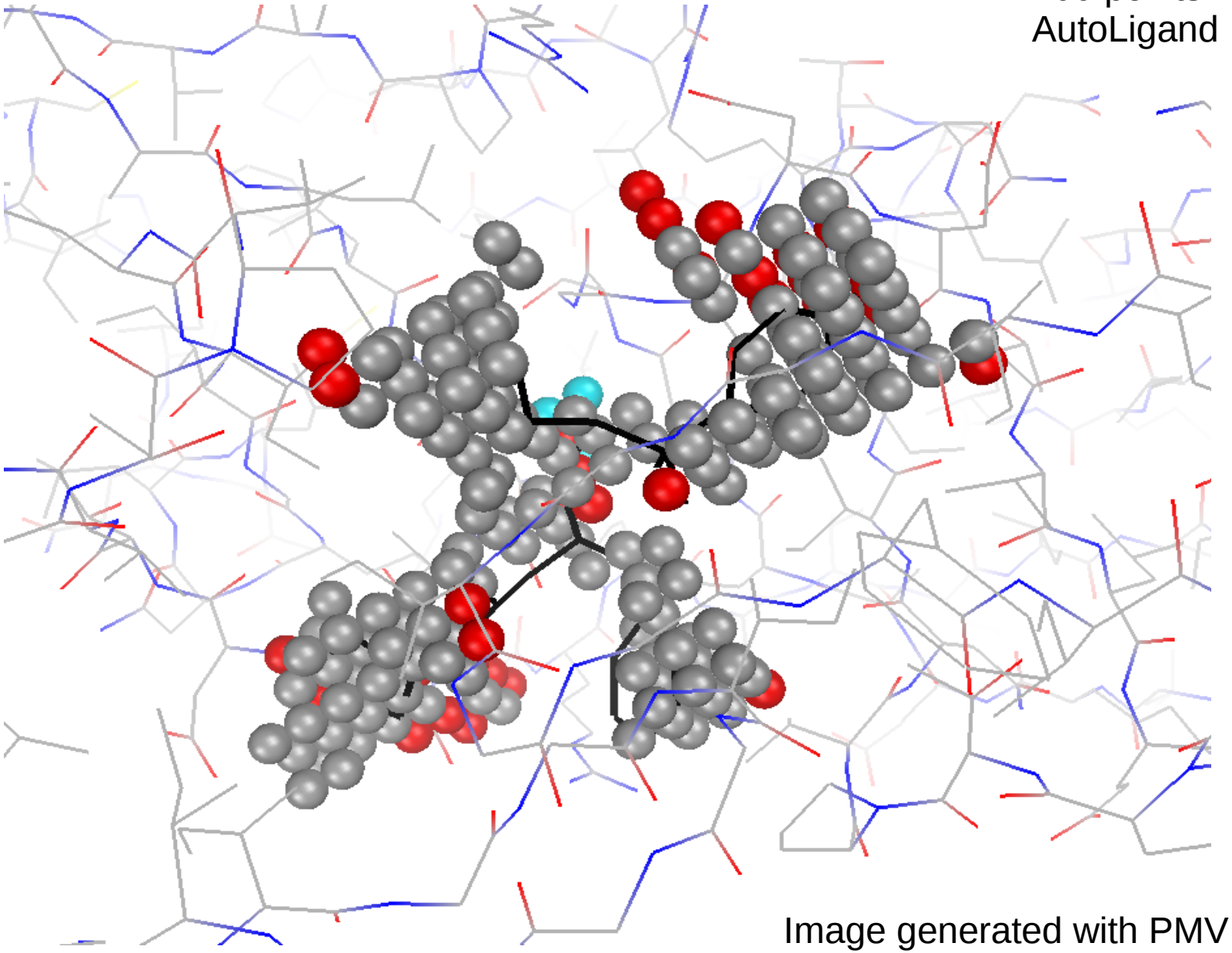


Image generated with PMV 1.4

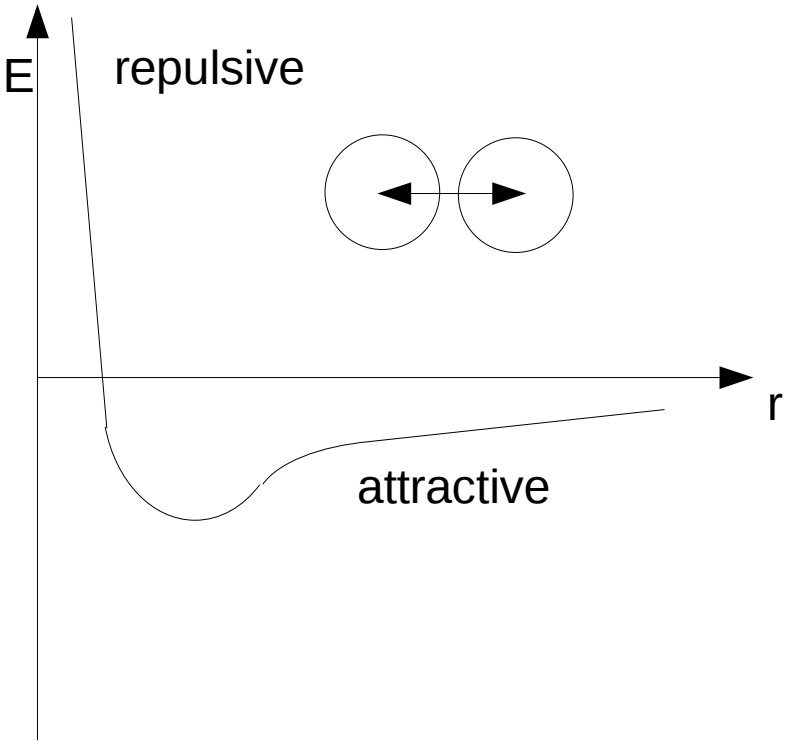
Algorithm

- Generate 1 Å grid using AutoGrid
- Pick starting points
  - scan whole protein for 10 best
  - or pick single point from GUI
- Fill loop:
  - Flood fill (initial volume fill)
  - Affinotaxis migration (optimization)
  - Pseudopod extensions (extended search)
- Output results

