

10/02/2024

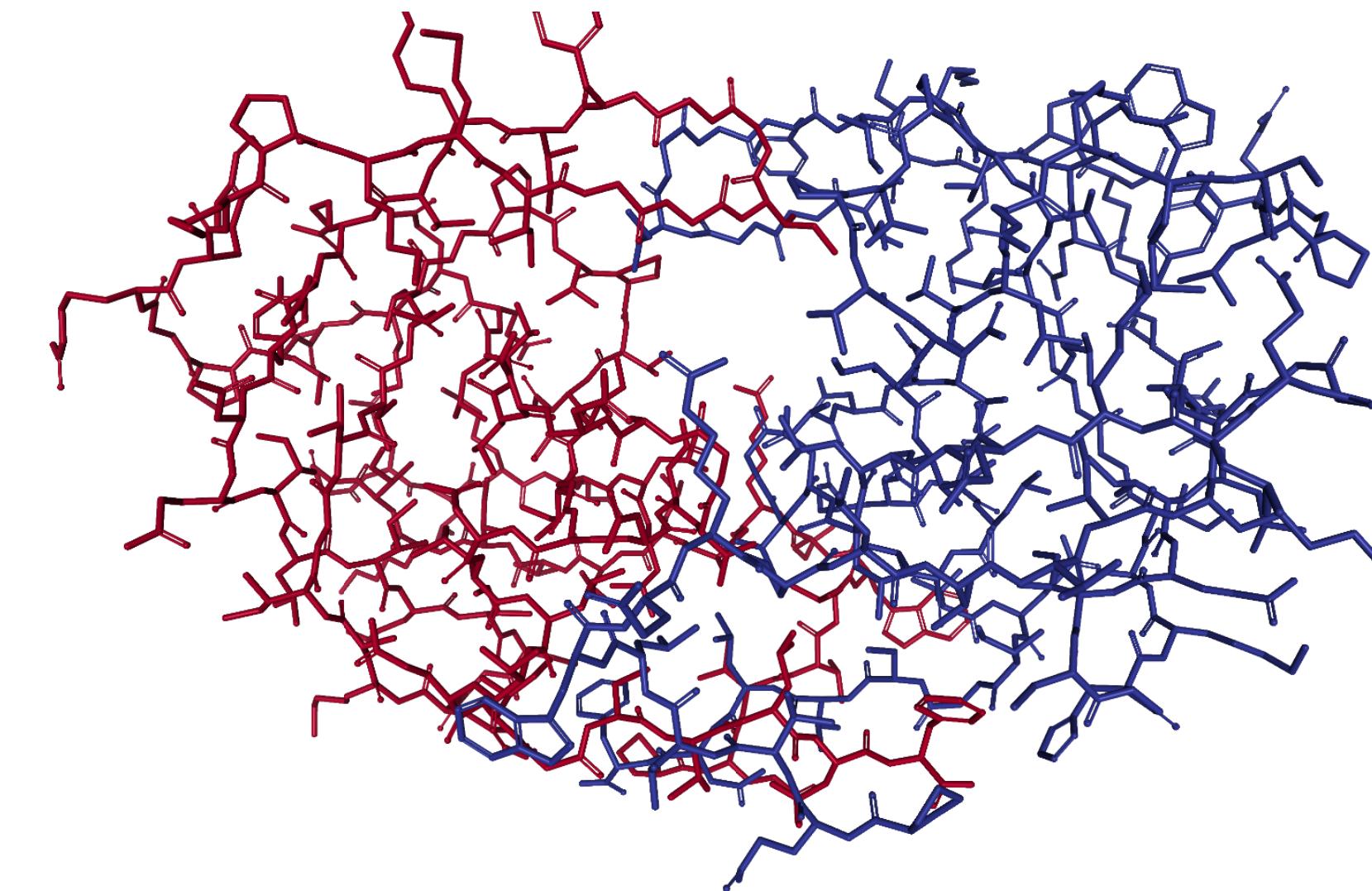
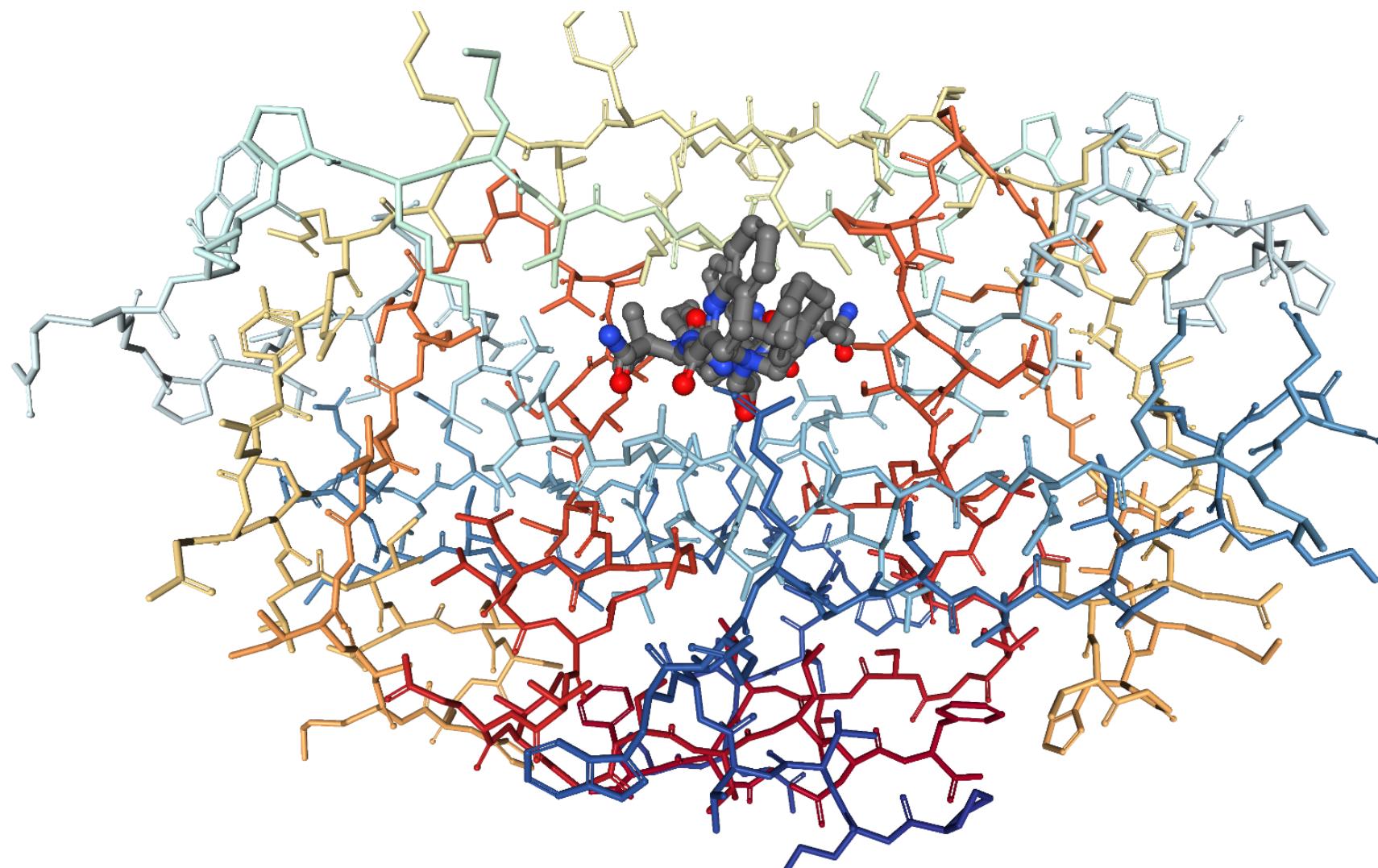
- Today's lecture is intended to help you achieve the following learning objective: Perform a molecular docking calculation and visualize the results. Explain the limitations of molecular docking.
- Introduction to Molecular Docking
- AutoDock Vina

Introduction to Molecular Docking

- At the end of this lecture, you should be able to address these questions:
 - What is molecular docking?
 - What is it good for?
 - How does it work?
 - What are its key approximations?

