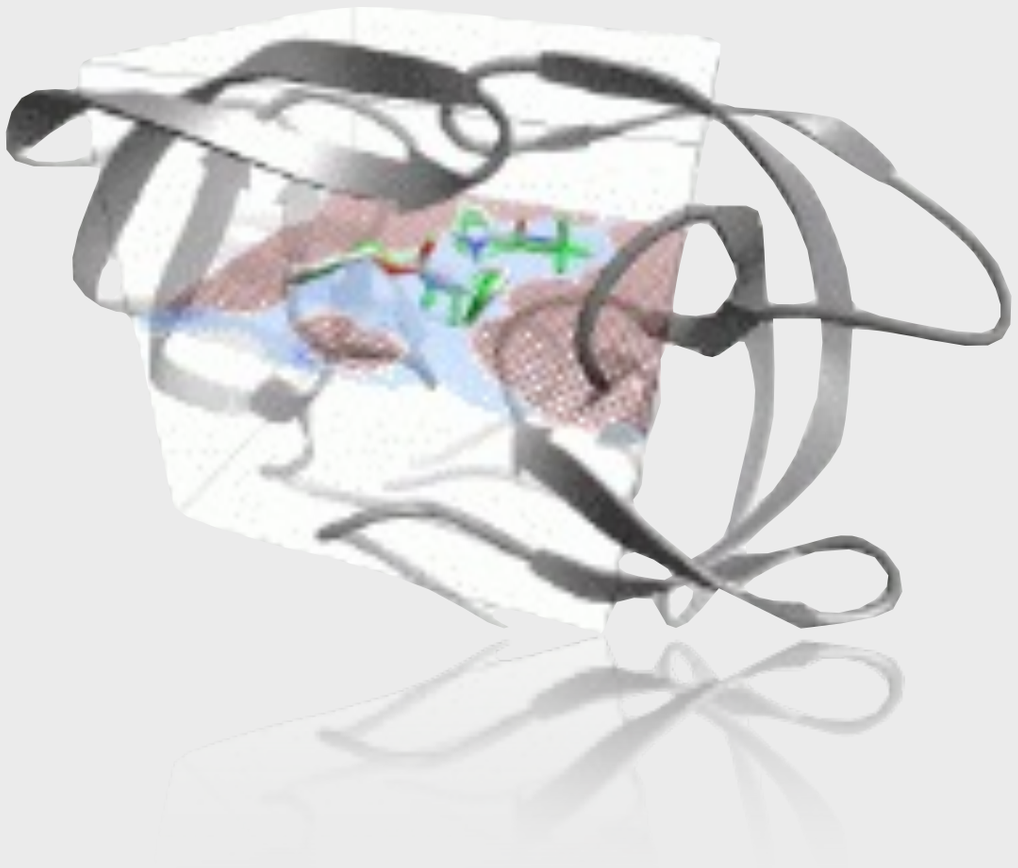
Docking of small molecules. Vina.



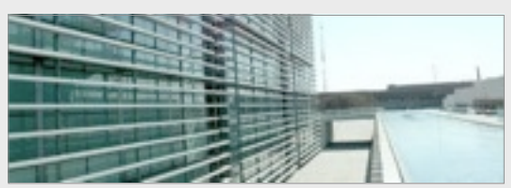
Marc A. Marti-Renom

http://bioinfo.cipf.es/sgu/

Structural Genomics Unit

Bioinformatics Department

Prince Felipe Resarch Center (CIPF), Valencia, Spain



PRINCIPE FELIPE

CENTRO DE INVESTIGACION

Objective

TO LEARN HOW-TO USE AutoDock VINA FOR DOCKING SMALL MOLECULES IN THE SURFACE OF A PROTEIN

Nomenclature

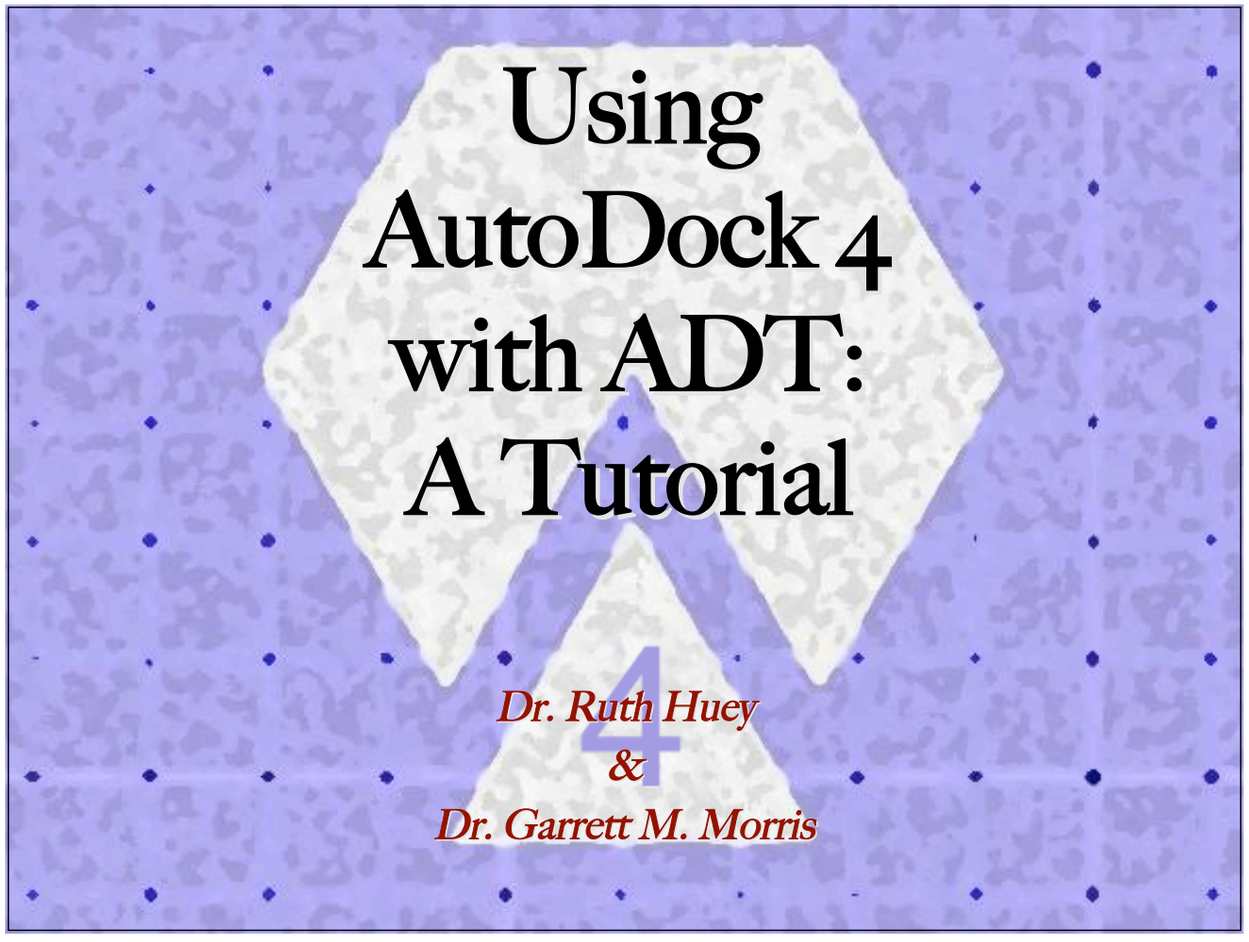
Ligand: Structure (usually a small molecule) that binds to the binding site.

Receptor: Structure (usually a protein) that contains the active binding site.

Binding site: Set of aminoacids (residues) that physically interact with the lingad (usually within 6 Ångstroms).

DISCLAIMER!

Credit should go to Dr. Oleg Trott, Dr. Ruth Huey and Dr. Garret M. Morris



Software News and Update
AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading

OLEG TROTT, ARTHUR J. OLSON
Department of Molecular Biology, The Scripps Research Institute, La Jolla, California

Received 3 March 2009; Accepted 21 April 2009
DOI 10.1002/jcc.21334
Published online in Wiley InterScience (www.interscience.wiley.com).

Abstract: AutoDock Vina, a new program for molecular docking and virtual screening, is presented. AutoDock Vina achieves an approximately two orders of magnitude speed-up compared with the molecular docking software previously developed in our lab (AutoDock 4), while also significantly improving the accuracy of the binding mode predictions, judging by our tests on the training set used in AutoDock 4 development. Further speed-up is achieved from parallelism, by using multithreading on multicore machines. AutoDock Vina automatically calculates the grid maps and clusters the results in a way transparent to the user.

© 2009 Wiley Periodicals, Inc. J Comput Chem 00: 000–000, 2009

Key words: AutoDock; molecular docking; virtual screening; computer-aided drug design; multithreading; scoring function

Introduction

Molecular docking is a computational procedure that attempts to predict noncovalent binding of macromolecules or, more frequently, of a macromolecule (receptor) and a small molecule (ligand) efficiently, starting with their unbound structures, structures obtained from MD simulations, or homology modeling, etc. The goal is to predict the bound conformations and the binding affinity.

The prediction of binding of small molecules to proteins is of particular practical importance because it is used to screen virtual libraries of drug-like molecules to obtain leads for further drug development. Docking can also be used to try to predict the bound conformation of known binders, when the experimental holo structures are unavailable.[1]

One is interested in maximizing the accuracy of these predictions while minimizing the computer time they take, because the computational resources spent on docking are considerable. For example, hundreds of thousands of computers are used for running docking in `FightAIDS@Home` and similar projects.[2]

Theory

In the spectrum of computational approaches to modeling receptor-ligand binding,

a. molecular dynamics with explicit solvent,
b. molecular dynamics and molecular mechanics with implicit solvent, and
c. molecular docking

can be seen as making an increasing trade-off of the representational detail for computational speed.[3]

Among the assumptions made by these approaches is the commitment to a particular protonation state of and charge distribution in the molecules that do not change between, for example, their bound and unbound states. Additionally, docking generally assumes much or all of the receptor rigid, the covalent lengths, and angles constant, while considering a chosen set of covalent bonds freely rotatable (referred to as active rotatable bonds here).

Importantly, although molecular dynamics directly deals with energies (referred to as force fields in chemistry), docking is ultimately interested in reproducing chemical potentials, which determine the bound conformation preference and the free energy of binding. It is a qualitatively different concept governed not only by the minima in the energy profile but also by the shape of the profile and the temperature.[4,5]

Docking programs generally use a scoring function, which can be seen as an attempt to approximate the standard chemical potentials of the system. When the superficially physics-based terms like the 6–12 van der Waals interactions and Coulomb energies are used in the scoring function, they need to be significantly empirically weighted, in part, to account for this difference between energies and free energies.[4,5]

Correspondence to: A.J. Olson; e-mail: olson@scripps.edu
Contract/grant sponsor: NIH; contract/grant number: 2R01GM069832

http://autodock.scripps.edu

http://vina.scripps.edu

O. Trott, A. J. Olson, Journal of Computational Chemistry (2009)

Summary

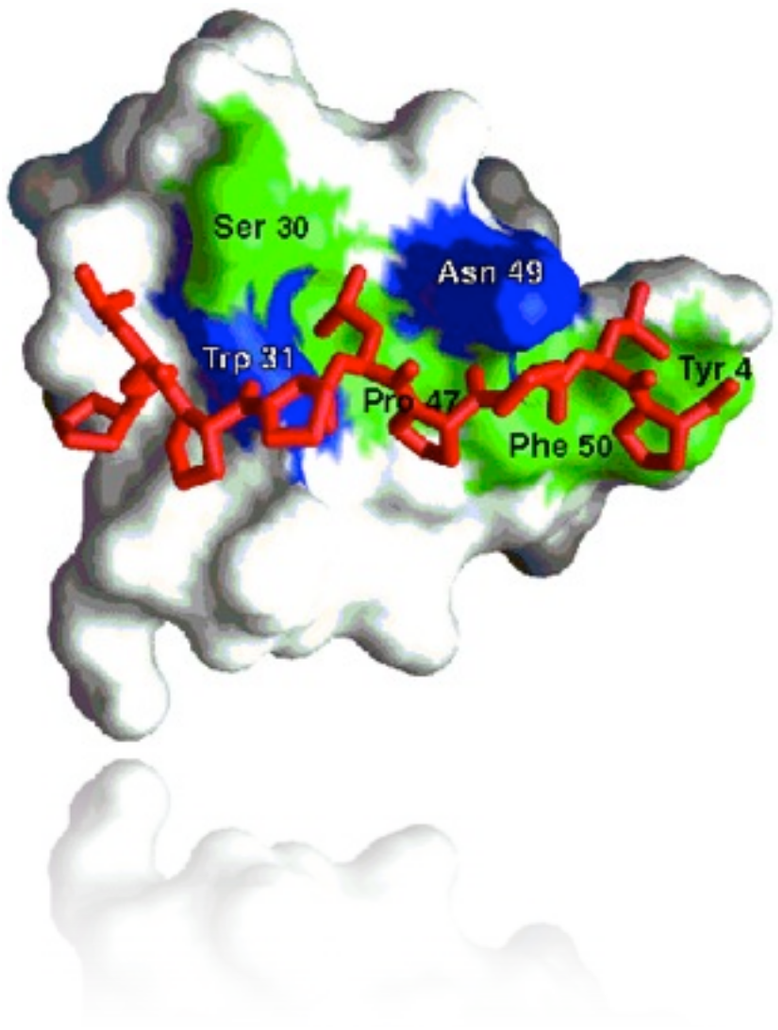
- INTRO
- DOCKING
- SEARCH METHODS
- EXAMPLE

- Vina 1.1.1 with ADT

What is docking?