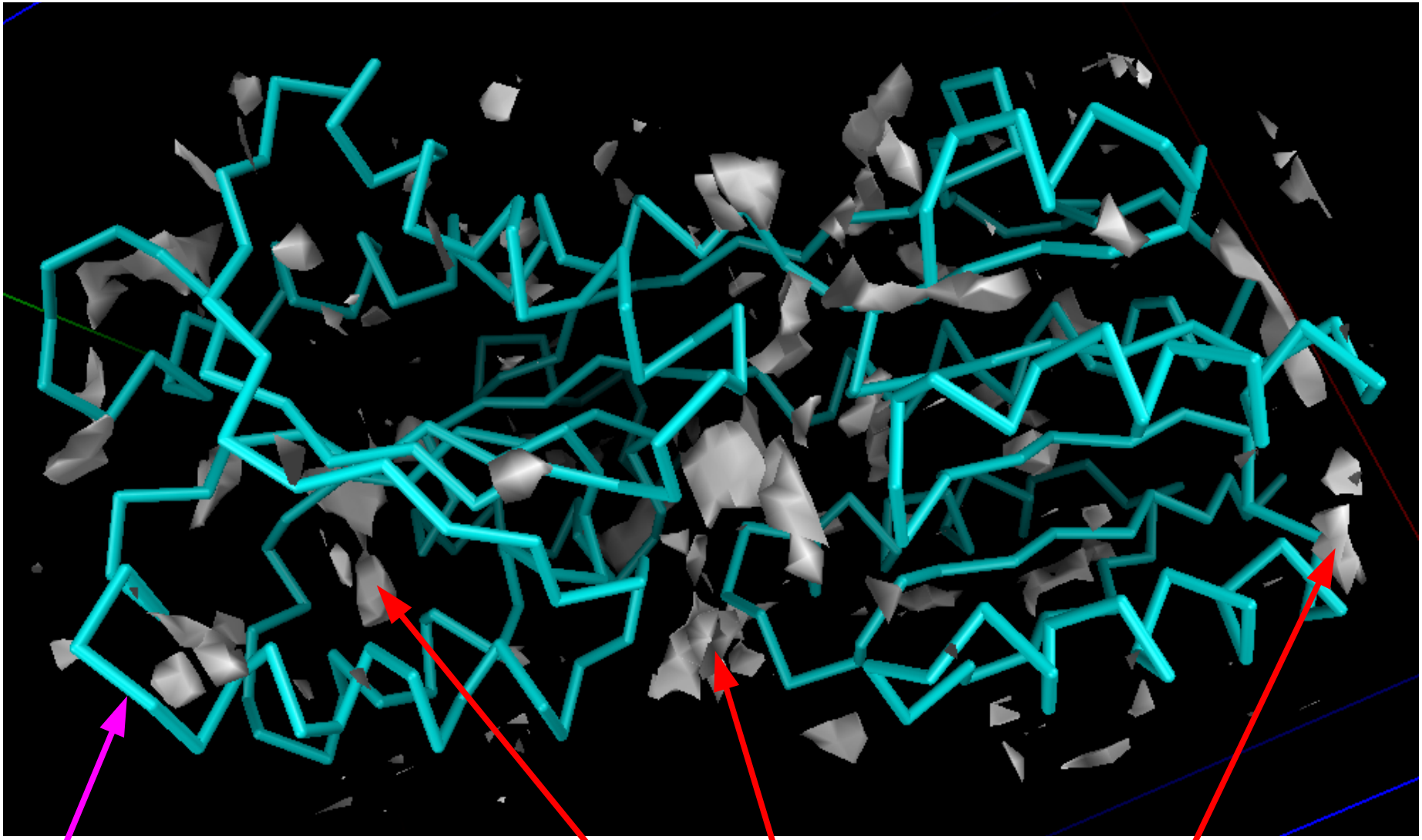
Reference: Harris, R. M., Olson, A. J., and Goodsell, D. S. Automated prediction of ligand-binding sites in proteins. Proteins 2008; 70:1506-1517.

AutoGrid Affinity Calculation

- Empirical force field free energy function with the following terms:
  - Van der Waal
  - Electrostatic
  - Hydrogen bond
  - Desolvation



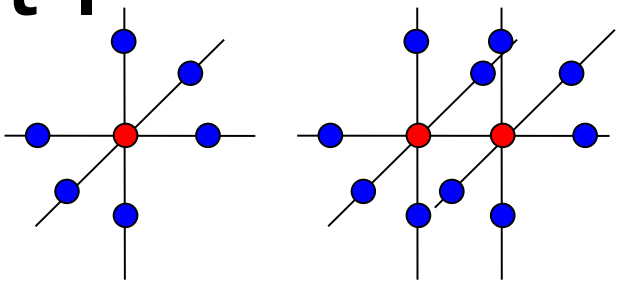
AutoGrid Carbon Affinity Map



HIV Protease backbone

Zones of high carbon atom affinity (gray shapes)