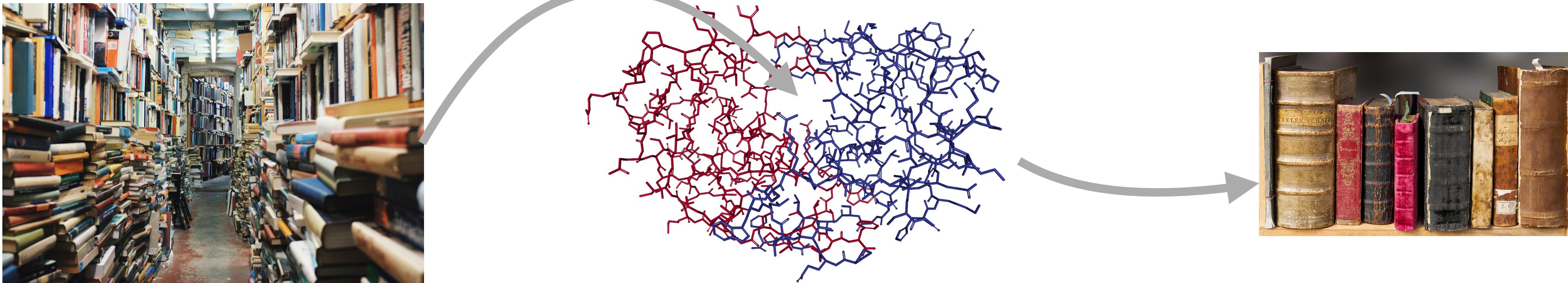
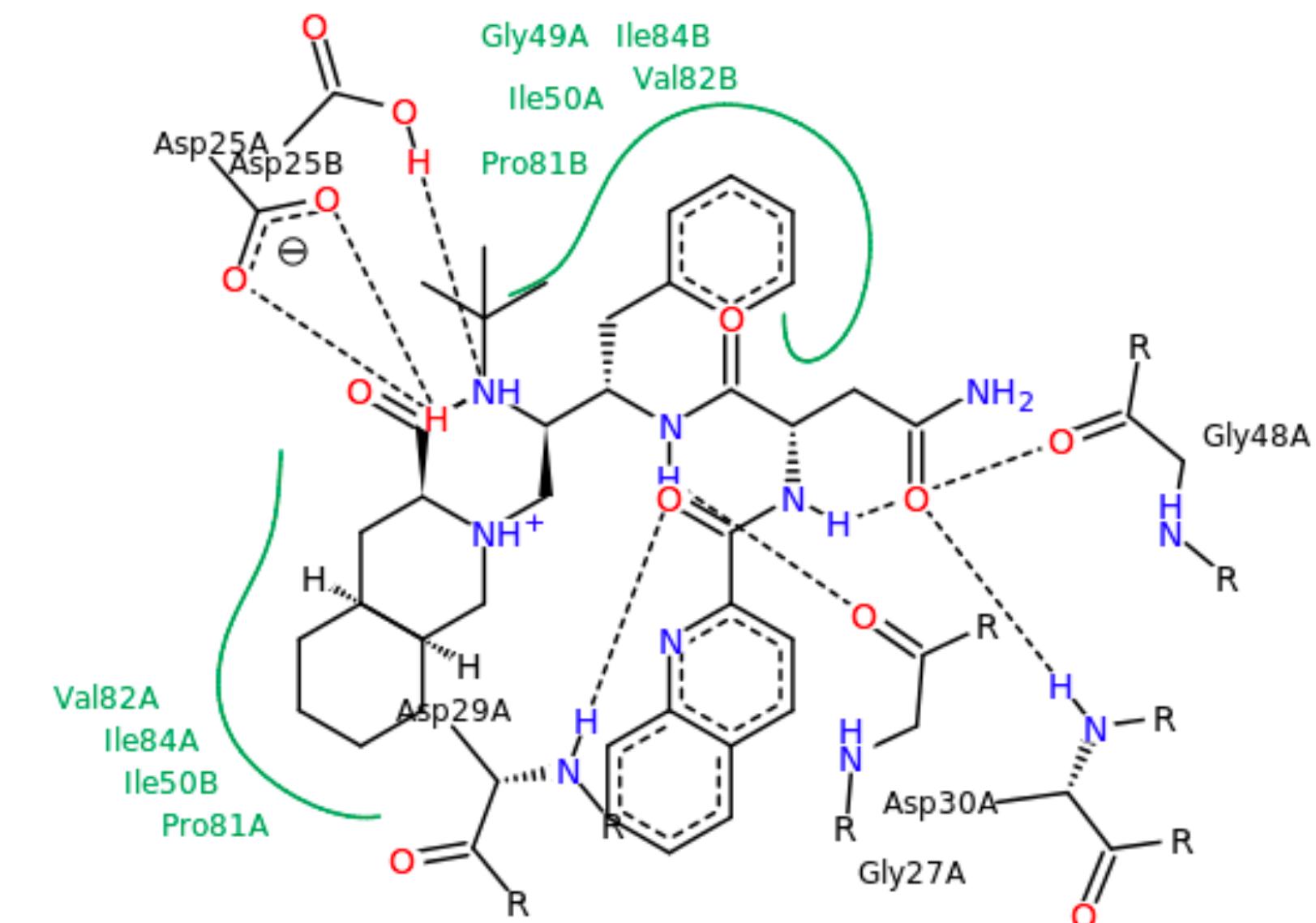
What is molecular docking?

- To predict
 - the 3D structure of a noncovalent complex
 - protein-ligand
 - protein-protein
 - the binding affinity of the partners (scoring)
- Prior to prediction, structure of binding partner(s) may be known, but
 - could be affected by binding
 - may be bound to different partners



What is molecular docking good for?