Predicting the best ways two molecules interact.

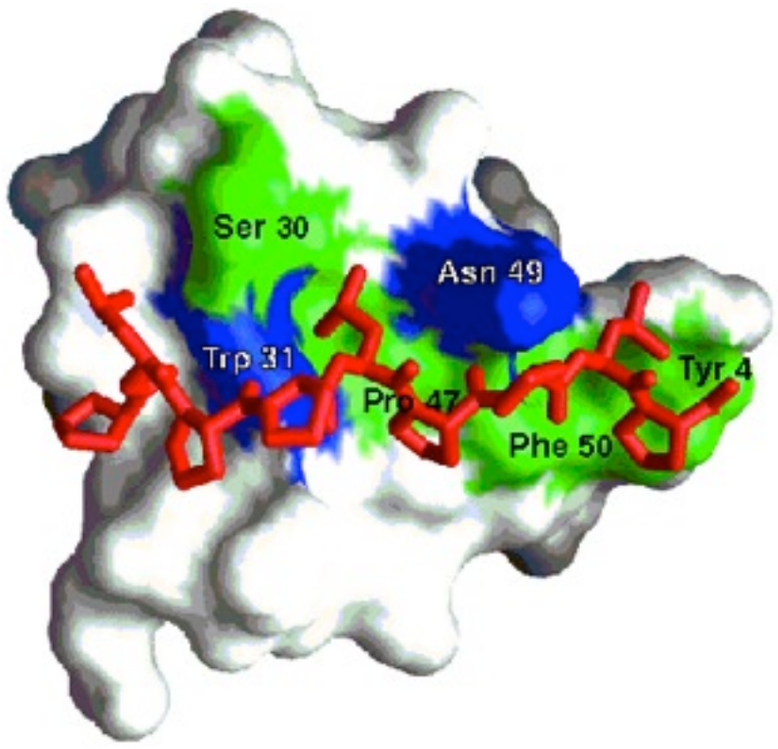
- Obtain the 3D structures of the two molecules
- Locate the best binding site (**Remember AnnoLyze? :-)**)
- Determine the best binding mode.



What is docking?

Predicting the **best** ways two molecules interact.

- We need to quantify or rank solutions
- We need a good scoring function for such ranking

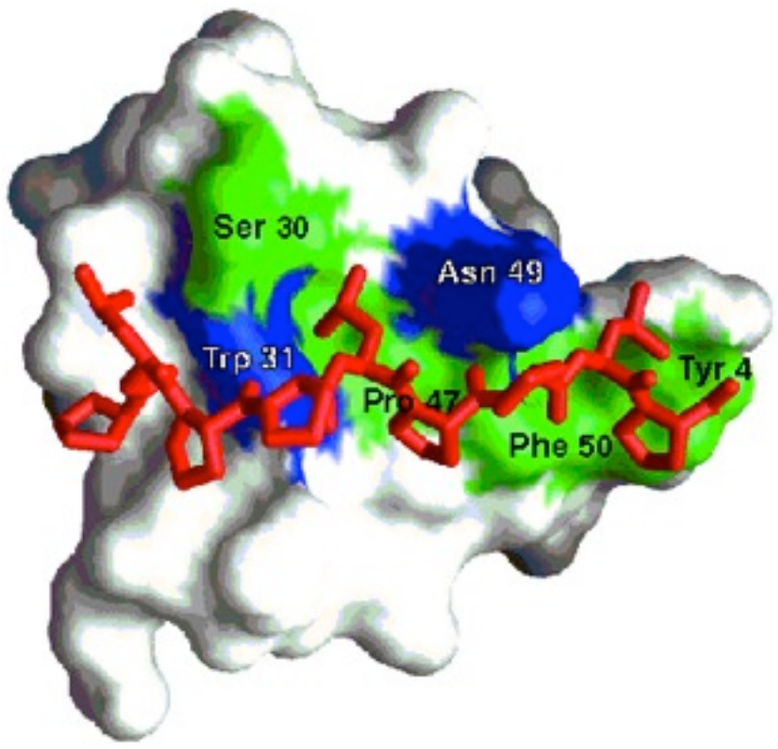


Scoring!

What is docking?

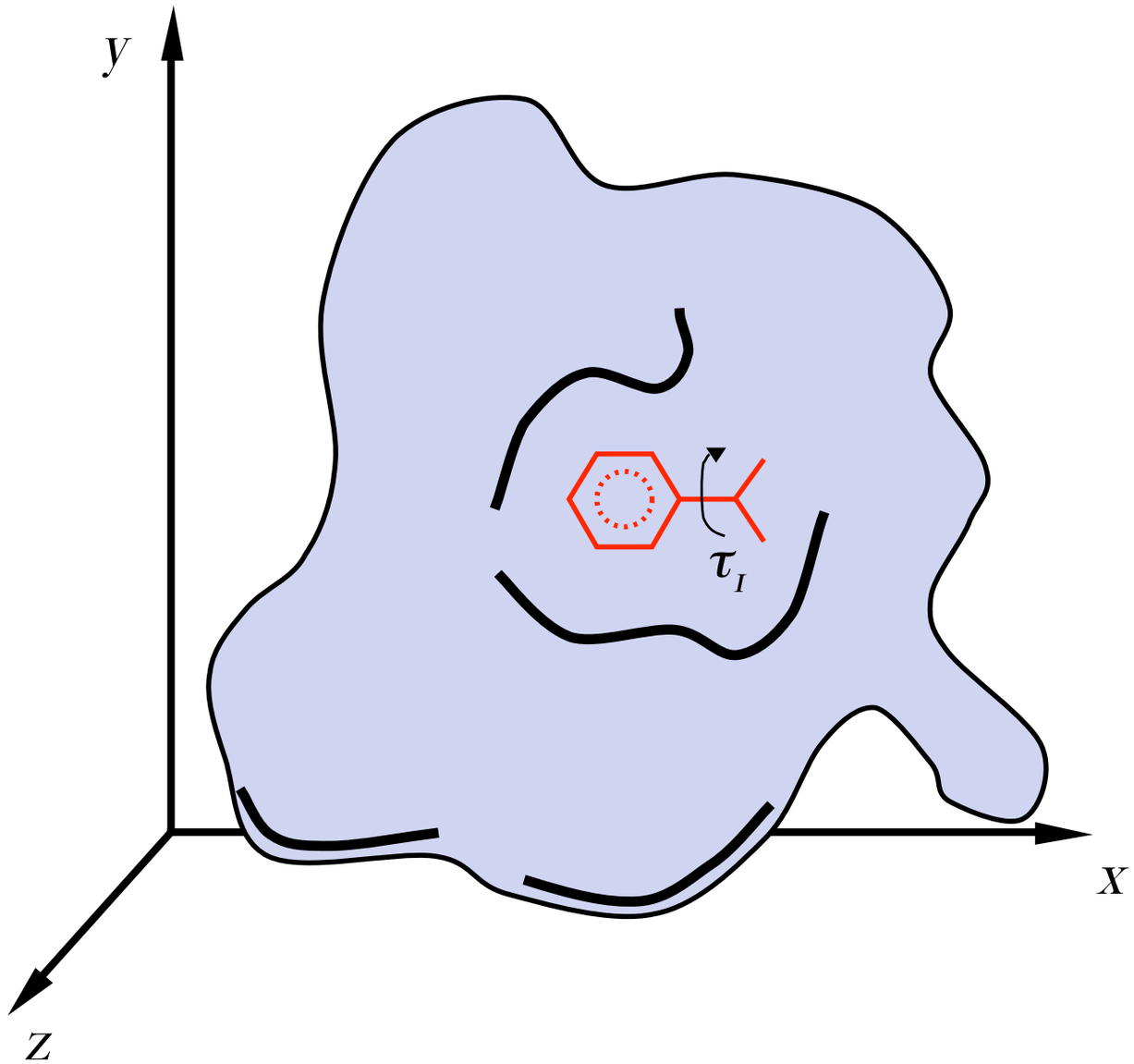
Predicting the best **ways** two molecules interact.

- X-ray and NMR structures are just ONE of the possible solutions
- There is a need for a search solution



Sampling!

REPRESENTATION



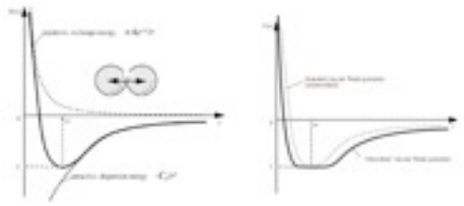
SCORING

AutoDock Vina

$$\Delta G_{binding} = \Delta G_{vdW} + \Delta G_{elec} + \Delta G_{hbond} + \Delta G_{desolv} + \Delta G_{tors}$$

