# Protein-Ligand Docking Evaluations

Thomas Funkhouser
Princeton University
CS597A, Fall 2005

Introduction

Protein-ligand docking:
- Given a protein and a ligand,
  determine the pose(s) and conformation(s)
  minimizing the total energy of the protein-ligand complex


http://www.molsoft.com/
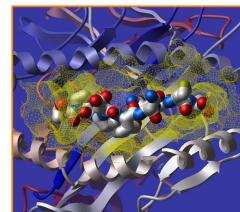
Introduction

Virtual screening:
- Given a protein and a database of ligands,
  use scores (produced by a docking tool) to determine
  which ligands are most likely to bind
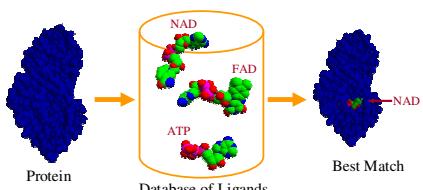
How Good Are Docking Programs?

Questions:
- Docking accuracy?
- Screening accuracy?
- Binding affinity prediction accuracy?
- Computation speed?

How Good Are Docking Programs?

Questions:
- Ø Docking accuracy?
- Screening accuracy?
- Binding affinity prediction accuracy?
- Computation speed?

## [Kellenberger04] Docking Study

8 Docking Programs:
- FRED (multiple conformers)
- DOCK (incremental construction)
- FLEXX (incremental construction)
- SLIDE (incremental construction)
- SURFLEX (incremental construction)
- GLIDE (Monte Carlo simulated annealing)
- QXP (Monte Carlo simulated annealing)
- GOLD (genetic algorithm)

[Kellenberger04] Docking Study
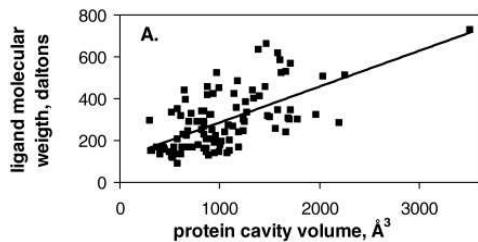
100 Protein-Ligand Complexes:


A.

[Kellenberger04] Docking Study

100 Protein-Ligand Complexes:

B.

[Kellenberger04] Docking Study

100 Protein-Ligand Complexes:

C.
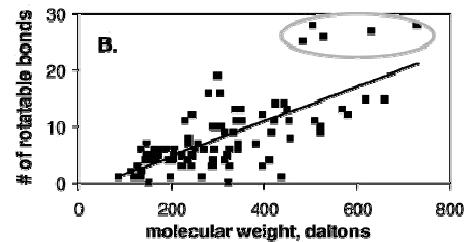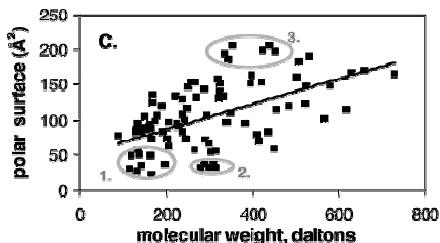
## [Kellenberger04] Docking Study

Docking accuracy:

For Ligand with Closest RMSD (best pose)

A.
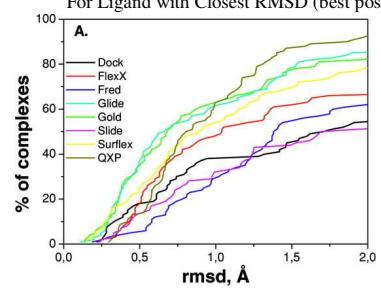
[Kellenberger04] Docking Study

Docking accuracy:

For Ligand Ranked First (top pose)
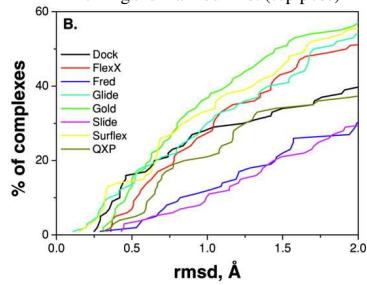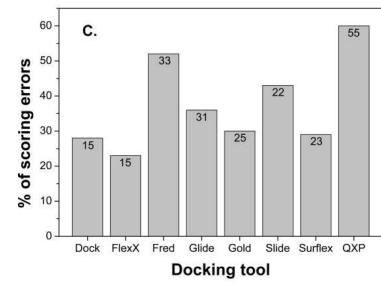
B.

[Kellenberger04] Docking Study

Docking accuracy:

C.