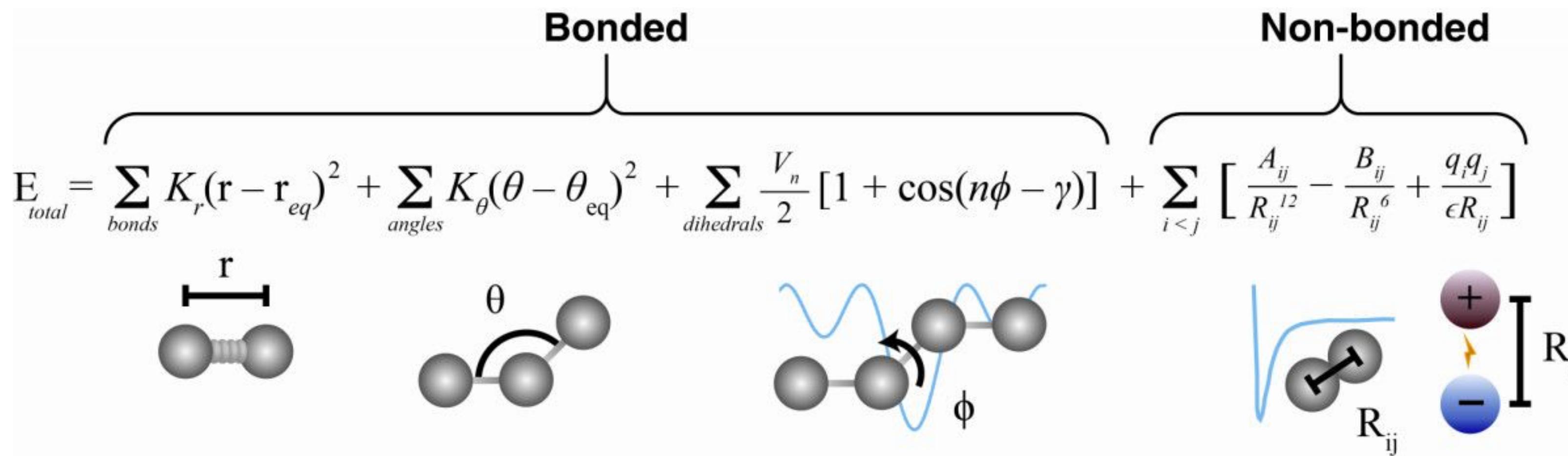
- Predicted structure can
 - explain behavior, e.g. mutagenesis
 - facilitate molecular design, e.g. of pharmaceuticals
 - Predicted scores can be used for virtual screening



How does docking work?

- Docking is optimization of a scoring function, $E(x)$
 - E can be the total potential energy or interaction energy
 - can be entirely physics-based or partly knowledge-based
 - physics-based are usually molecular mechanics energies
 - x is a vector describing the molecular coordinates
- Optimization algorithms include
 - anchor-and-grow in UCSF DOCK, the original docking program
 - genetic algorithm in AutoDock, the most popular docking program
 - Fast fourier transform, especially for fragment and protein-protein docking

Scoring is based on molecular mechanics



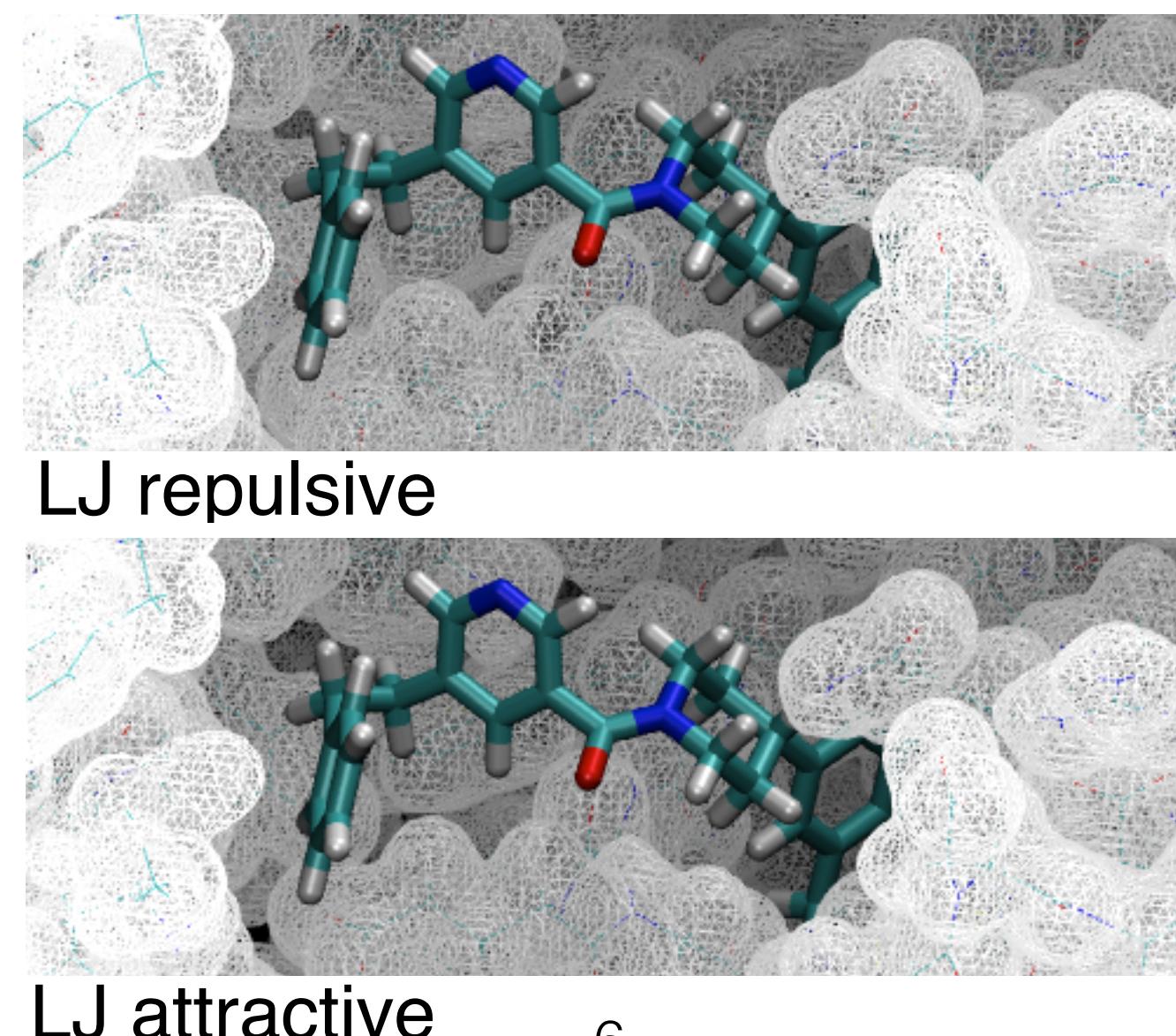
From Figure 3 of
Durant and McCammon, 2011