- ΔG_{vdW}

  12-6 Lennard-Jones potential



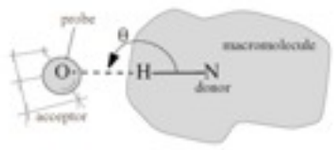
- ΔG_{elec}

  Coulombic with Solmajer-dielectric

$$\varepsilon(r) = A + \frac{B}{1 + ke^{-\lambda B r}}$$

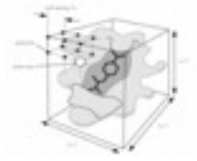
- ΔG_{hbond}

  12-10 Potential with Goodford Directionality



- ΔG_{desolv}

  Stouten Pairwise Atomic Solvation Parameters

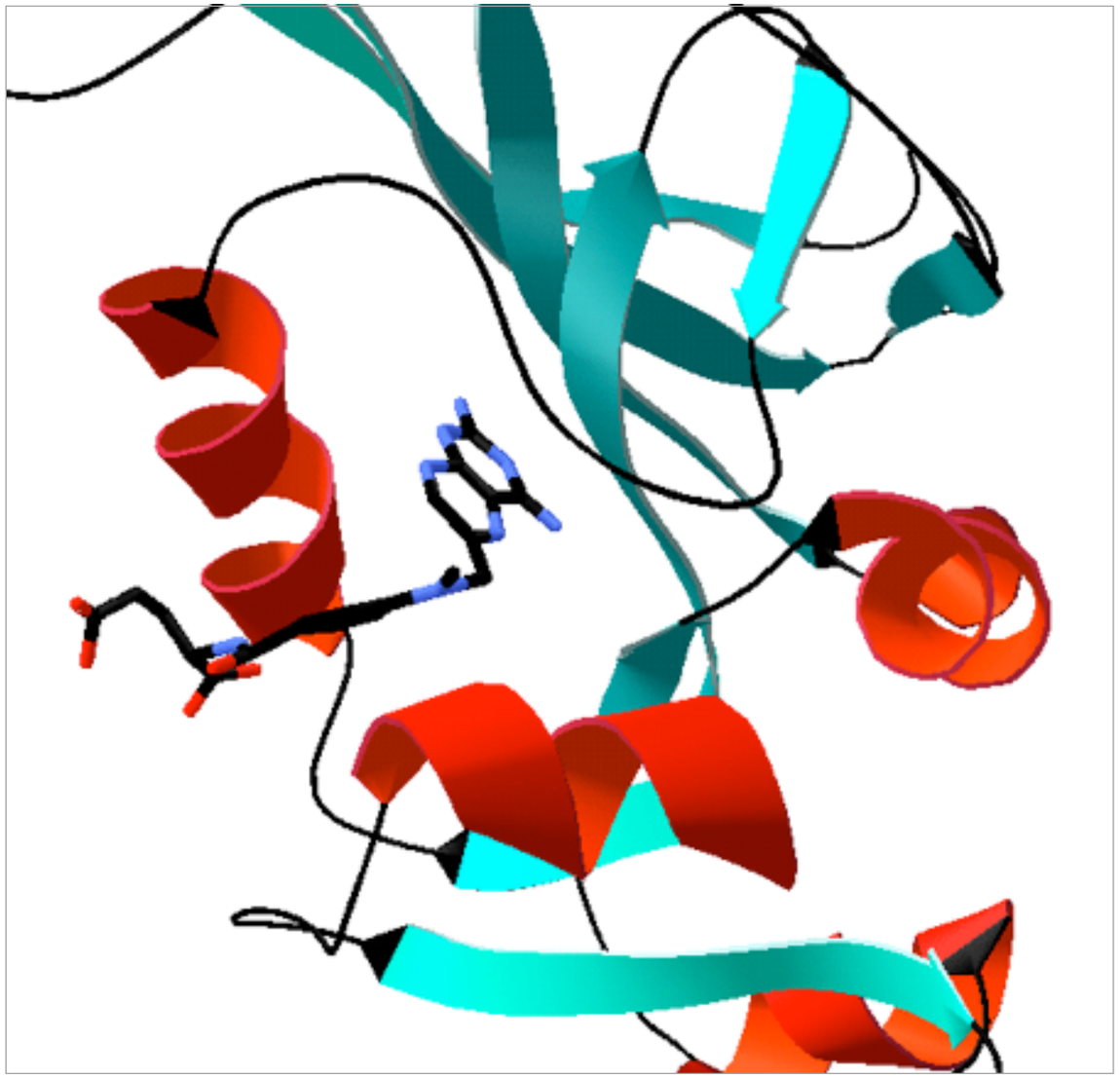


- ΔG_{tors}

  Number of rotatable bonds

PROBLEM!

Very CPU time consuming...



Dihidrofolate reductase with a metotrexate (4dfr.pdb)

$$N=T^{360/i}$$

N: number of conformations

T: number of rotable bonds

I: incremental degrees

Metotrexato

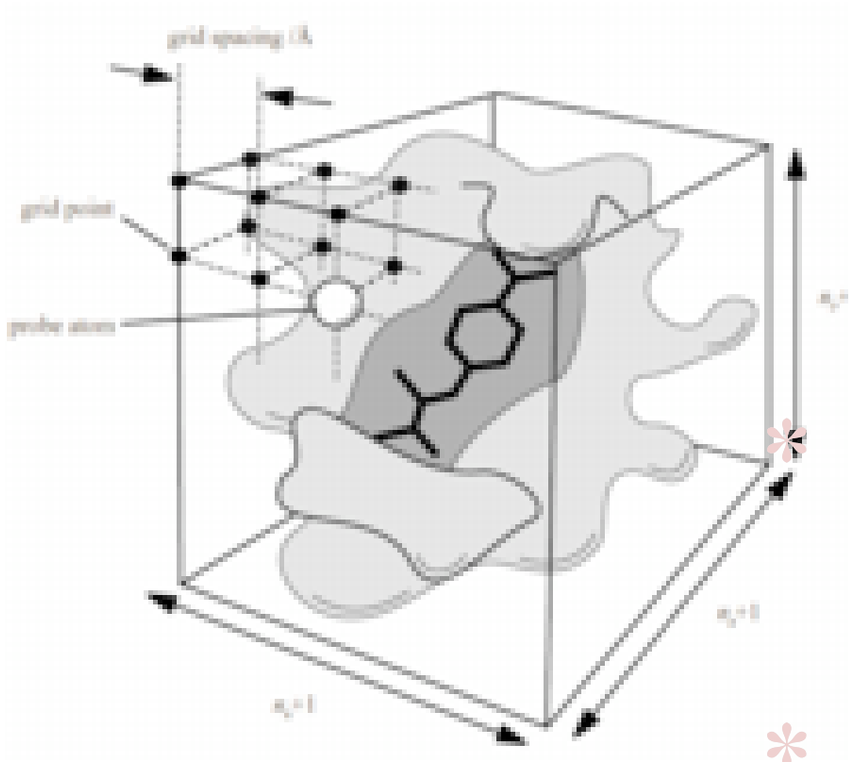
10 rotable bonds

30° increments (discrete)

10^{12} plausible conformations!

SOLUTION

Use of grid maps!

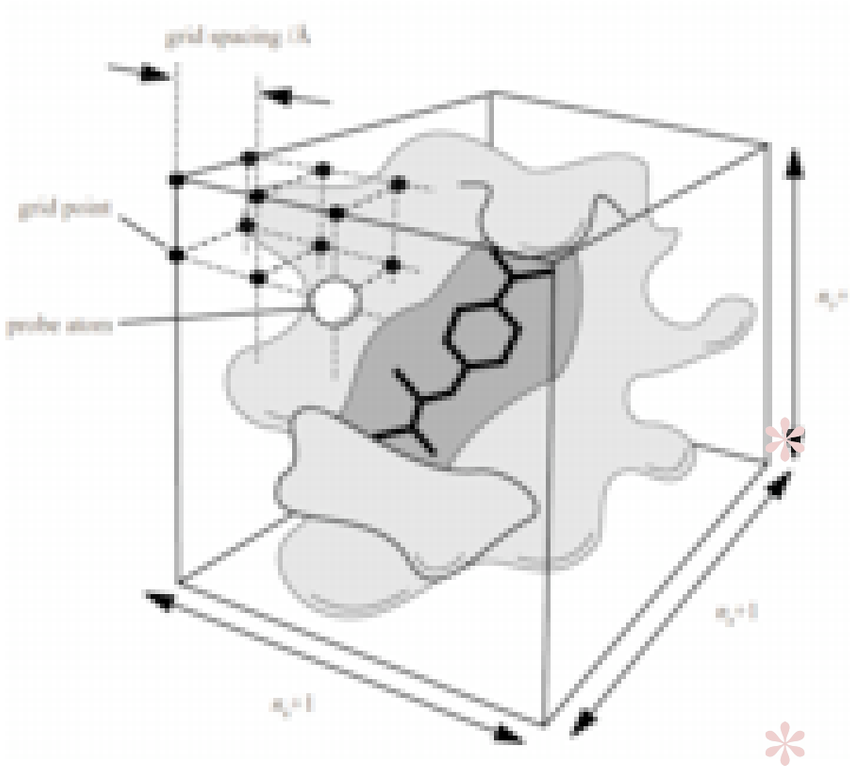


- Saves lots of time (compared to classical MM/MD)
- Need to map each atom to a grid point
- Limits the search space!

AutoGrid Vina

Use of grid maps!

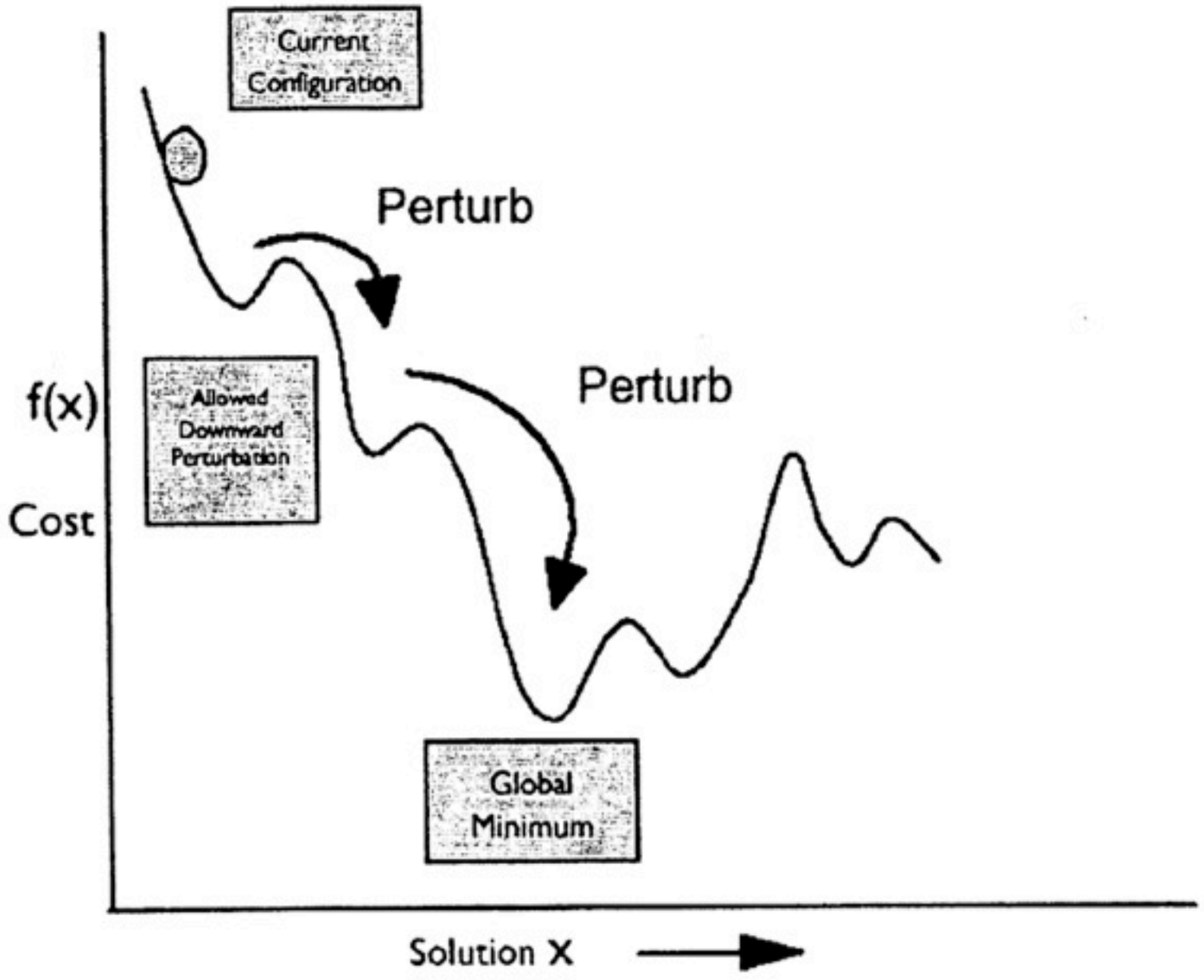
- Center of grid *
  - center of ligand
  - center of receptor
  - a selected atom or coordinate
- Box dimension *
- Grid resolution (spacing)
  - default 0.375 Angstroms
- Number of grid points (dimension)
  - use ONLY even numbers
- MAKE SURE ALL LIGAND IS INSIDE GRID AND CAN MOVE!