Fill Loop – part I



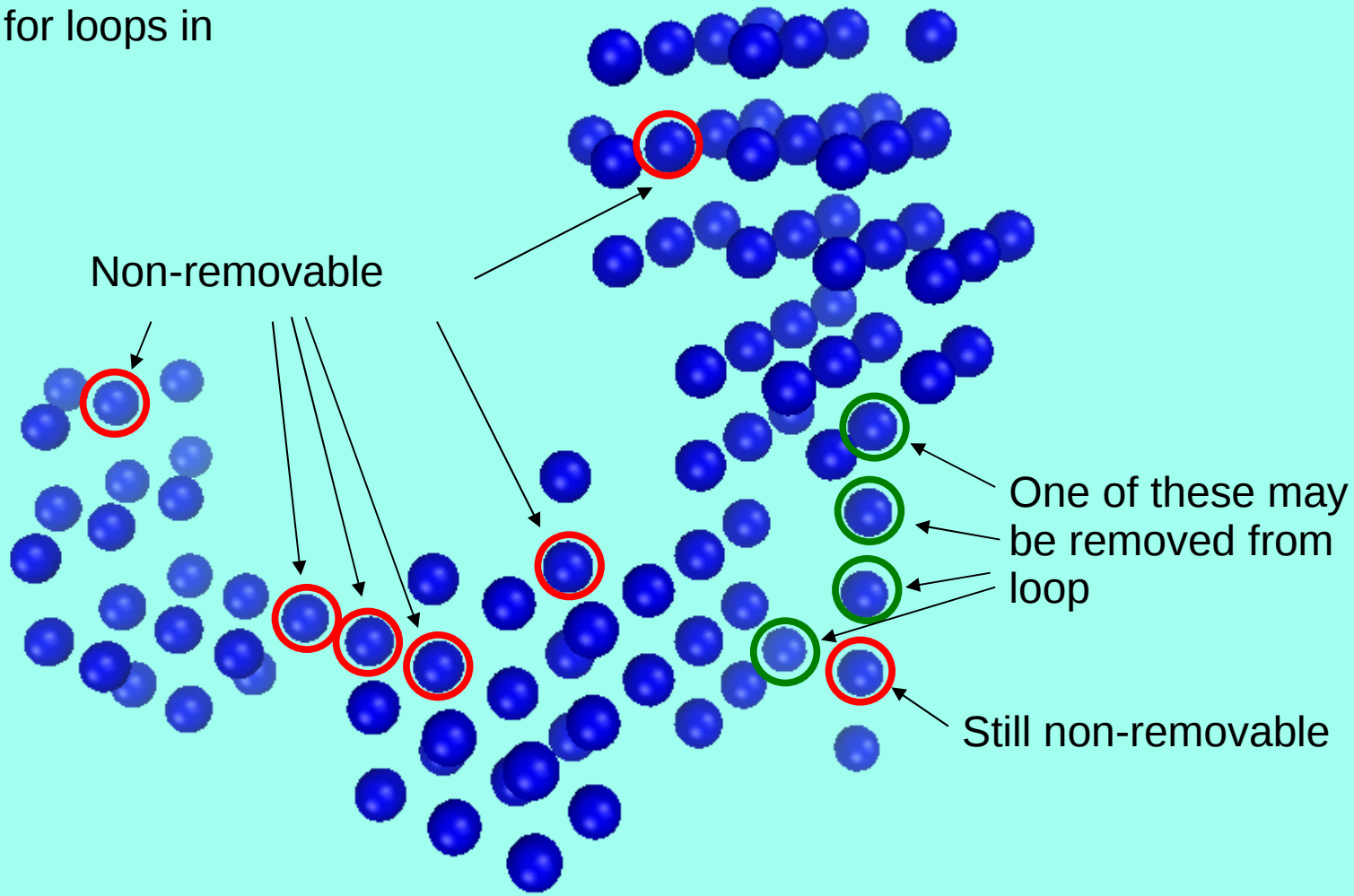
- Fill initial volume (flood fill):
  - Add starting point to Fill list (red point)
  - Generate list of orthogonal Neighbor points (6 blue points)
  - Find best Neighbor point (sort on affinity energy) and add to Fill list (change from blue to red)
  - Remove that point from Neighbor list and generate new Neighbors (now have 10 blue points)
  - This produces a shell of Neighbor points around the Fill list
  - Continue adding points until user specified number is reached

Fill Loop – part II

- Affinotaxis (volume migration to areas of higher affinity)
  - Compare best Neighbor point to worst Fill point, if better – add Neighbor point and remove Fill point
  - But! Need to check if point to be removed is safe to remove – if not, remove 2nd worst Fill point, etc.
  - Assign State to point in two passes
    - 1st pass: assign removable/non-removable state
    - 2nd pass: re-evaluate non-removable state point to see if in loop structure and thus removable