AMBER interaction energies

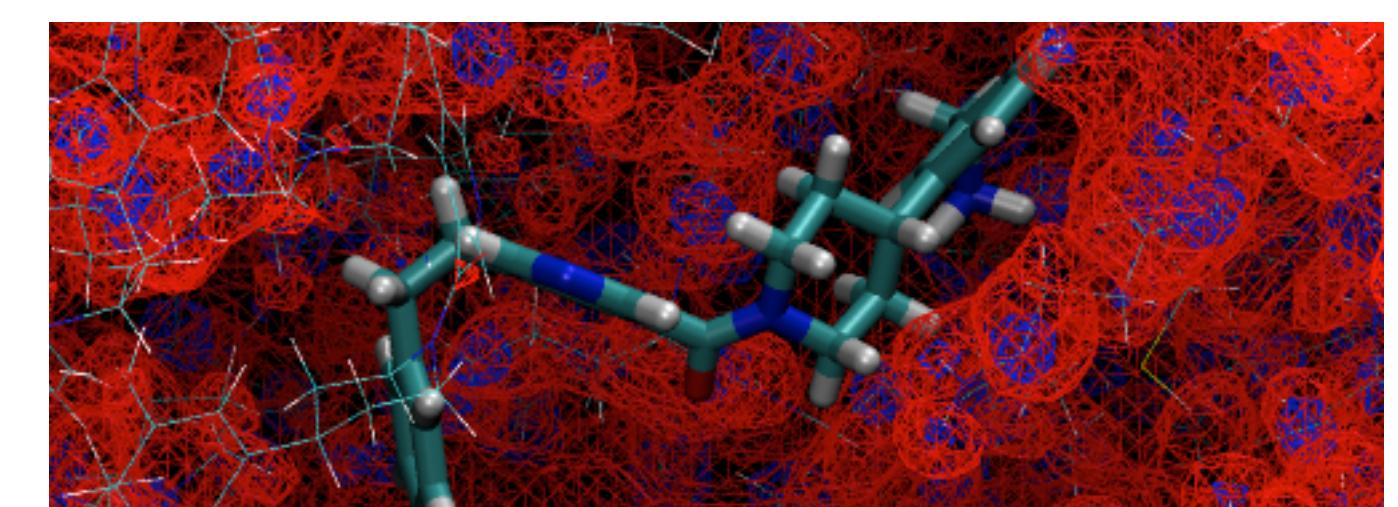
$$E = \sum_{i=1}^{lig} \sum_{j=1}^{rec} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + 332.0 \frac{q_i q_j}{Dr_{ij}} \right],$$

$$A_{ij} = \sqrt{A_{ii}} \sqrt{A_{jj}} \quad \text{and} \quad B_{ij} = \sqrt{B_{ii}} \sqrt{B_{jj}},$$

[Meng, Shoichet, and Kuntz, 1992]



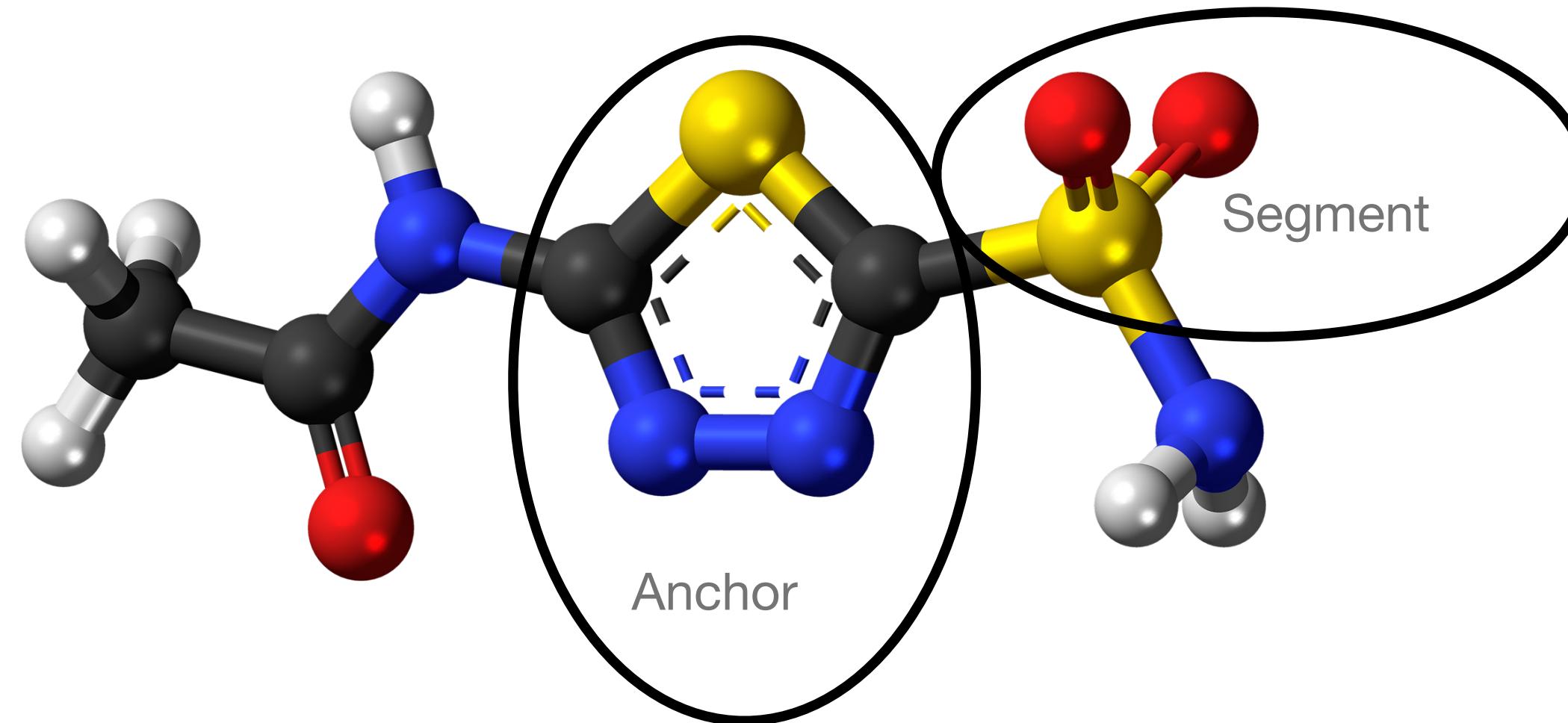
Nonbonded interactions are often interpolated from precomputed grids



$E(x)$ is often rugged, requiring special optimization

This principle can be understood considering climbing a mountain peak. For minimization, flip everything upside down.