With VINA much simplified (*)

Search algorithms

Simulated Annealing

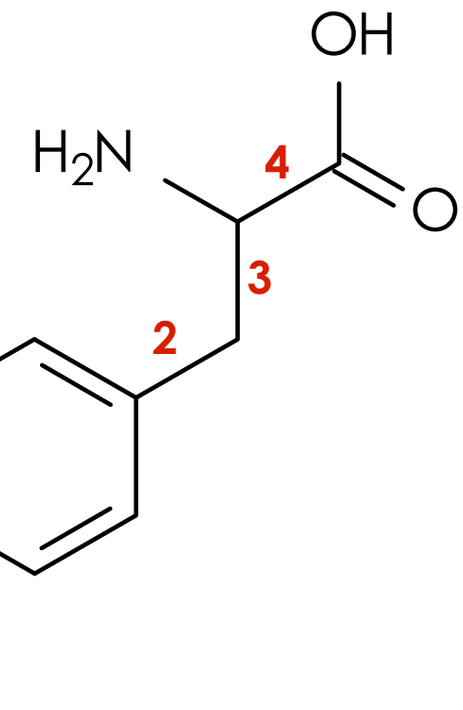


Search algorithms

Genetic Algorithm

Use of a Genetic Algorithm as a sampling method

- Each conformation is described as a set of rotational angles.
- 64 possible angles are allowed to each of the bond in the ligand.
- Each plausible dihedral angle is codified in a set of binary bits (2^6=64)
- Each conformation is codified by a so called chromosome with 4 × 6 bits (0 or 1)



111010.010110.001011.010010

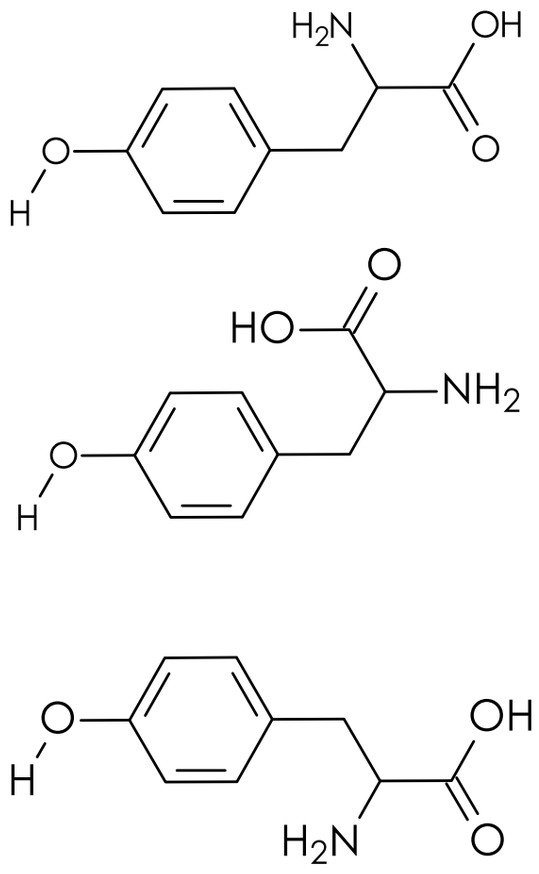
Φ_1 Φ_2 …

$$\Phi_1 = 1\times2^5 + 1\times2^4 + 1\times2^3 + 0\times2^2 + 1\times2^1 + 0\times2^0 = 58°$$

Search algorithms

Genetic Algorithm

Population (*ie*, set of chromosomes or configurations)



011010.010110.011010.010111 ← Chromosome

111010.010110.001011.010010

001010.010101.000101.010001

101001.101110.101010.001000

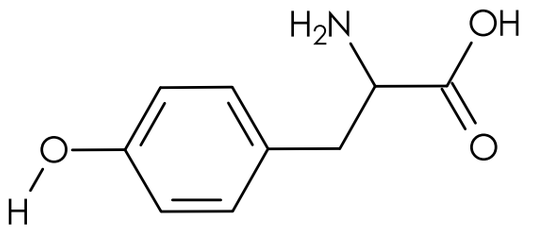
001010.101000.011101.001011

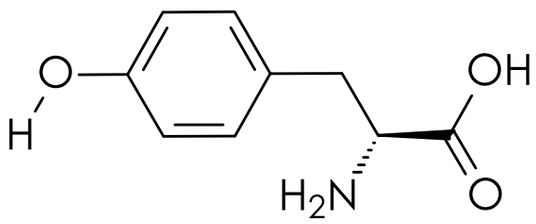
Gene

Search algorithms

Genetic Algorithm

Genetic operators...





011010.010110.011010.010111

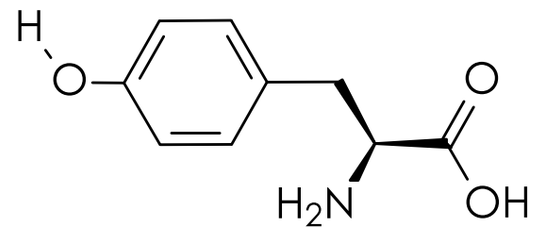
Single mutation

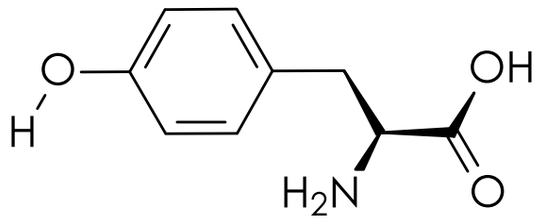
011010.011110.011110.010111

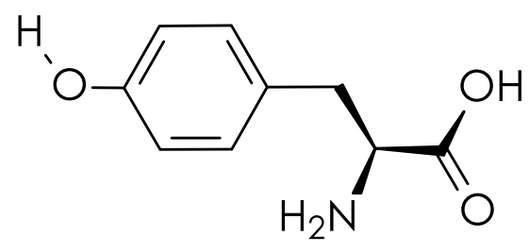
Search algorithms

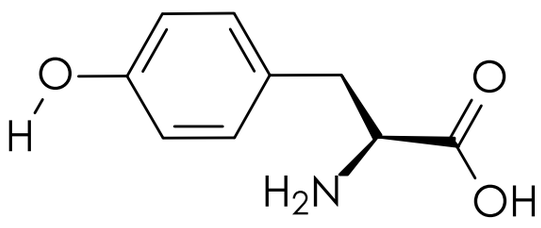
Genetic Algorithm

Genetic operators...









001010.010101.000101.010001

011010.010110.011010.010111

Recombination

001010.010101.011010.010111

011010.010110. 000101.010001

Search algorithms

Genetic Algorithm

Genetic operators...

011010.010110.011010.010111
111010.010110.001011.010010
001010.010101.000101.010001
101001.101110.101010.001000
001010.101000.011101.001011

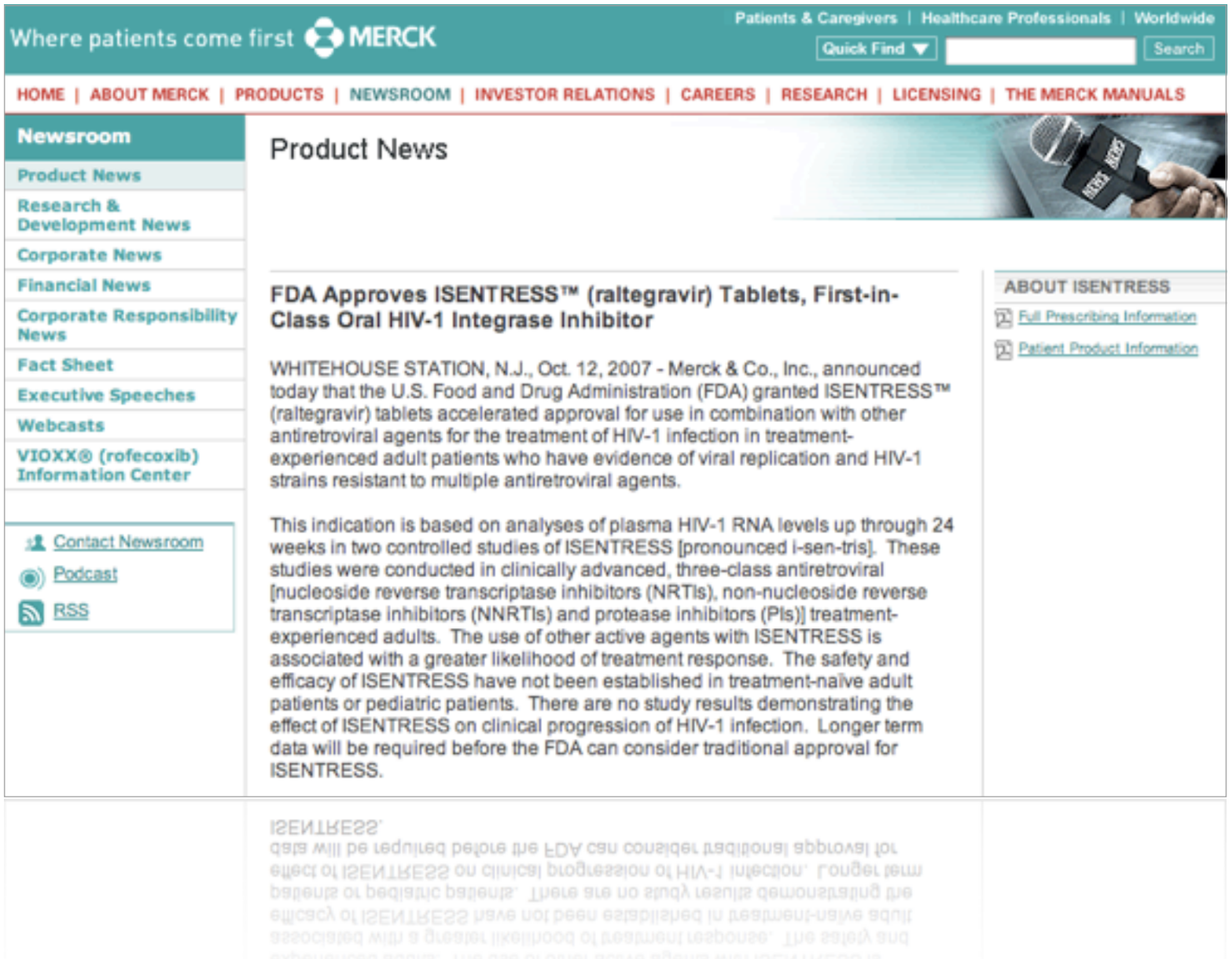
Migration

111110.010010.011110.010101
101010.110110.011011.011010
001010.010101.000101.010001
101101.101010.101011.001100
011010.100000.011001.101011

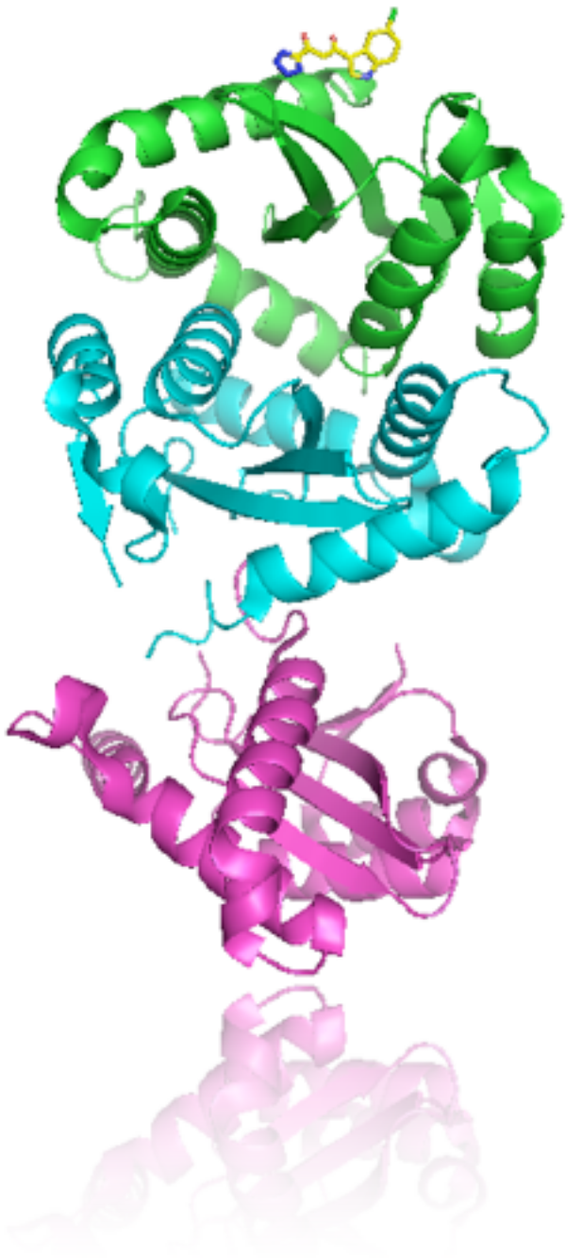
AutoDock Example

Discovery of a novel binding trench in HIV Integrase

Schames, J.R., R.H. Henchman, J.S. Siegel, C.A. Sotriffer, H. Ni, and J.A. McCammon, Discovery of a novel binding trench in HIV integrase. J Med Chem, 2004. 47(8): 1879-81