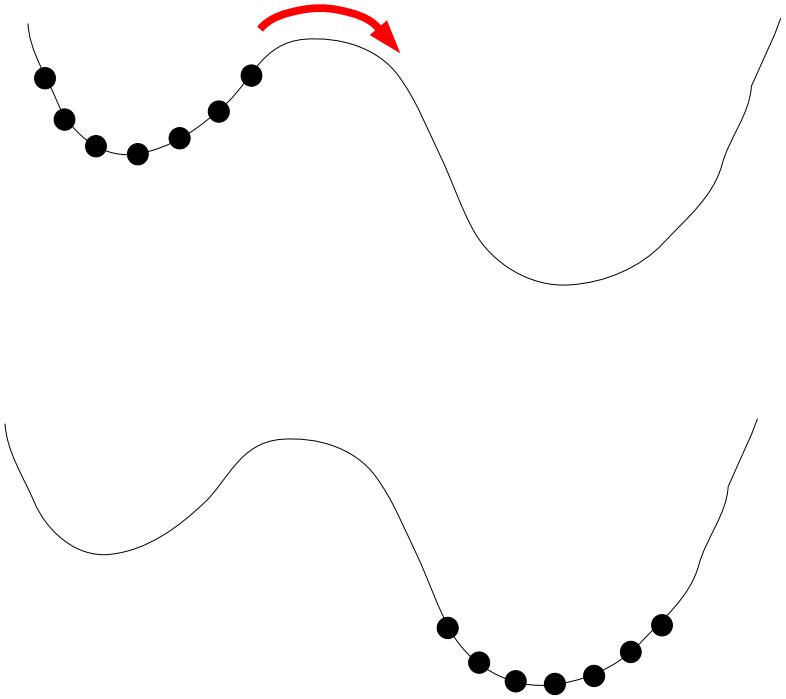
Second pass non-removables

Only 11 points need to be re-evaluated for loops in 2nd pass



Fill Loop – part III

The code must avoid local minima



Fill Loop – part III

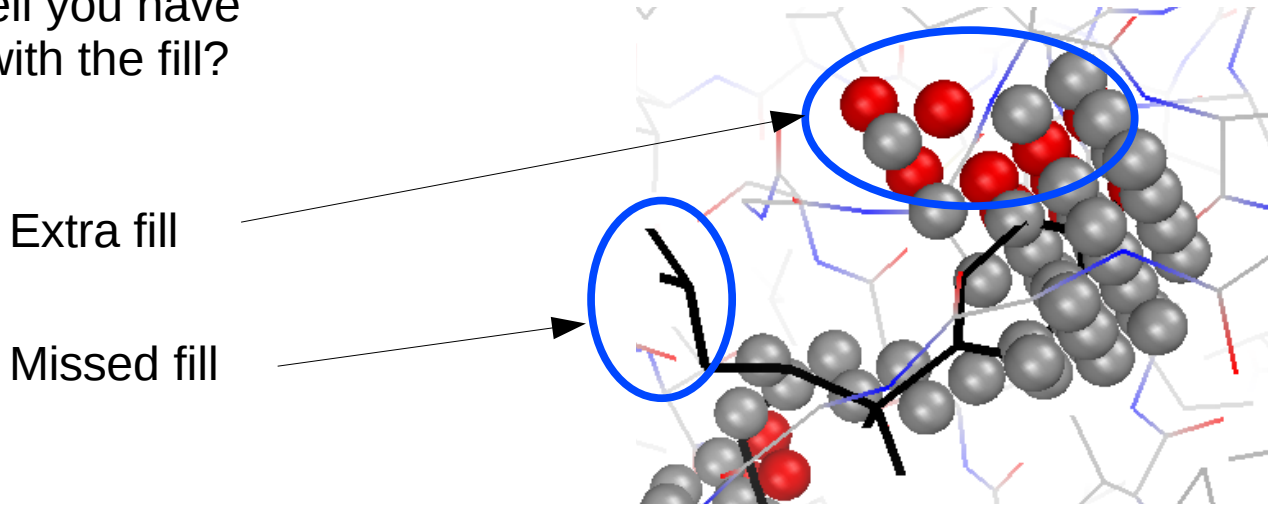
- Pseudopod extension searching (extend search “over the hill” for better affinity)
  - Select best point in 6 Å shell about each point in turn
  - Extend a set of points from Fill point to the selected point and sum total affinity energy
  - Compare total energy of pseudopod to total energy of same number of removable worst points
  - If extended ray is better, add to Fill and remove worst points – then loop back to flow migration and pseudopod projection again

Testing the code

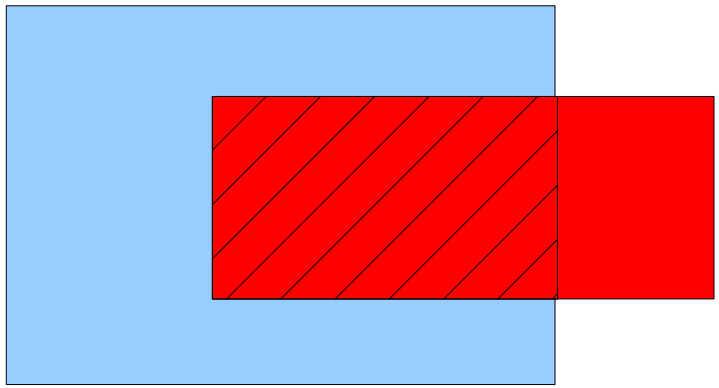
- Screen AutoDock calibration set of 187 receptor/ligand systems – count the % of the ligand overlapped by the volume fill
- Ligand site is found in 95% of the systems
  - 29% of systems overlap 100% of ligand
  - 64% of systems overlap between 20% and 99%
  - 2% of systems overlap < 20% of ligand
  - 5% of systems have no overlap of ligand (a miss)
- Sounds Great?

Better Scoring of the Tests

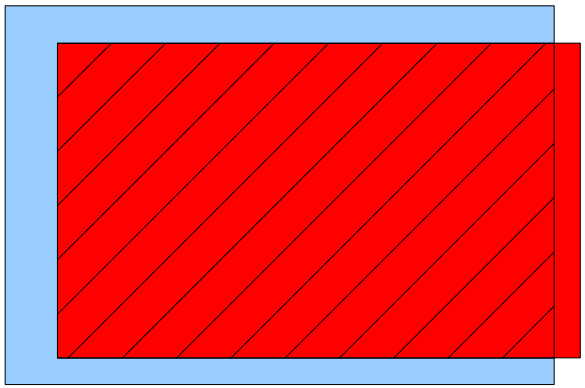
How do you tell how well you have overlapped the ligand with the fill?