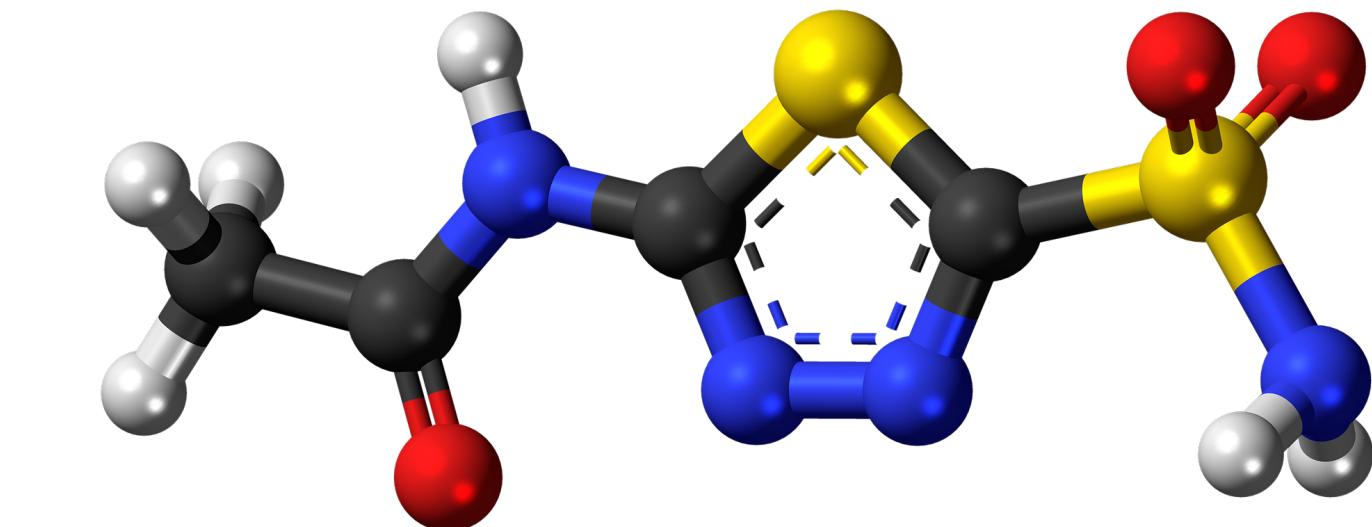
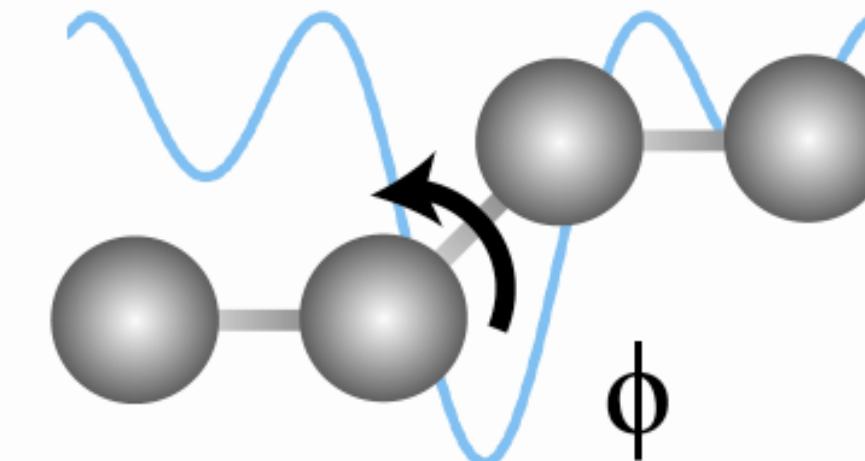
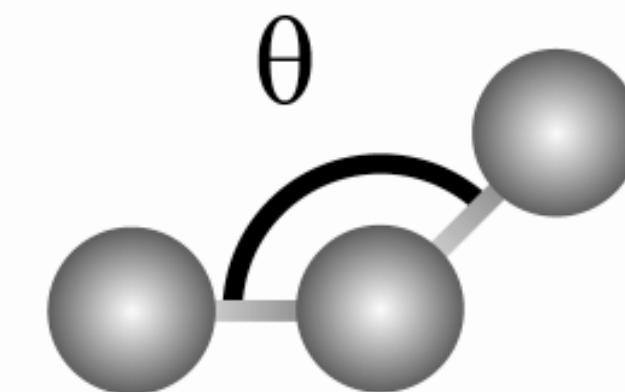
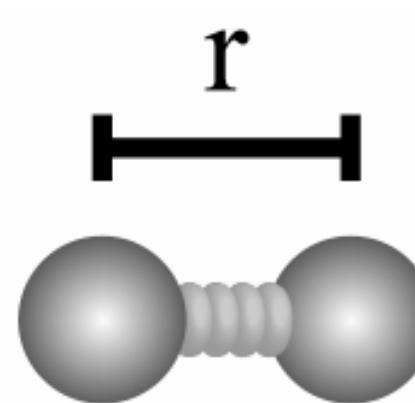
UCSF DOCK is based on anchor-and-grow



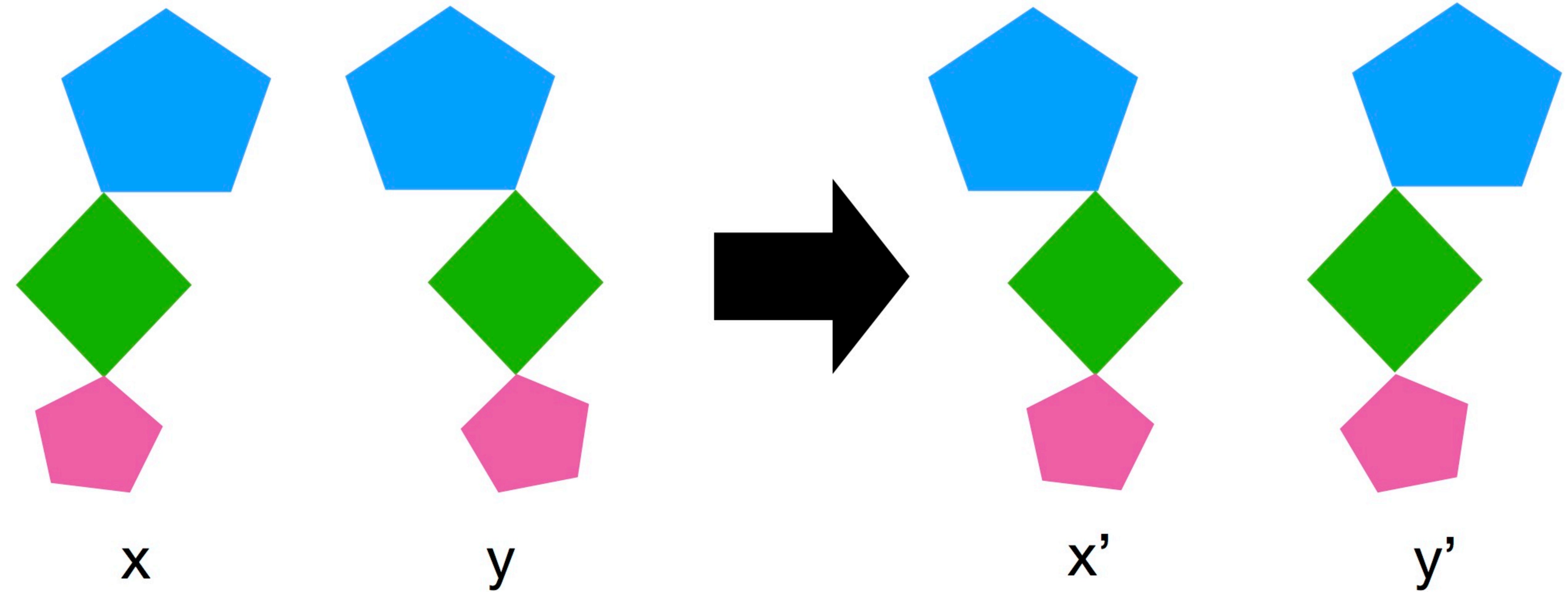
- Receptor spheres and rigid “anchor” in ligand
 - represented as graph of atoms separated by distances
 - docking is search for isomorphic subgraph
- Until the molecule is complete, segments are iteratively
 - added to the anchor and
 - pruned if the energy is too high
- Complete structures are locally minimized

AutoDock uses a genetic algorithm

- Population of structures
 - represented by torsions. bond length and angles assumed constant.
 - evolve over generations
- Generations iterate
 - mapping & fitness evaluation. mapping x and calculating $E(x)$.
 - selection. fitter individuals reproduce more.
 - crossover. torsions swapped between individuals. enable global search.
 - mutation. small changes to individuals. permit local search.