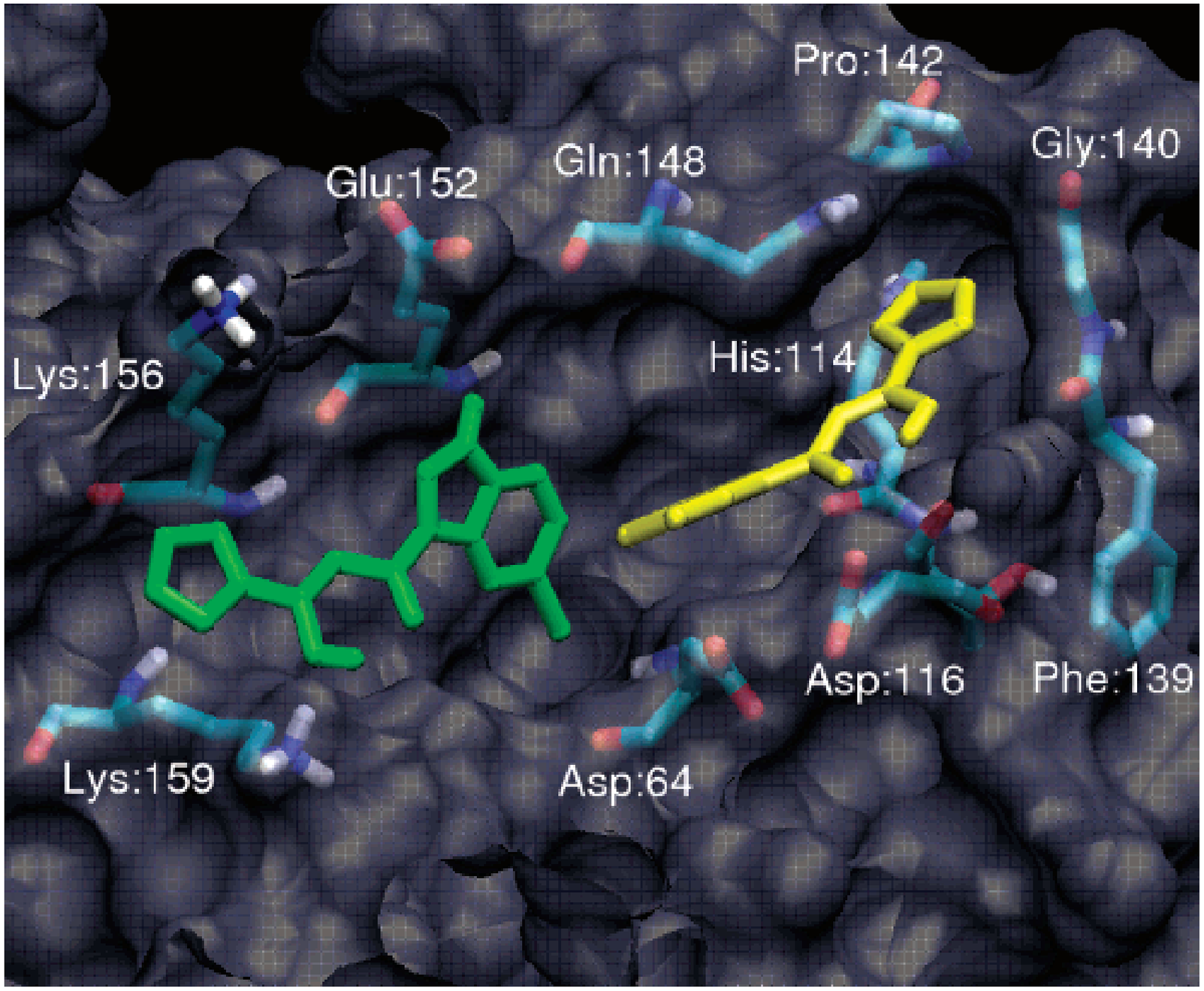
ISENTRESS example



- One structure known with 5CITEP
  - Not clear (low resolution)
  - Binding near to DNA interacting site
  - Loop near the binding
- Docking + Molecular Dynamics
  - AMBER snapshots
  - AutoDock flexible torsion thetetrazolering and indole ring.

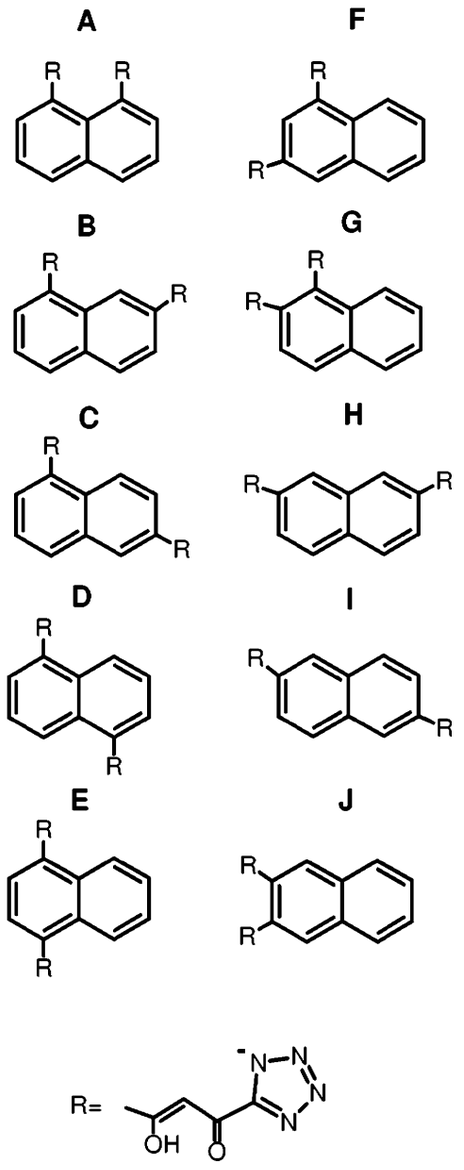
Schames, J.R., R.H. Henchman, J.S. Siegel, C.A. Sotriffer, H. Ni, and J.A. McCammon, Discovery of a novel binding trench in HIV integrase. J Med Chem, 2004. 47(8): 1879-81

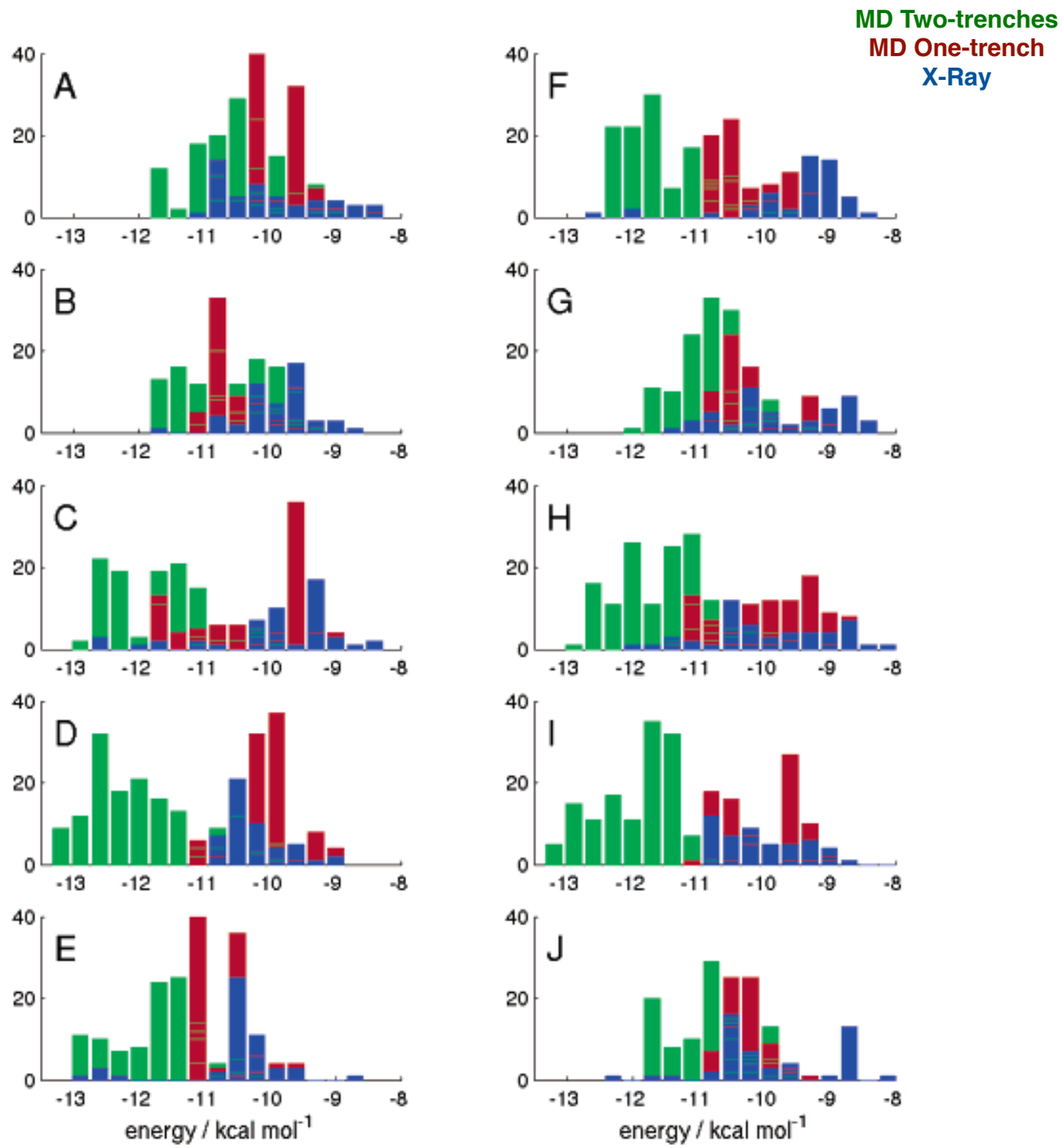
ISENTRESS example



Schames, J.R., R.H. Henchman, J.S. Siegel, C.A. Sotriffer, H. Ni, and J.A. McCammon, Discovery of a novel binding trench in HIV integrase. J Med Chem, 2004. 47(8): 1879-81