Answer: Use an Intersection over Union



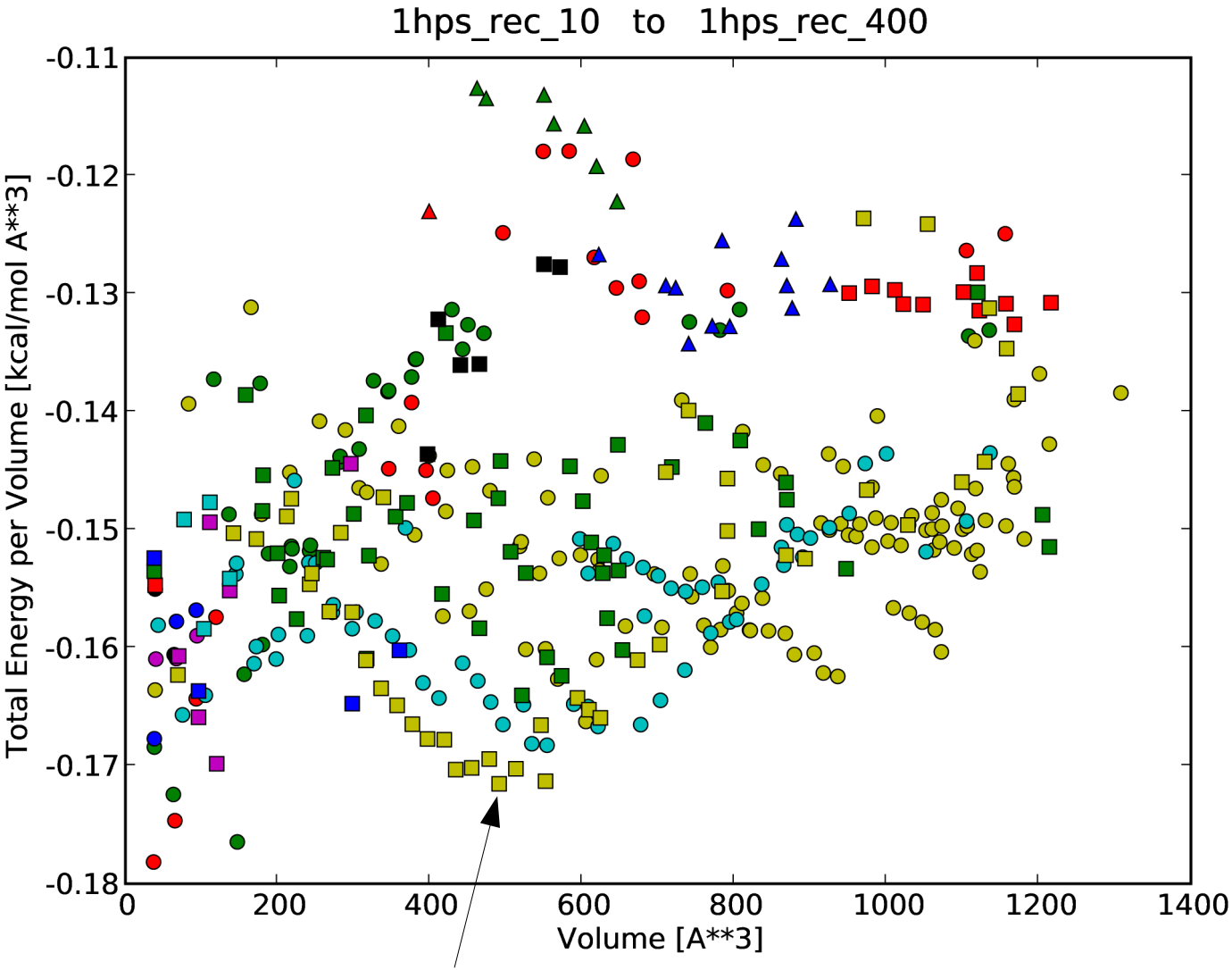
Low I/U



High I/U

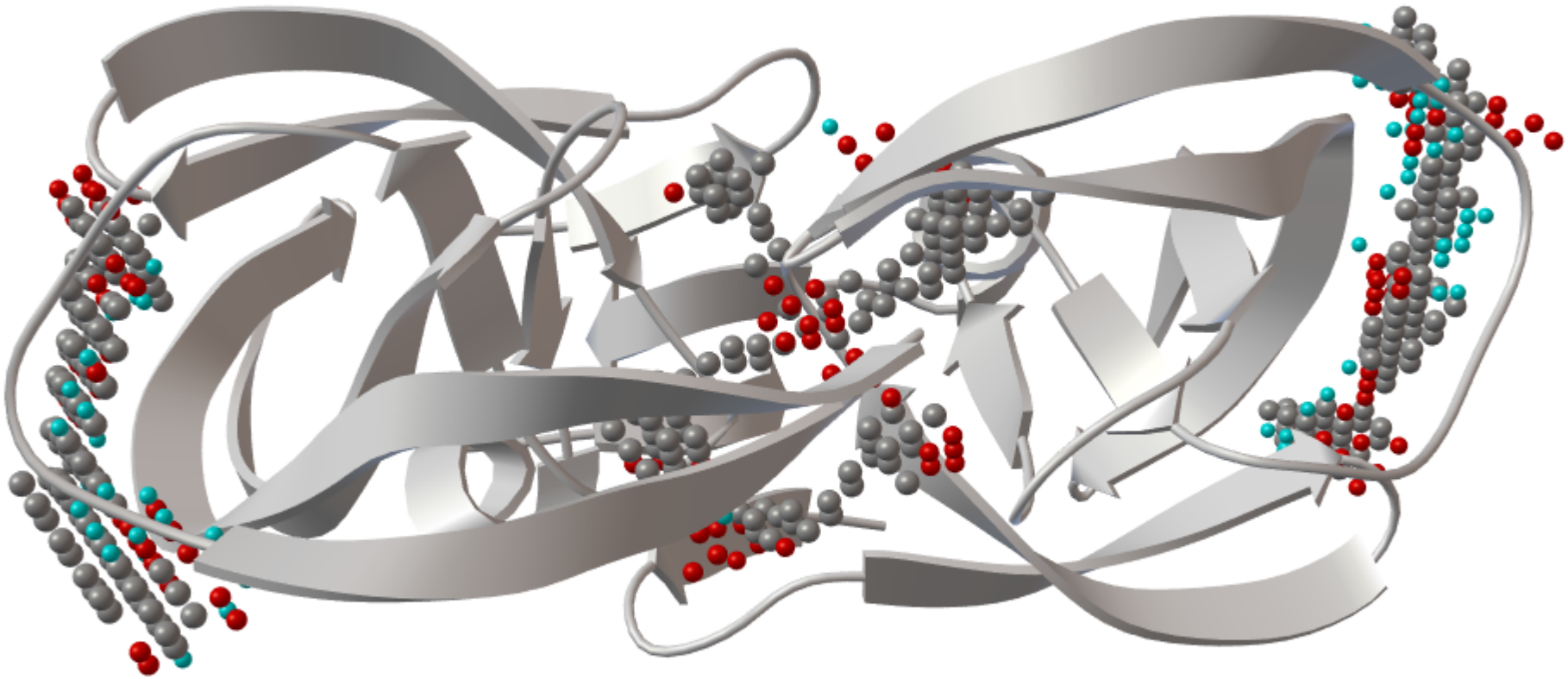
AutoLigand finds binding sites with I/U >= 0.2 only 85% of the time.