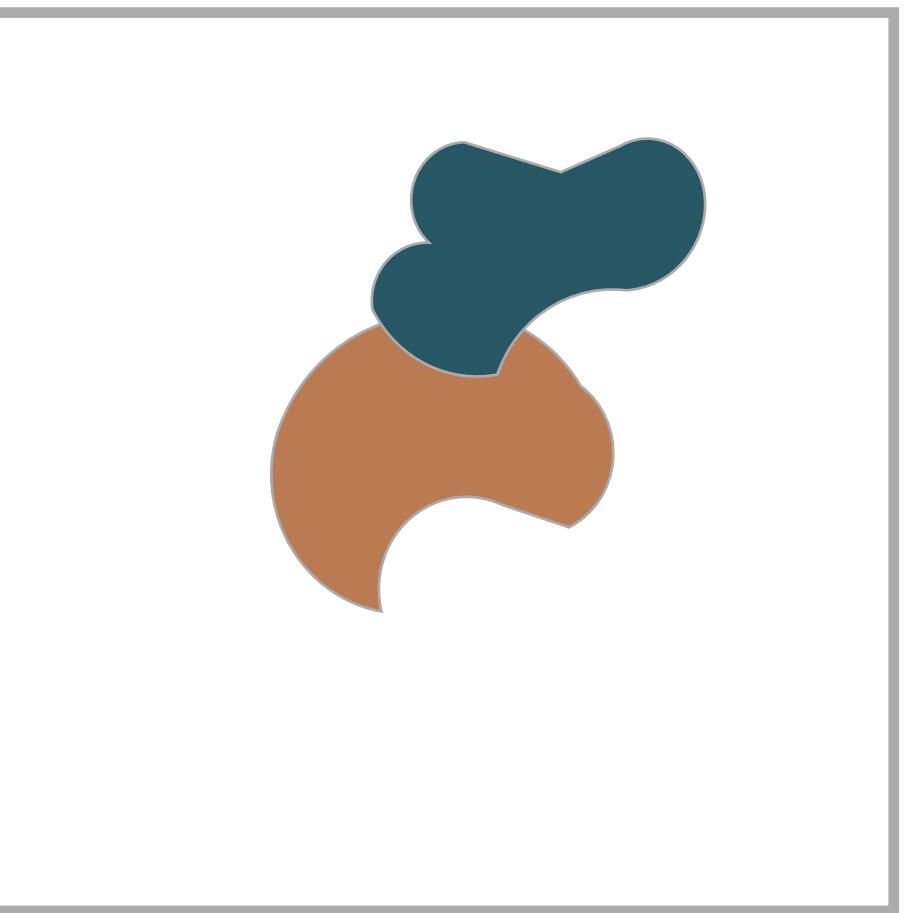


Crossover

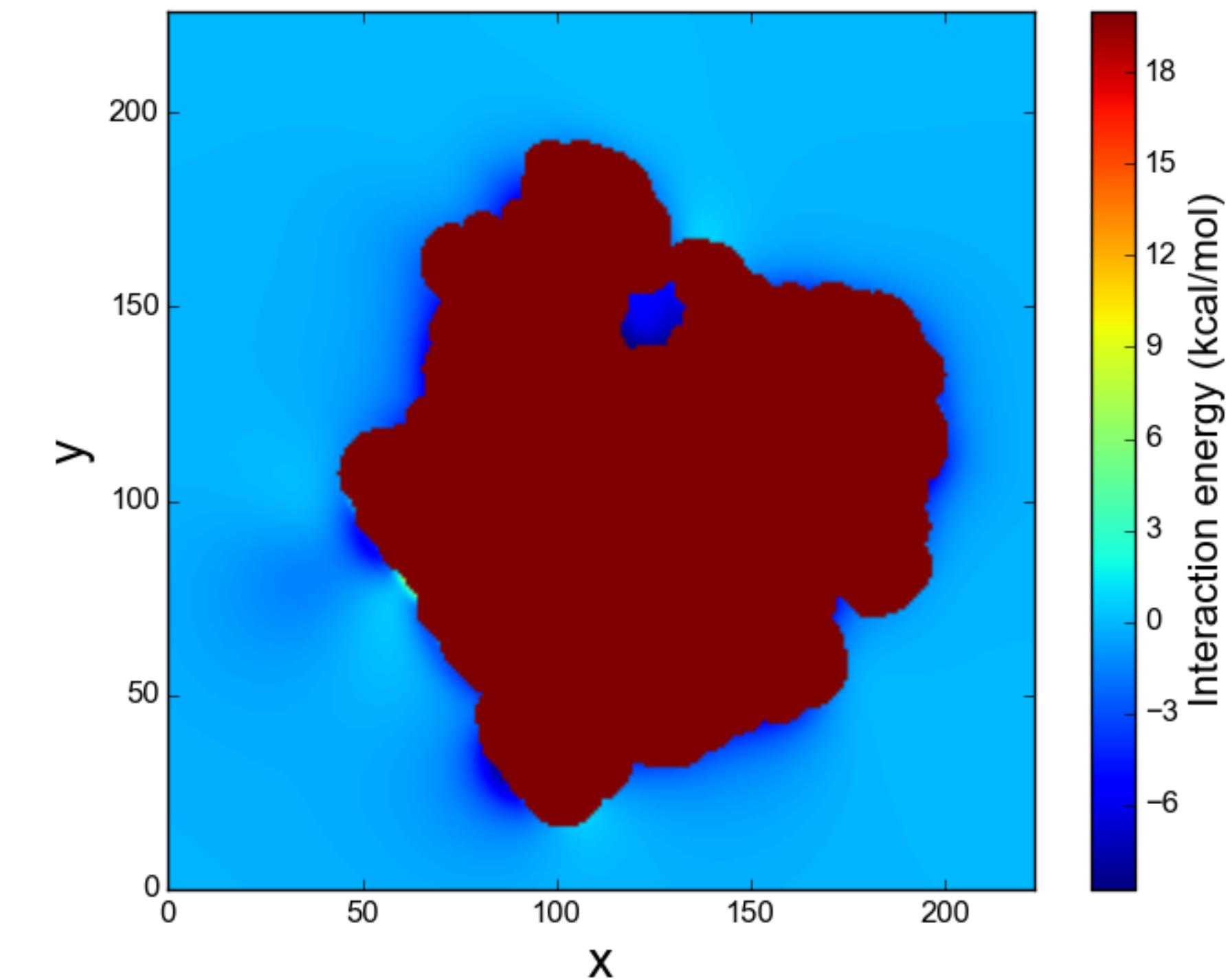
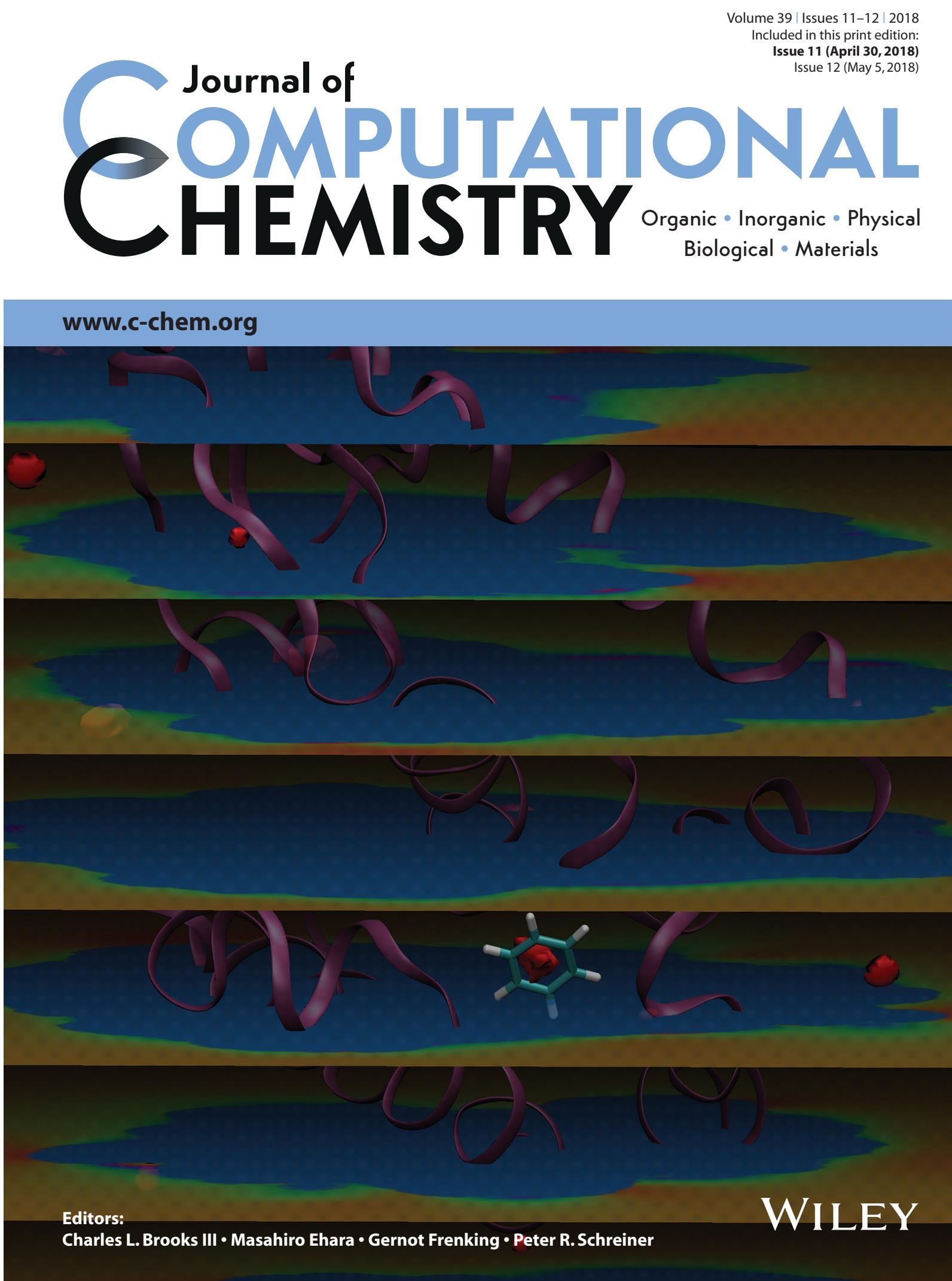