[Kellenberger04] Docking Study

Docking accuracy:

[Kellenberger04] Docking Study

Docking accuracy:

[Kellenberger04] Docking Study

Docking accuracy:

[Kellenberger04] Docking Study

Docking accuracy:

Only 5 most flexible ligands
(>25 rotatable bonds)
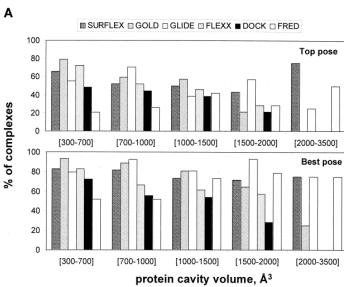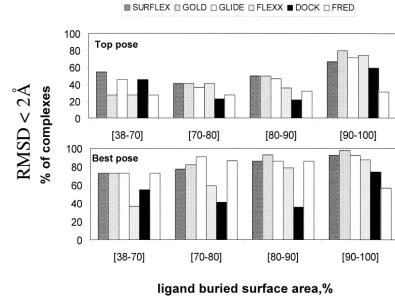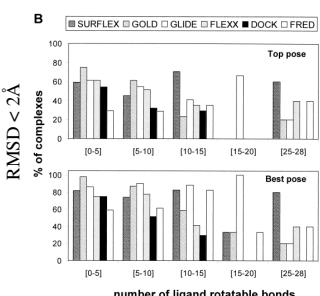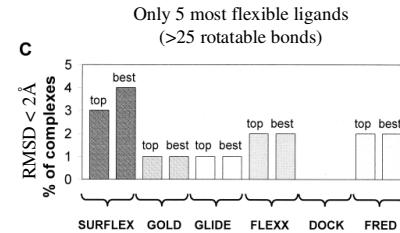
## [Kellenberger04] Docking Study
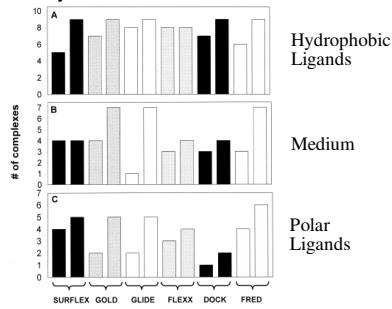
Docking accuracy:

How Good Are Docking Programs?

Questions:
- Docking accuracy?
Ø Screening accuracy?
- Binding affinity prediction accuracy?
- Computation speed?

[Kellenberger04] Screening Study

Dock 1000 ligands into HIV-1 TK
- 10 known TK inhibitors
- 990 randomly chosen "drug-like" molecules

Measure how often TK inhibitors are highly ranked

[Kellenberger04] Screening Study

Screening accuracy:

TABLE I. Description of Hit Lists generated by 8 Docking Tools on the Thymidine Kinase Example
A hit list is generated from the top-scoring compounds selected at a given threshold.

	Top 2.5%		Top 5%		Top 10%	
	Hit Rate[a]	Yield[b]	Hit Rate	Yield	Hit Rate	Yield
DOCK	0	0	2	10	3	30
FLEXX	8	20	8	40	8	80
FRED	0	0	0	0	2	20
GLIDE	8	20	10	50	6	60
GOLD	4	10	8	40	10	100
SLIDE	0	0	0	0	0	0
SURFLEX[c]	16	40	16	80	10	100
QXP	0	0	4	20	2	20

[a]Hit rate: (AH/TH) × 100
[b]Yield: (AH/A) × 100, where TH is the total number of compounds in the hit list, AH the number of true hits in the hit list, and A a total number of true hits in the library.
[c]Figures reported for SURFLEX were obtained by using a protein penetration threshold value of −6.

Precision   Recall

[Kellenberger04] Screening Study
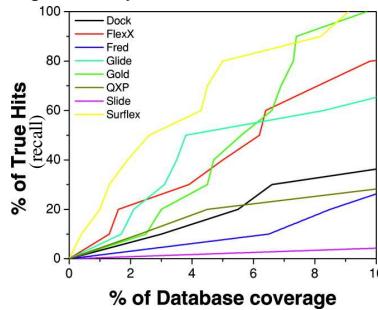
Screening accuracy:

How Good Are Docking Programs?

Questions:
- Docking accuracy?
- Screening accuracy?
- Binding affinity prediction accuracy?
- ØComputation speed?

[Kellenberger04] Docking Study

Computation time:

Program	Average[a]	Minimum[b]	Maximum[c]
FRED[d]	18	0.1	193
DOCK	46	1	667
FLEXX	67	2	595
QXP	108	37	378
SLIDE	118	1	1743
SURFLEX	135	9	1460
GOLD	137	55	479
GLIDE	234	9	2825

How Good Are Docking Programs?