Results for HIV (complex)



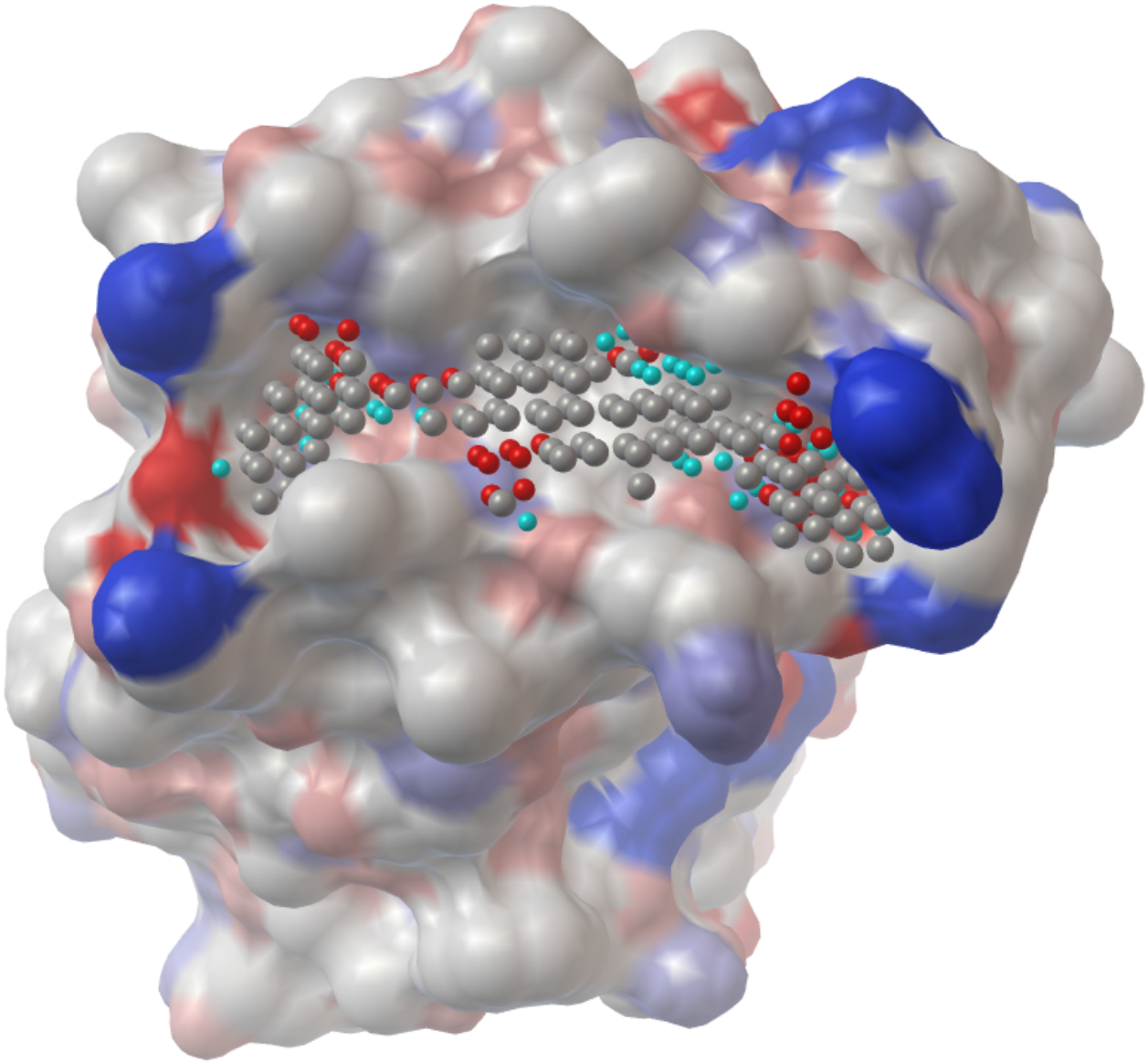
Optimal ligand binding site and size

HIV Results

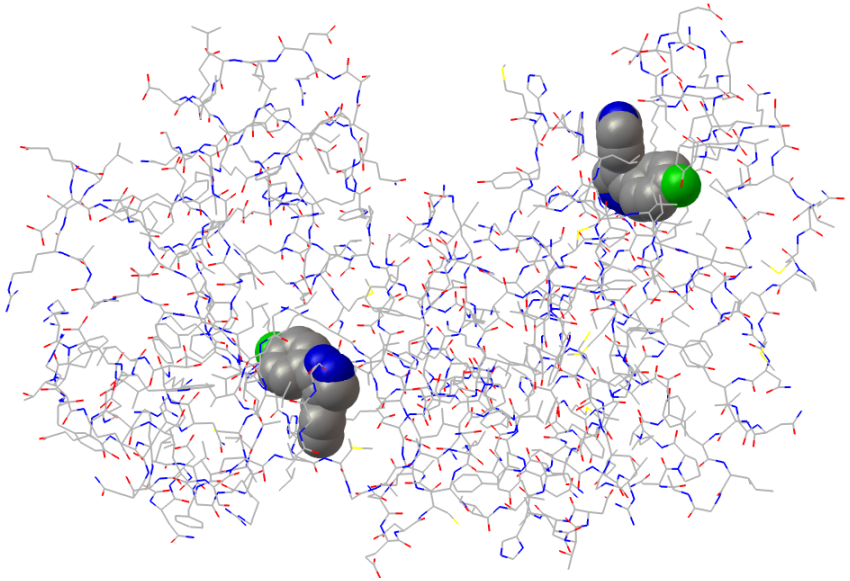


1hps – with active site and exosites filled

HIV exo site

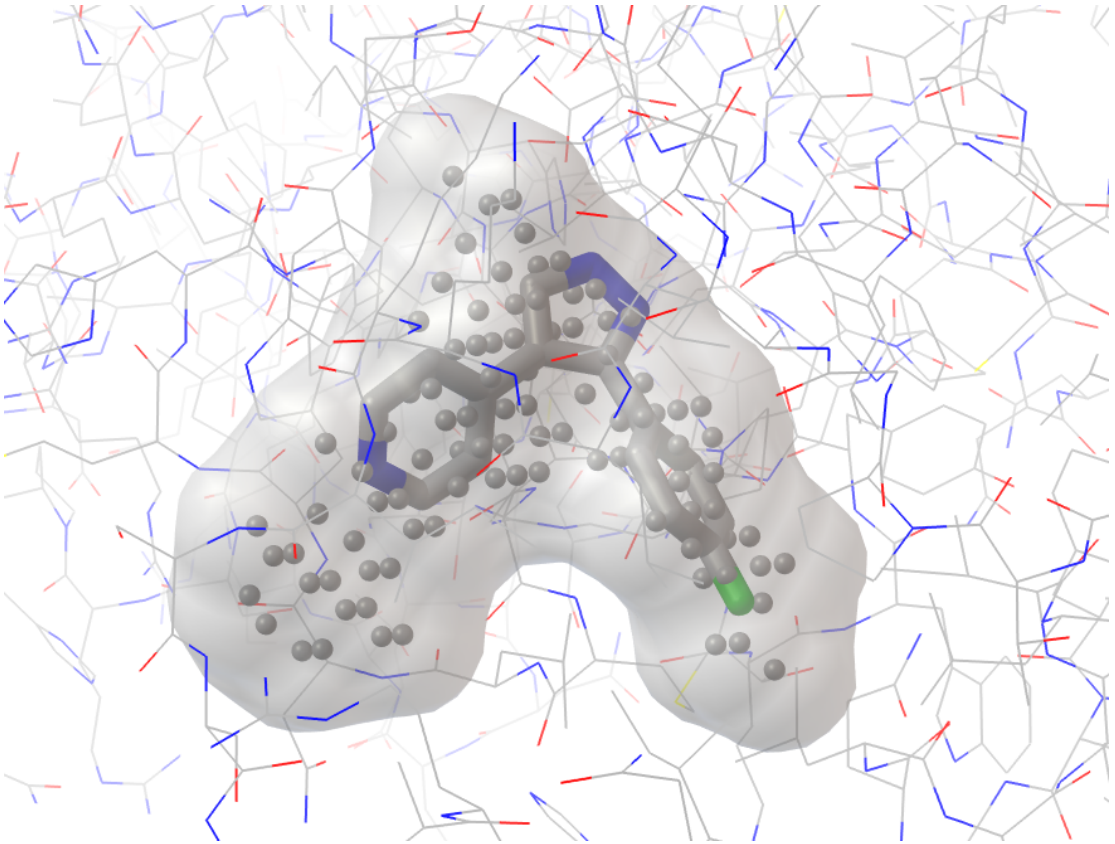


More Results

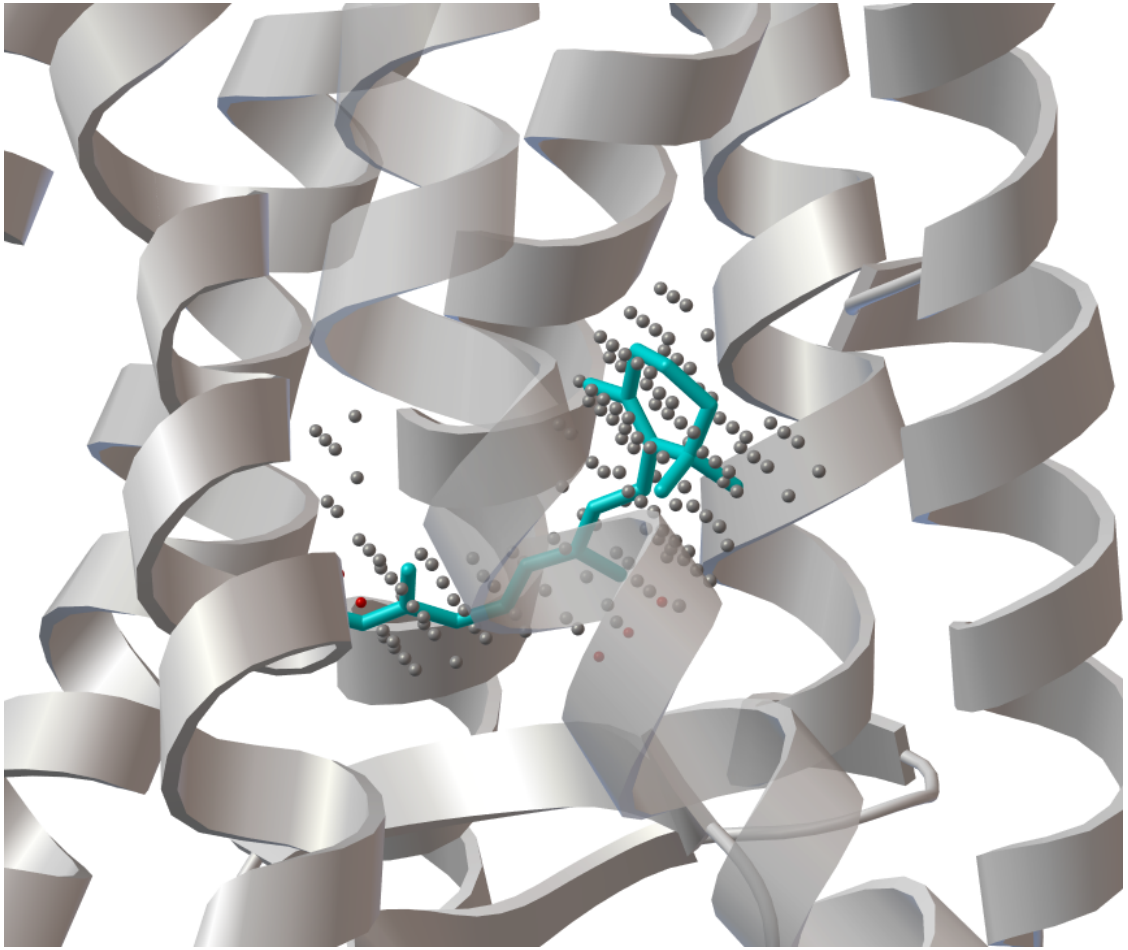