FFT is common in docking

- Both molecules are represented by a 3D grid
- FFT correlation calculation
 - gives scores as molecules are translated relative to one another
 - is much faster than direct calculation
- Scores can be based on
 - shape complementarity [Katchalski-Katir et al, 1991]
 - van der Waals and electrostatics
- Strategy makes most sense for rigid binding



FFT can also be used to estimate binding ΔG



2D cross section of the interaction energy

[Nguyen, Zhou, and Minh, 2018]

How well does docking work?

- How well does it work?
 - Fairly successful at generating binding poses [Damm-Ganamet et al, 2013]. Usually successful (~80%) at ranking them.
 - Poorly correlated with binding free energies [Warren et al, 2006].
 - Unreliable at separating actives from decoys [Cross et al, 2009].
 - Sometimes successful at virtual screening. Hit rates < 20%.

How can it get better?

- How can it be better?
 - Free energy includes both enthalpy and entropy. Docking scores usually exclude entropy.
 - Water often mediates protein-ligand interactions. Docking usually does not consider it.
 - Polarizability. Ligands can adapt to the protein environment. Docking usually does not consider this.

Review

- What is molecular docking?
- What is it good for?
- How does it work?
- What are its key approximations?

References

- Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humblet, C. Comparison of Several Molecular Docking Programs: Pose Prediction and Virtual Screening Accuracy. *Journal of Chemical Information and Modeling* 2009, 49 (6), 1455–1474. <https://doi.org/10.1021/ci900056c>.
- Damm-Ganamet, K. L.; Smith, R. D.; Dunbar, J. B.; Stuckey, J. A.; Carlson, H. A. CSAR Benchmark Exercise 2011-2012: Evaluation of Results from Docking and Relative Ranking of Blinded Congeneric Series. *Journal of Chemical Information and Modeling* 2013, 53 (8), 1853–1870. <https://doi.org/10.1021/ci400025f>.
- Durrant, J. D.; McCammon, J. A. Molecular Dynamics Simulations and Drug Discovery. *BMC Biol* 2011, 9 (1), 71. <https://doi.org/10.1186/1741-7007-9-71>, adapted under the CC BY 2.0 license.
- Katchalski-Katzir, E.; Shariv, I.; Eisenstein, M.; Friesem, A. a; Aflalo, C.; Vakser, I. a. Molecular Surface Recognition: Determination of Geometric Fit between Proteins and Their Ligands by Correlation Techniques. *Proceedings of the National Academy of Sciences of the United States of America* 1992, 89 (6), 2195–2199. <https://doi.org/10.1073/pnas.89.6.2195>.
- Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. Automated Docking with Grid-Based Energy Evaluation. *Journal of Computational Chemistry* 1992, 13 (4), 505–524.
- Nguyen, T. H.; Zhou, H.-X.; Minh, D. D. L. Using the Fast Fourier Transform in Binding Free Energy Calculations. *Journal of Computational Chemistry* 2018, 39, 621–636. <https://doi.org/10.1002/jcc.25139>.
- Warren, G. L.; Andrews, C. V. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F; Senger, S.; et al. A Critical Assessment of Docking Programs and Scoring Functions. *Journal of Medicinal Chemistry* 2006, 49 (20), 5912–5931. <https://doi.org/10.1021/jm050362n>.

AutoDock Vina

- This module will be a tutorial on AutoDock Vina
 - docking program developed at The Scripps Research Institute
 - <http://vina.scripps.edu>
 - After this module, you should be able to set up and run a molecular docking calculation using AutoDock Tools and AutoDock Vina

Why AutoDock Vina?