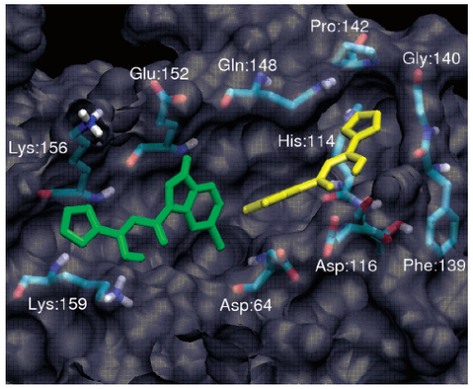
ISENTRESS example

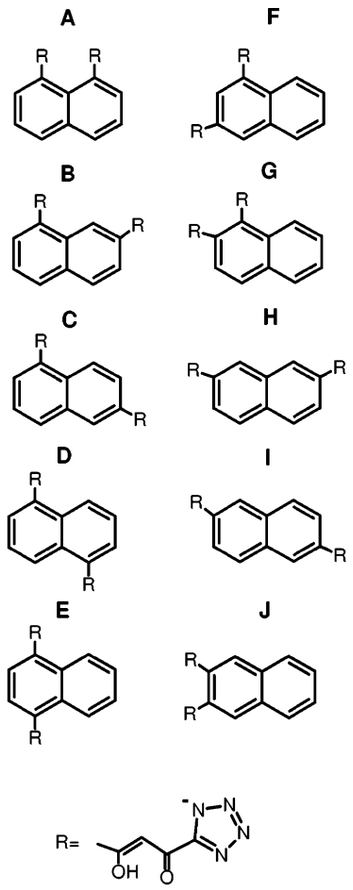


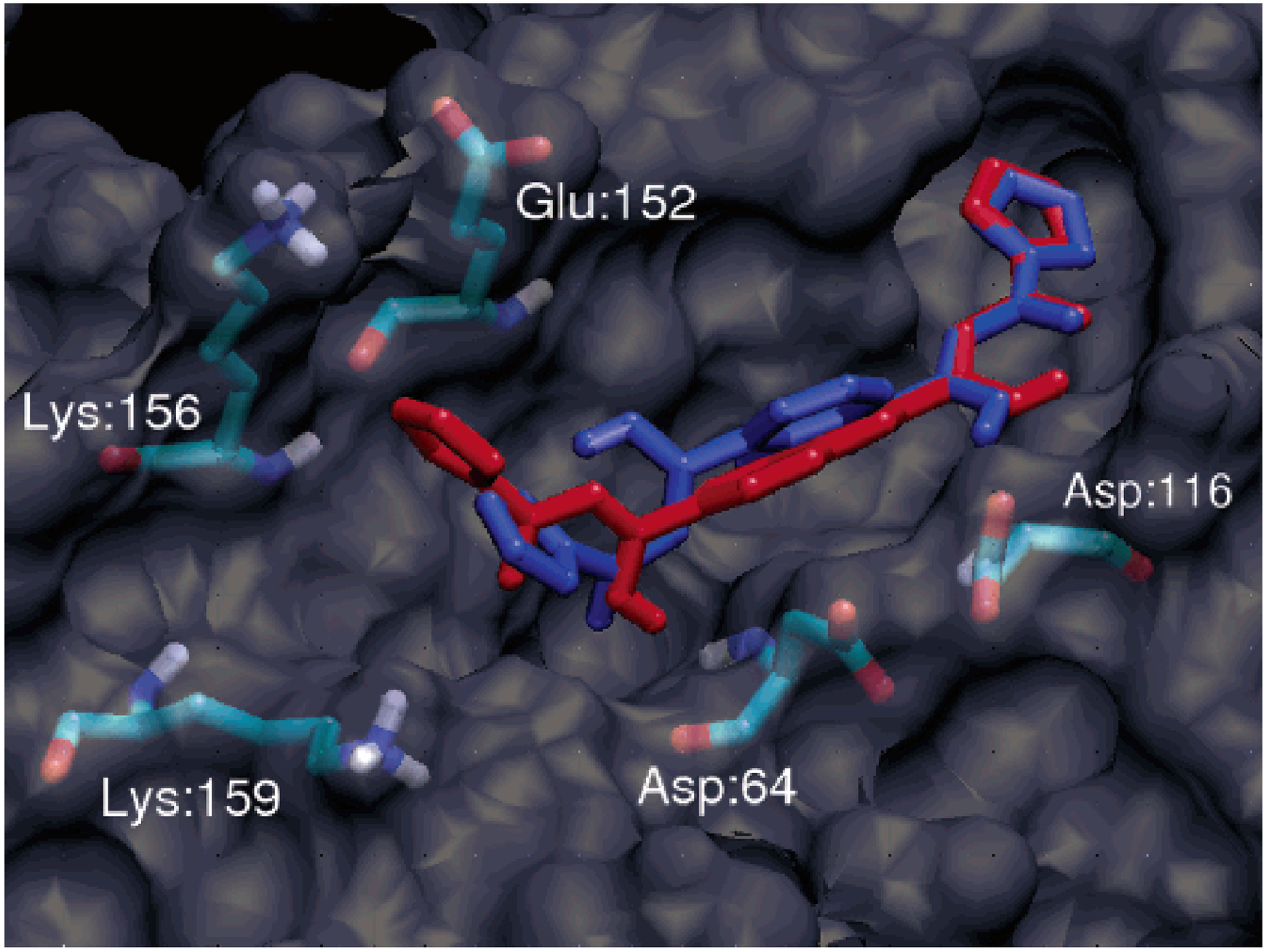


Schames, J.R., R.H. Henchman, J.S. Siegel, C.A. Sotriffer, H. Ni, and J.A. McCammon, Discovery of a novel binding trench in HIV integrase. J Med Chem, 2004. 47(8): 1879-81

ISENTRESS example







Schames, J.R., R.H. Henchman, J.S. Siegel, C.A. Sotriffer, H. Ni, and J.A. McCammon, Discovery of a novel binding trench in HIV integrase. J Med Chem, 2004. 47(8): 1879-81

ISENTRESS example

Newsroom

- Product News
- Research & Development News
- Corporate News
- Financial News
- Corporate Responsibility News
- Fact Sheet
- Executive Speeches
- Webcasts
- VIOXX® (rofecoxib) Information Center

Contact Newsroom

Podcast

RSS

Product News

FDA Approves ISENTRESS™ (raltegravir) Tablets, First-in-Class Oral HIV-1 Integrase Inhibitor

WHITEHOUSE STATION, N.J., Oct. 12, 2007 - Merck & Co., Inc., announced today that the U.S. Food and Drug Administration (FDA) granted ISENTRESS™ (raltegravir) tablets accelerated approval for use in combination with other antiretroviral agents for the treatment of HIV-1 infection in treatment-experienced adult patients who have evidence of viral replication and HIV-1 strains resistant to multiple antiretroviral agents.

This indication is based on analyses of plasma HIV-1 RNA levels up through 24 weeks in two controlled studies of ISENTRESS [pronounced i-sen-tris]. These studies were conducted in clinically advanced, three-class antiretroviral [nucleoside reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs) and protease inhibitors (PIs)] treatment-experienced adults. The use of other active agents with ISENTRESS is associated with a greater likelihood of treatment response. The safety and efficacy of ISENTRESS have not been established in treatment-naïve adult patients or pediatric patients. There are no study results demonstrating the effect of ISENTRESS on clinical progression of HIV-1 infection. Longer term data will be required before the FDA can consider traditional approval for ISENTRESS.

ABOUT ISENTRESS

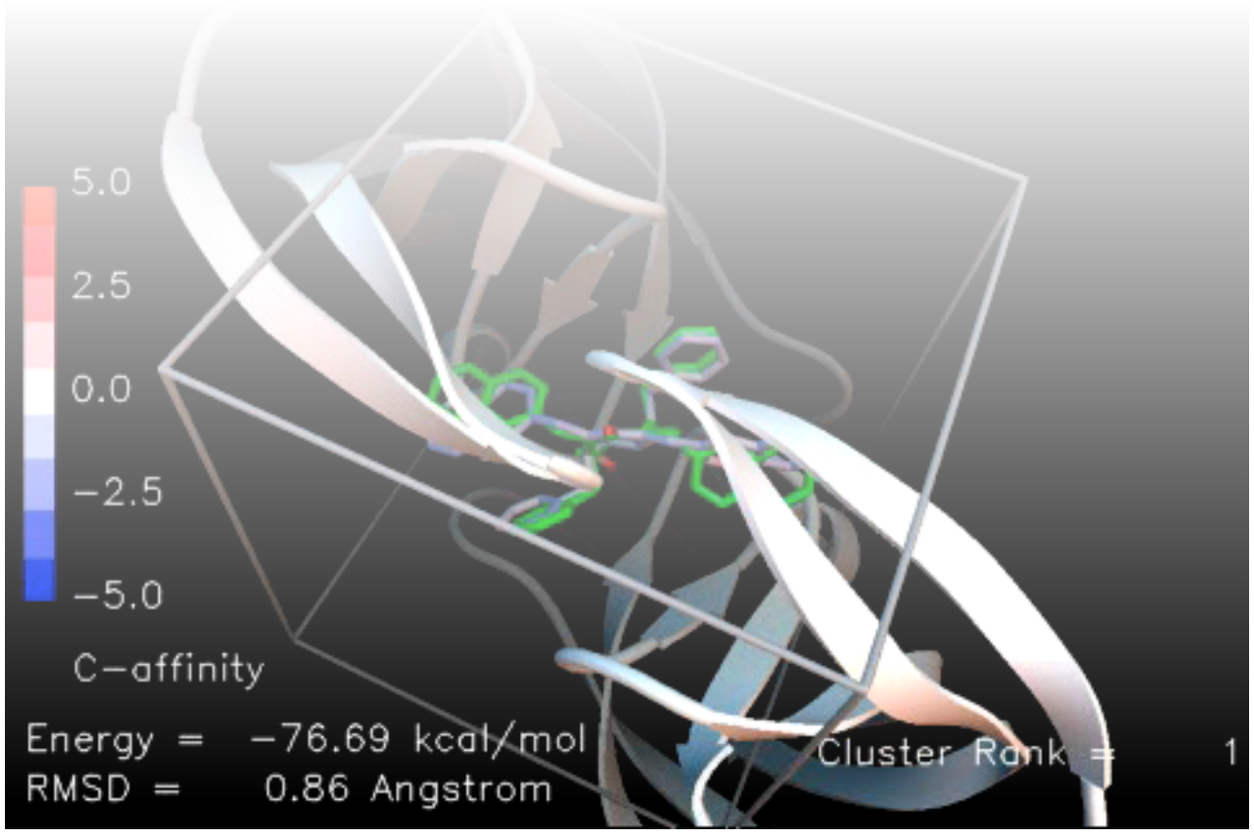
Full Prescribing Information

Patient Product Information

Vina 1.1.1

Goodsell, D. S. and Olson, A. J. (1990), Automated Docking of Substrates to Proteins by Simulated Annealing Proteins:Structure, Function and Genetics., 8: 195-202.
Morris, G. M., et al. (1996), Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4 J. Computer-Aided Molecular Design, 10: 293-304.
Morris, G. M., et al. (1998), Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function J. Computational Chemistry, 19: 1639-1662.
Huey, R., et al. (2007), A Semiempirical Free Energy Force Field with Charge-Based Desolvation J. Computational Chemistry, 28: 1145-1152.

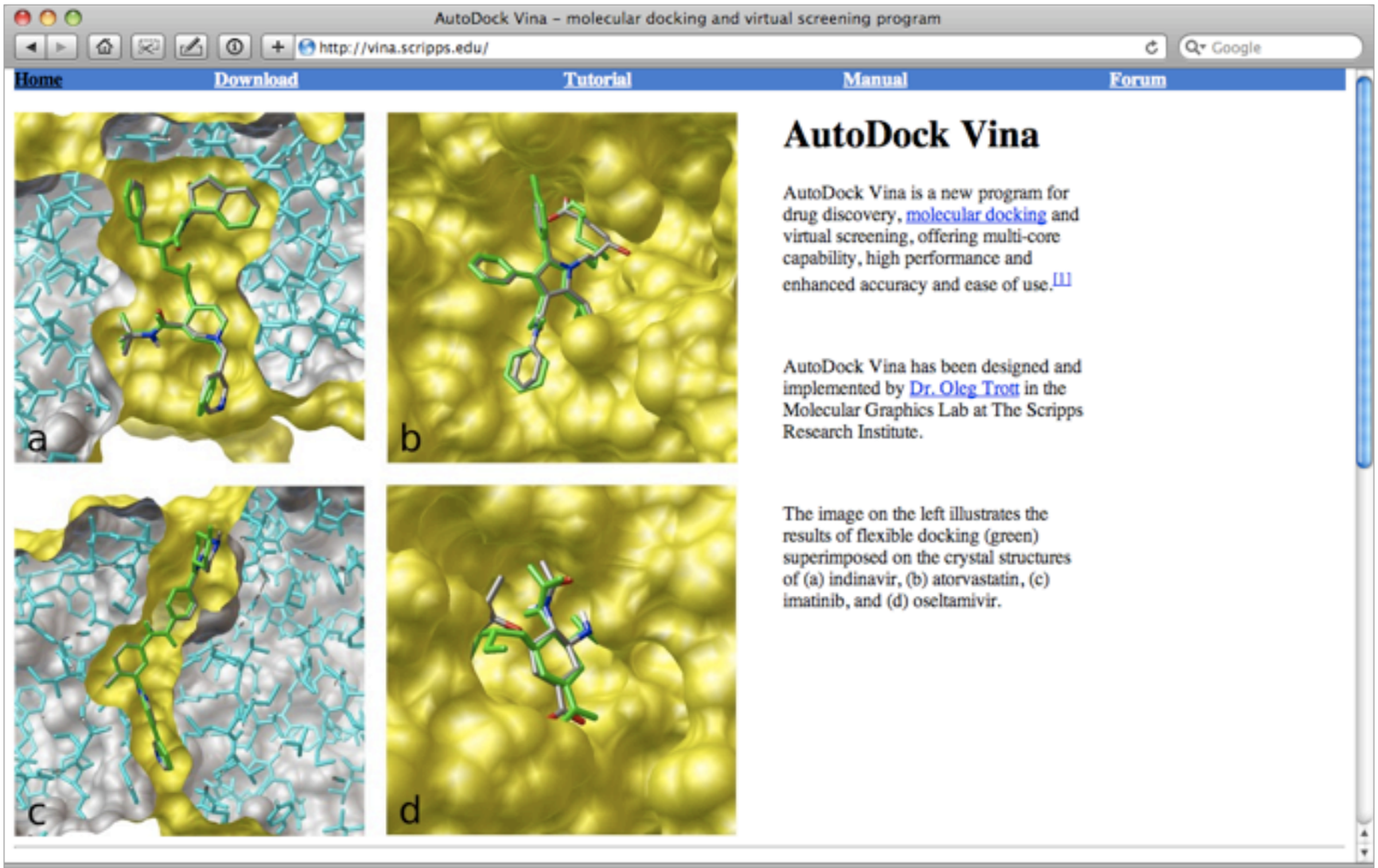
Vina 1.1.1



Goodsell, D. S. and Olson, A. J. (1990), Automated Docking of Substrates to Proteins by Simulated Annealing Proteins:Structure, Function and Genetics., 8: 195-202.
Morris, G. M., et al. (1996), Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4 J. Computer-Aided Molecular Design, 10: 293-304.
Morris, G. M., et al. (1998), Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function J. Computational Chemistry, 19: 1639-1662.
Huey, R., et al. (2007), A Semiempirical Free Energy Force Field with Charge-Based Desolvation J. Computational Chemistry, 28: 1145-1152.