p38a High Throughput Screen found second site

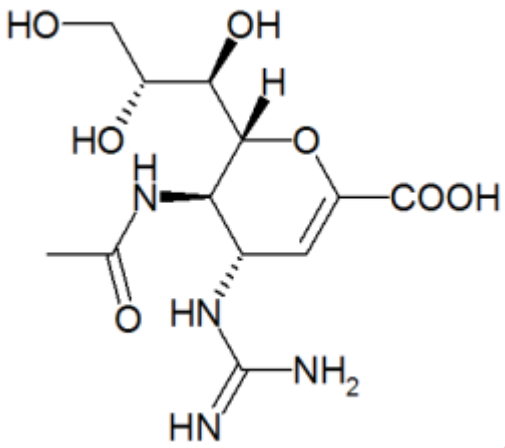
AutoLigand result
for second site



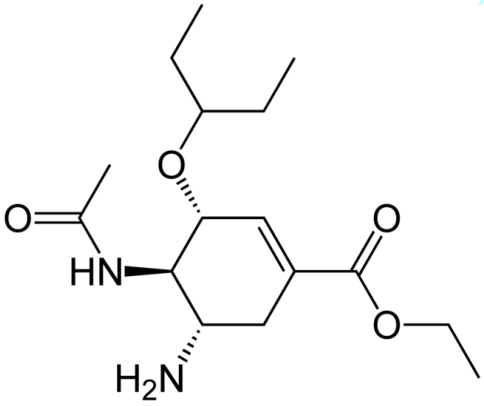
Rhodopsin Binding Site



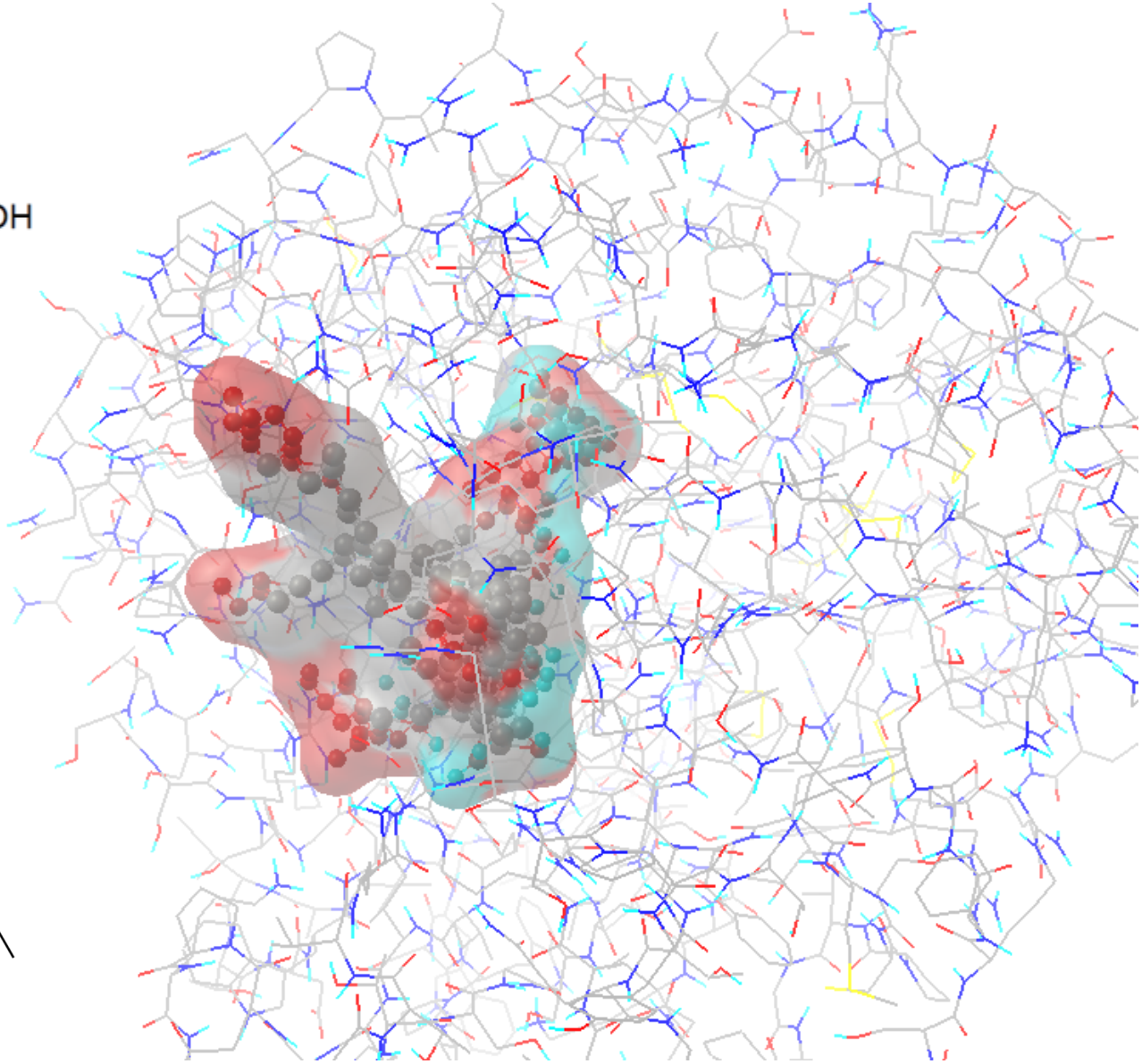
Swine Flu – N1 Binding Site



Relenza



Tamiflu



Conclusions

- AutoLigand is effective in finding ligand binding sites without a priori knowledge of ligand
- AutoLigand can find the size, shape, and atom type best fit for ligand sites
- AutoLigand can be used as a drug development tool by indicating regions to modify existing drugs or suggesting new drug structures altogether

Acknowledgments

- Art Olson
- David Goodsell
- Albert Beuscher
- Mike Pique
- Michel Sanner