- There are many molecular docking programs
- Why AutoDock Vina?
 - Free
 - Works on multiple platforms
 - Fast
 - Very popular
 - >10,000 citations of primary reference [Trott and Olson, 2010]
 - >1,400 citations to primary reference of AutoDock 3 & 4 [Morris et al, 1998]
 - Available on PSC bridges and other NSF supercomputers

Comparison of AutoDock Versions

AutoDock 4

Lennard-Jones sterics

Lamarkian genetic algorithm

Pre-calculated grid maps

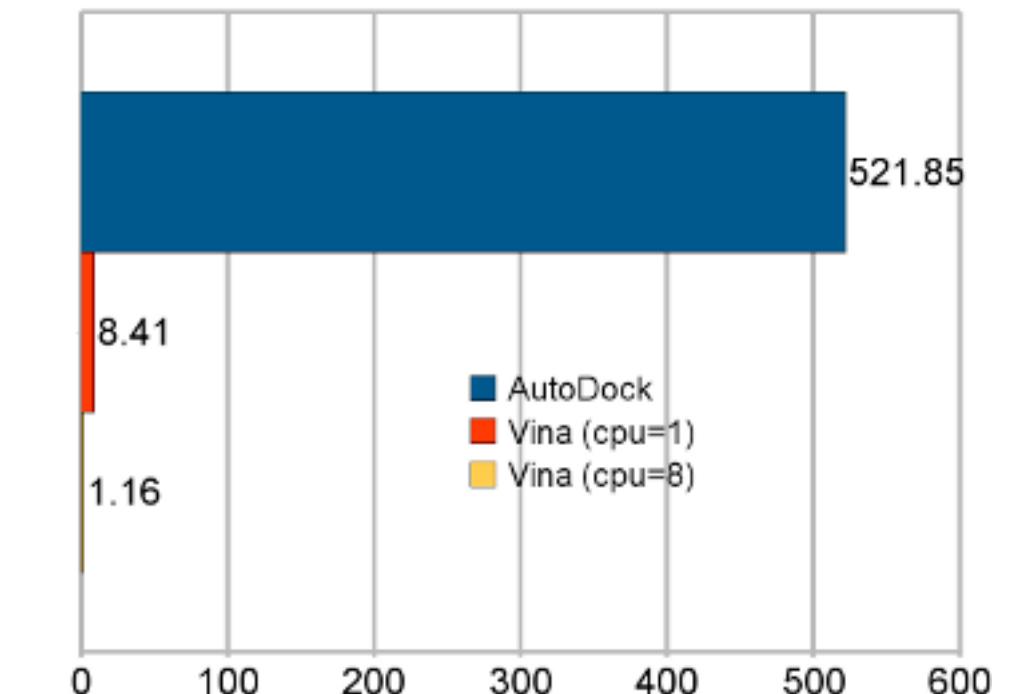
AutoDock Vina

2 orders of magnitude faster.
Parallelized code.

Gaussian sterics

Gradient-based optimizer

On-the-fly grid maps



Average time per receptor-ligand pair on the test set.
“AutoDock” refers to AutoDock 4, and “Vina” to
AutoDock Vina 1.

The Vina Scoring Function

$$c = \sum_{i < j} f_{t_i t_j}(r_{ij}),$$

- General form of terms is , a summation over all pairs of atoms that can move relative to each other
- Steric terms are the sum of two Gaussians and a repulsion,

$$\text{gauss}_1(d) = e^{-(d/0.5\text{\AA})^2}$$

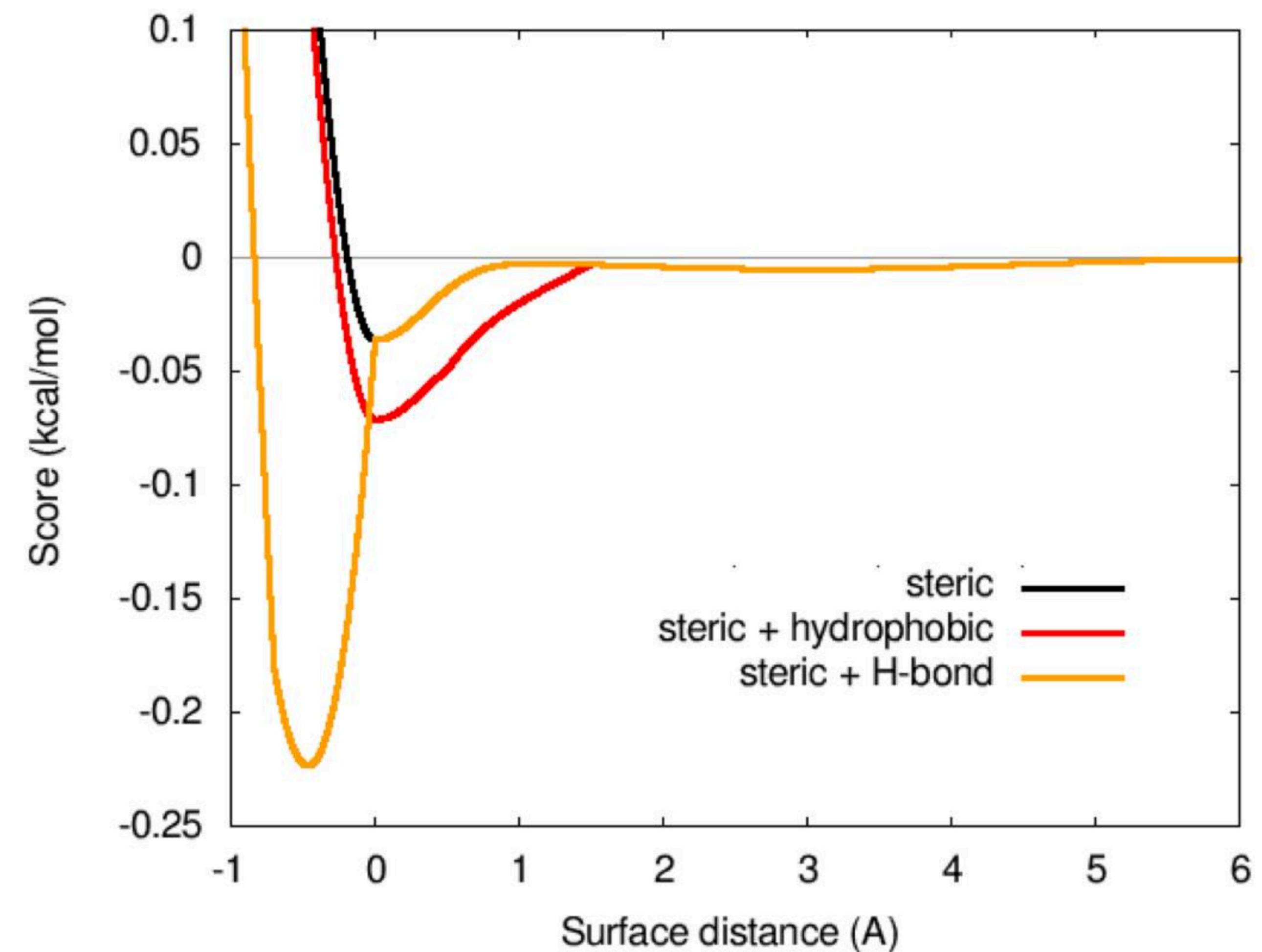
$$\text{gauss}_2(d) = e^{-((d-3\text{\AA})/2\text{\AA})^2}$$

$$\text{repulsion}(d) = \begin{cases} d^2, & \text{if } d < 0 \\ 0, & \text{if } d \geq 0 \end{cases}$$

- Hydrophobic and H-bond terms are linear interpolations

The Vina Scoring Function

Weight	Term
-0.0356	gauss ₁
-0.00516	gauss ₂
0.840	repulsion
-0.0351	hydrophobic
-0.587	hydrogen bonding
<hr/>	
0.0585	N_{rot}



References

- Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *Journal of Computational Chemistry* 1998, 19 (14), 1639–1662.
- Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization and Multithreading. *Journal of Computational Chemistry* 2010, 31 (2), 455–461.