Vina 1.1.1

Where to get help...



http://vina.scripps.edu

Vina 1.1.1

Alternatives

Progressive building

- FLEXX
- DOCK
- GROW
- GroupBUILD
- LUDI
- LEGEND
- SPROUT
- BUILDER
- GENSTAR

Conformational search

- MIMUMBA
- COBRA
- WIZRAD

Binding site description

- GRID

Genetic algorithms

- GOLD
- Others

Virtual screening

- AutoDOCK

Molecular dynamcis

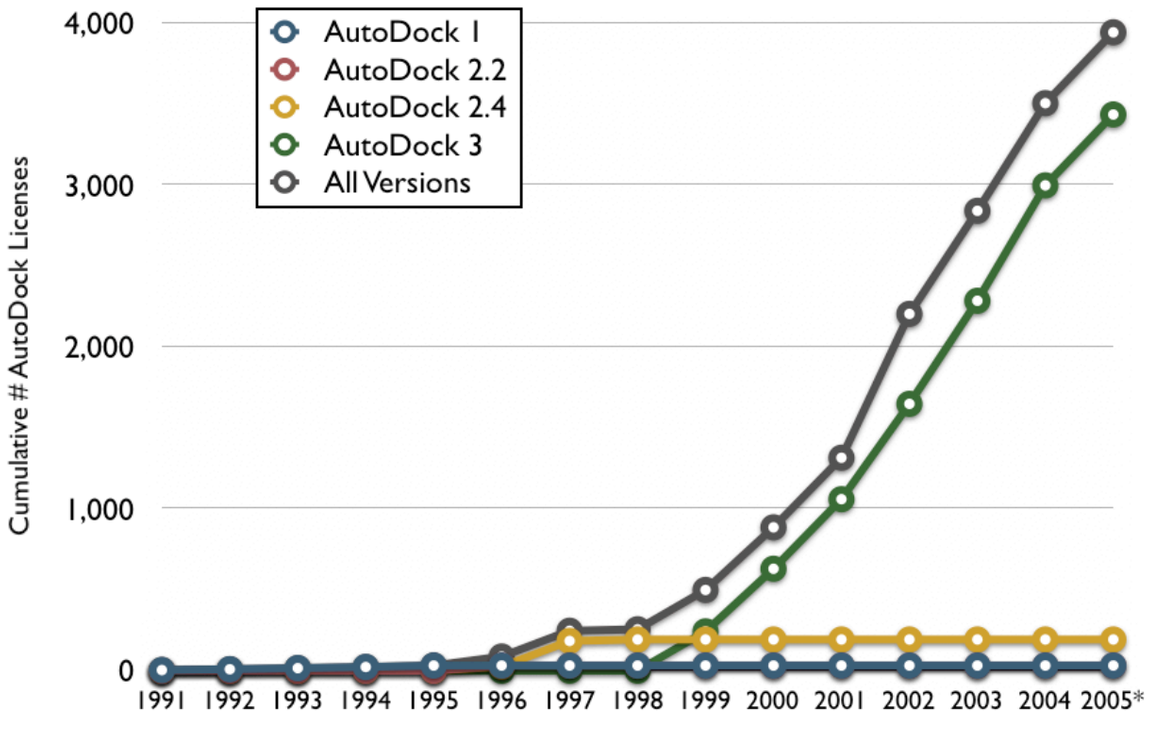
- MCSS
- CONCEPTS

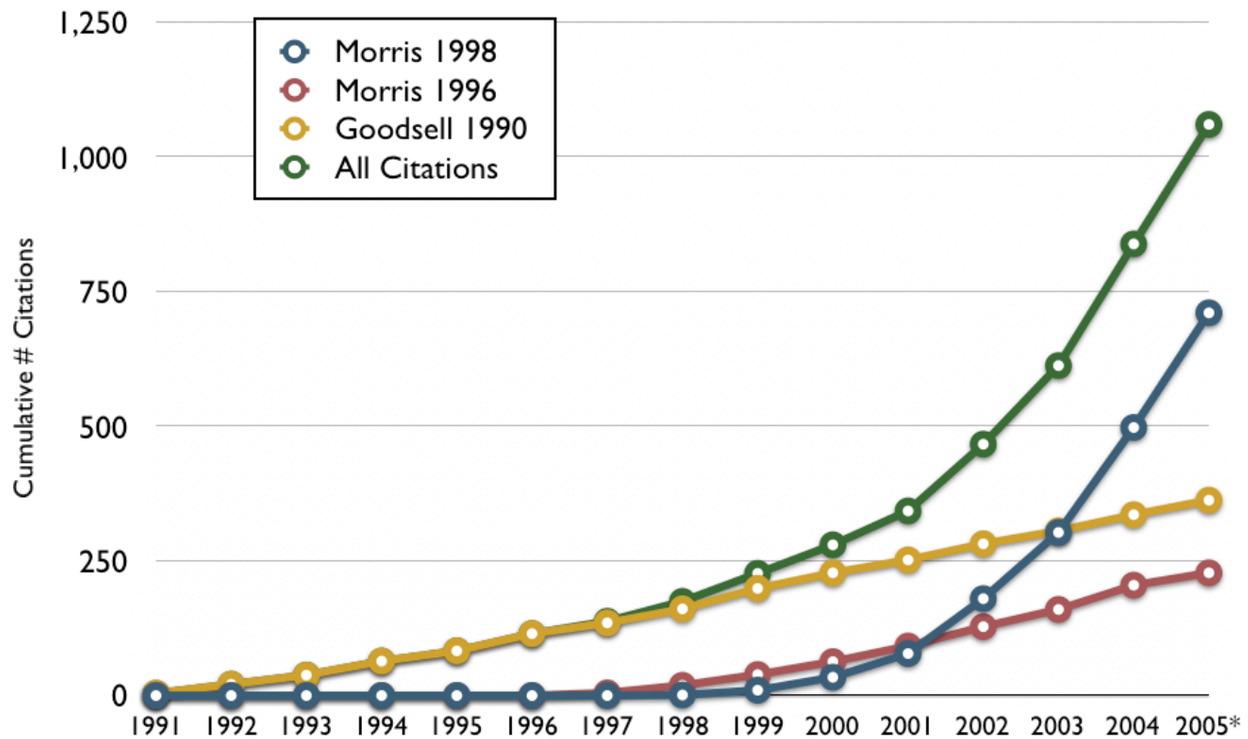
Databases

- CAVEAT
- FOUNDATION
- CLIX
- NEWLEAD
- LEAPFROG

AutoDock 4.0

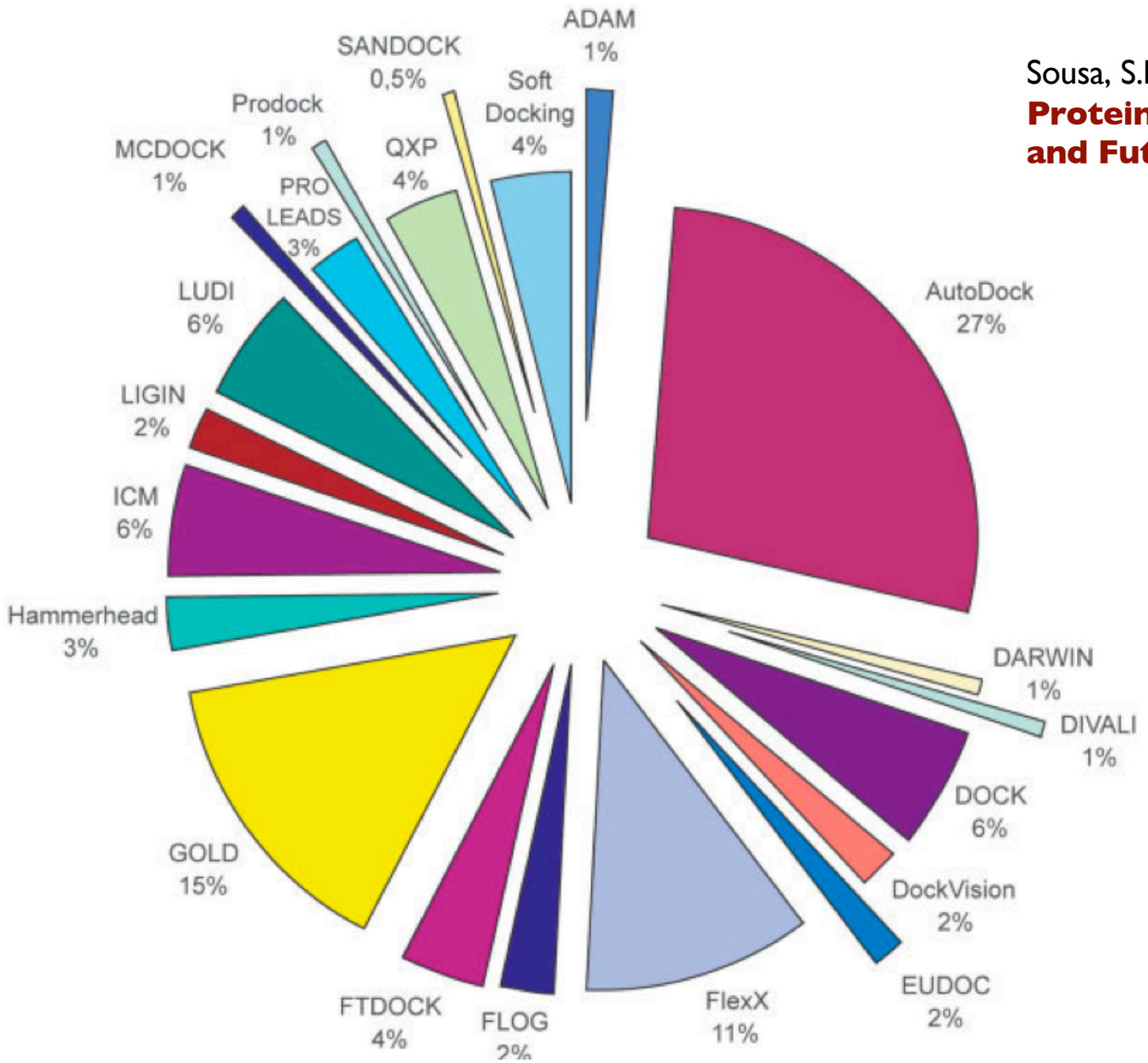
Why AutoDock over others





AutoDock 4.0

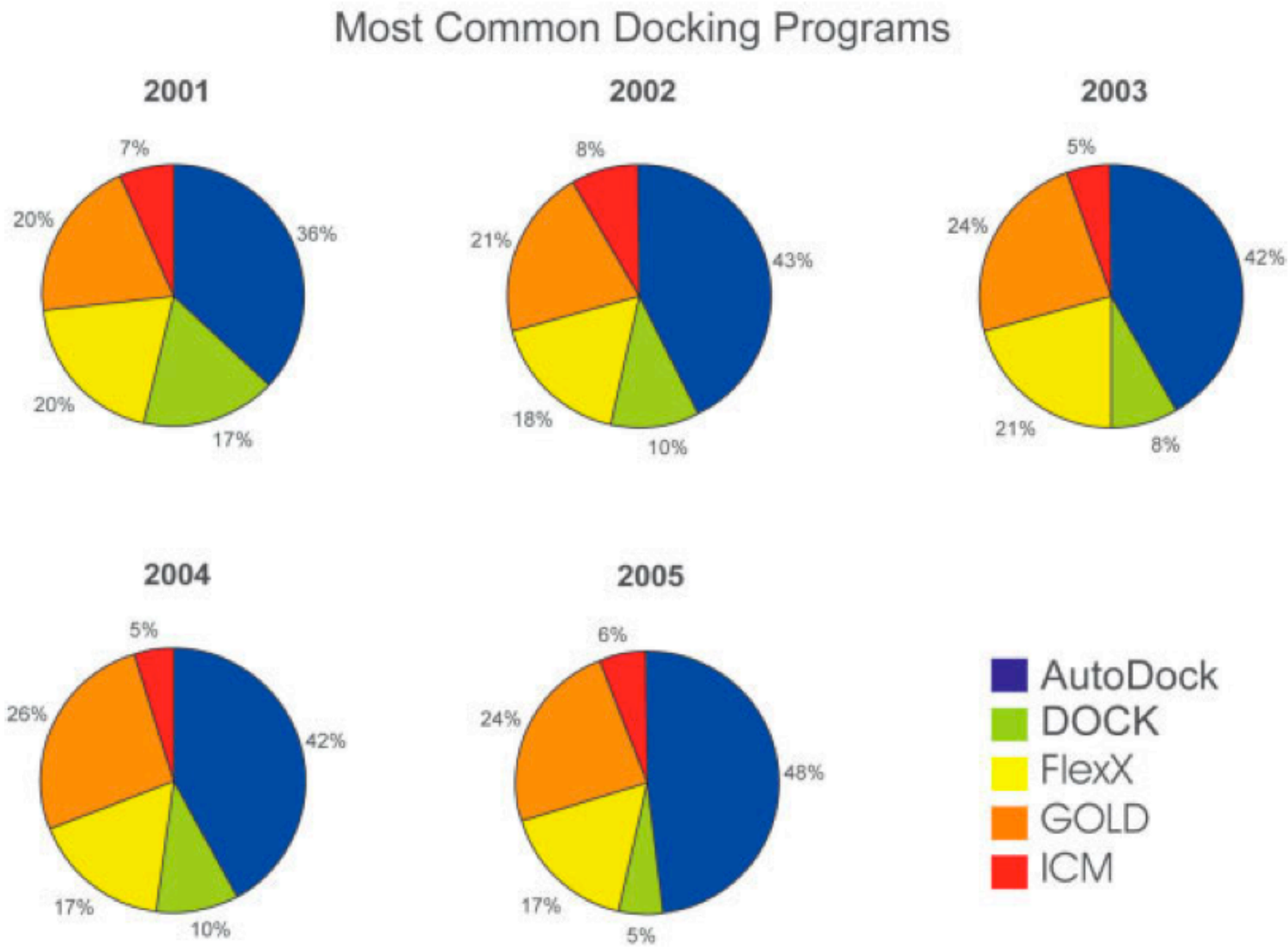
Why AutoDock over others



Sousa, S.F., Fernandes, P.A. & Ramos, M.J. (2006)
Protein-Ligand Docking: Current Status and Future Challenges *Proteins*, **65**:15-26

AutoDock 4.0

Why AutoDock over others



AutoDock / Vina

Practical considerations

- What problem does AutoDock solve?
  - *Flexible* ligands (4.0 *flexible* protein).
- What range of problems is feasible?
  - Depends on the search method:
    - LGA > GA >> SA >> LS
    - SA : can output trajectories, D < about 8 torsions.
    - LGA : D < about 8-32 torsions.
- When is AutoDock not suitable?
  - No 3D-structures are available;
  - Modelled structure of poor quality;
  - Too many (32 torsions, 2048 atoms, 22 atom types);
  - Target protein too flexible.

Vina 1.1.1

Things to know before using AutoDock

Ligand:

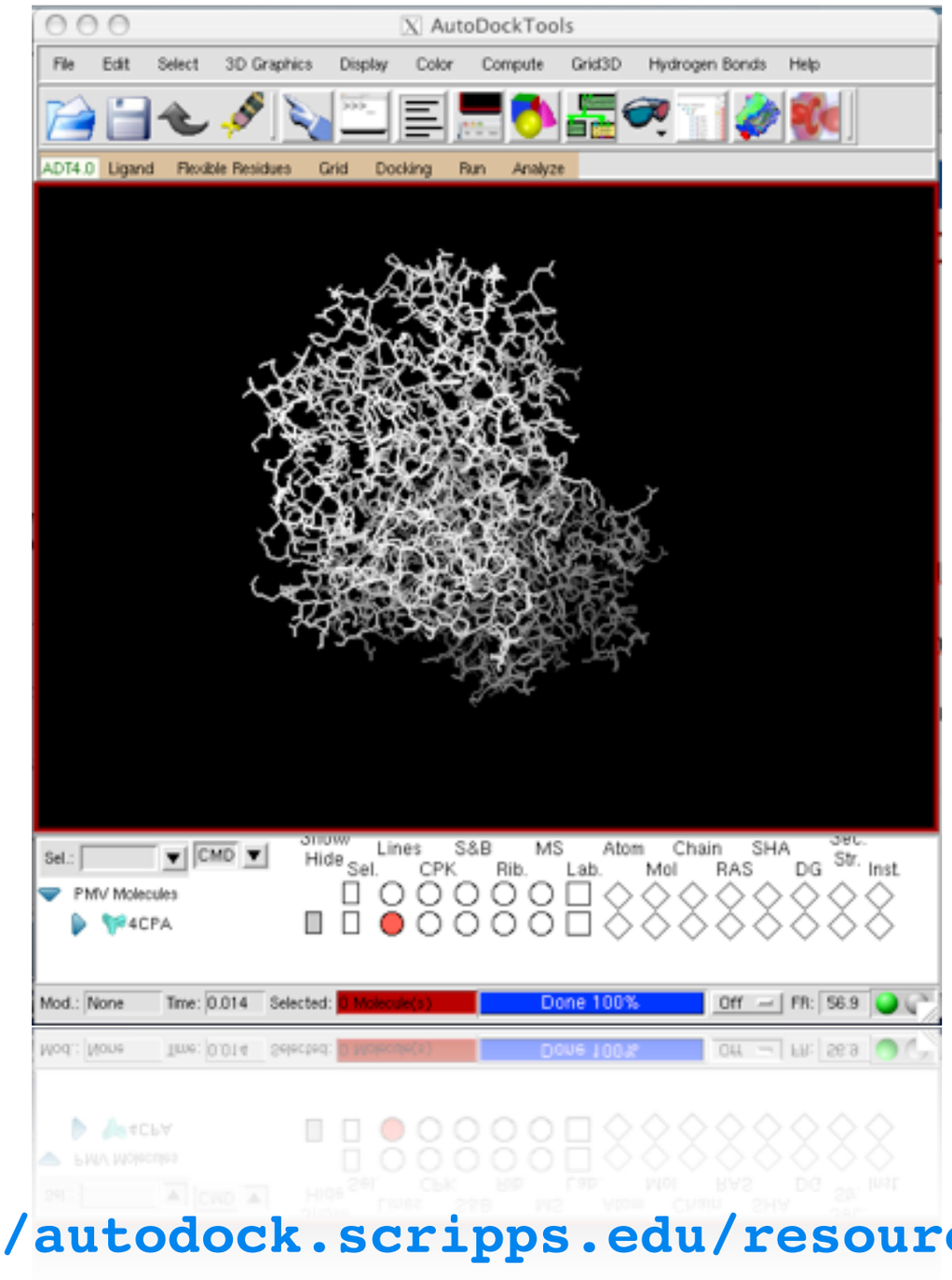
- Add all hydrogens, compute Gasteiger charges, and merge non-polar H; also assign AutoDock 4 atom types
- Ensure total charge corresponds to tautomeric state
- Choose torsion tree root & rotatable bonds

Macromolecule:

- Add all hydrogens, compute Gasteiger charges, and merge non-polar H; also assign AutoDock 4 atom types
- Assign Stouten atomic solvation parameters
- Optionally, create a flexible residues PDBQT in addition to the rigid PDBQT file
- Compute AutoGrid maps

Vina 1.1.1

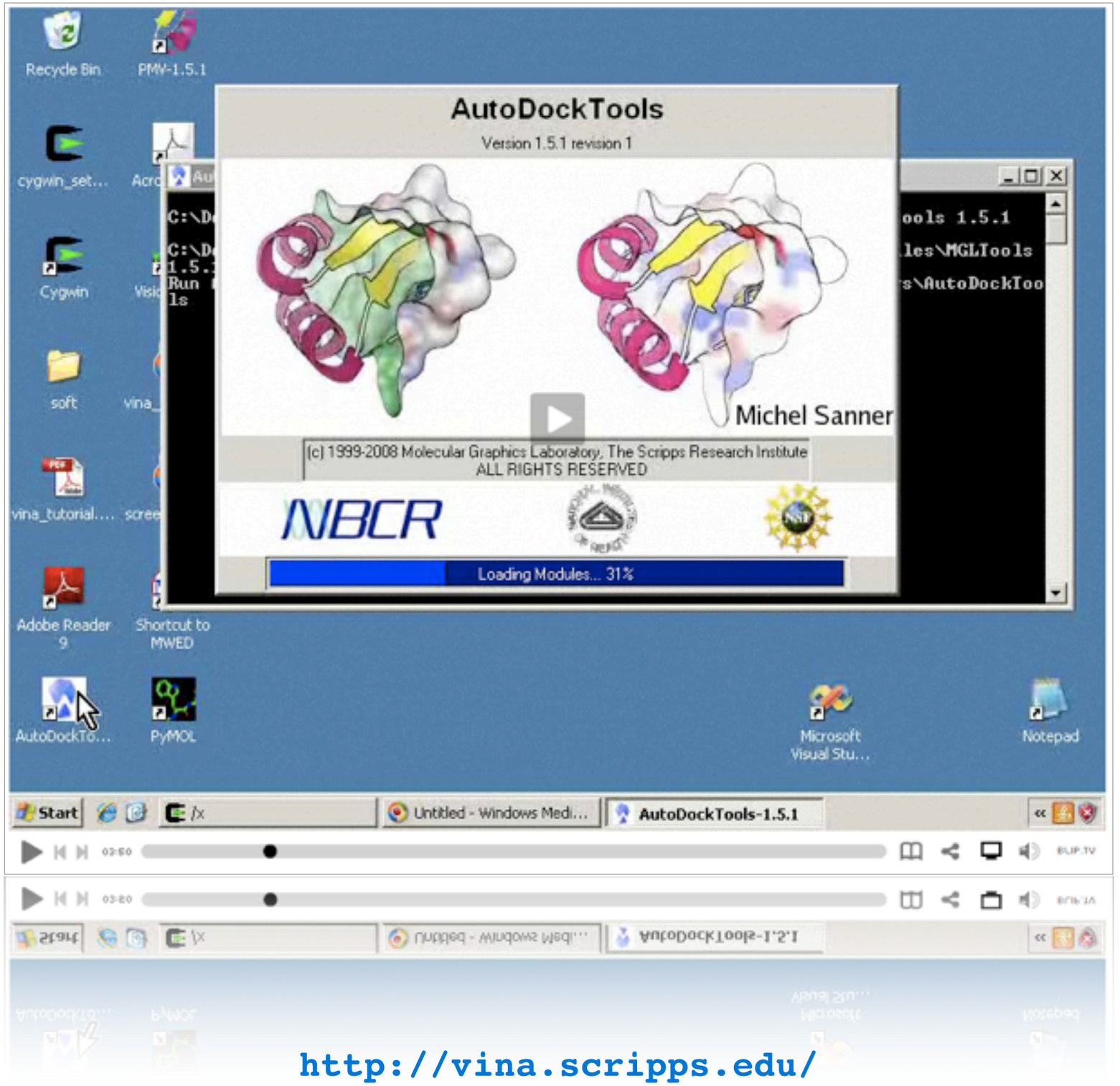
Good that we have AutoDock Tools (ATD)



http://autodock.scripps.edu/resources/adt

Vina 1.1.1

Good we have a nice tutorial



http://vina.scripps.edu/

Acknowledgements

This presentation was based on:

“Using AutoDock 4 with ADT. A tutorial”
by Dr. Ruth Huey and Dr. Garret M. Morris

Vina Tutorial
by Dr. Oleg Trott