Questions:
- Docking accuracy?
- Screening accuracy?
- ØBinding affinity prediction accuracy?
- Computation speed?

[Marsden04] Study

Compare predicted and measured binding energies

Empirical methods:
- Gold [Jones97]
- DOCK [Kuntz82]
- ChemScore [Eldridge97]

Knowledge-based methods:
- PMF [Muegge99]
- Bleep [Mitchell99]

[Marsden04]

[Marsden04] Study

Figure 7. GOLD calculated log K_d *vs.* Experimental log K_d

$r^2 = 0.20$
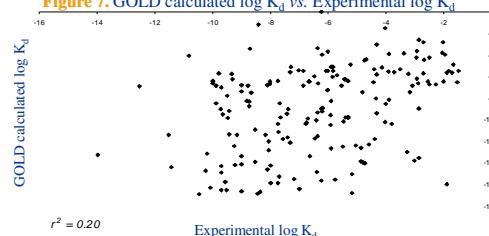
GOLD calculated log K_d

Experimental log K_d

[Marsden04]

[Marsden04] Study

Figure 8. DOCK calculated log K_d *vs.* experimental log K_d

$r^2 = 0.20$
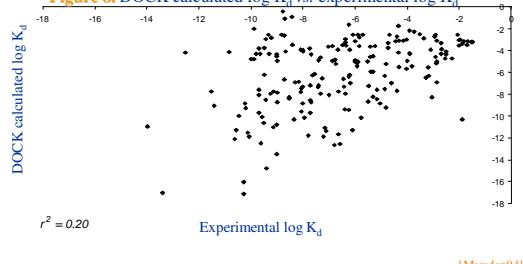
DOCK calculated log K_d

Experimental log K_d

[Marsden04]

[Marsden04] Study

Figure 9. ChemScore calculated log K_d *vs.* experimental log K_d

$r^2 = 0.18$

ChemScore calculated log K_d
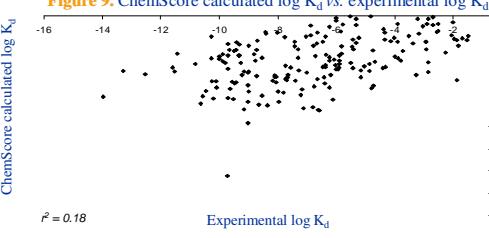
Experimental log K_d

[Marsden04]

[Marsden04] Study

Figure 5. PMF calculated log K_d *vs.* experimental log K_d

$r^2 = 0.11$
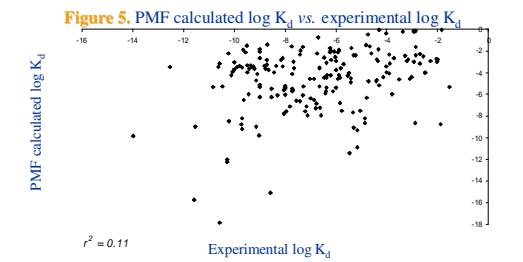
PMF calculated log K_d

Experimental log K_d

[Marsden04]

[Marsden04] Study

Figure 6. BLEEP calculated log K_d *vs.* Experimental log K_d

$r^2 = 0.32$
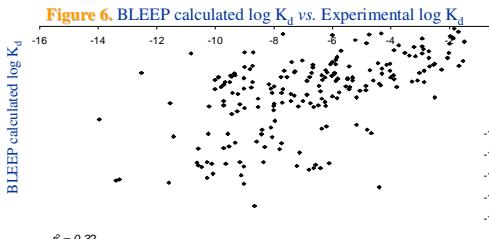
BLEEP calculated log K_d

Experimental log K_d

[Marsden04]

[Marsden04] Study

Table 1. Correlations Between Experimental and Calculated log K_d Values Given by Five Scoring Functions.

	Dataset	No. of complexes	BLEEP		PMF		GOLD		DOCK		ChemScore	
			R_s	r^2	R_s	r^2	R_s	r^2	R_s	r^2	R_s	r^2
	All	205	0.59	0.32	0.31	0.11	0.50	0.20	0.43	0.20	0.45	0.18
A	Serine proteinases	35	0.82	0.74	0.75	0.51	0.61	0.47	0.83	0.69	0.13	0.00
B	Metalloproteinases	25	0.72	0.44	0.45	0.33	0.44	0.14	0.42	0.20	0.43	0.17
C	Carbonic anhydrase ii	18	0.53	0.47	0.46	0.32	0.42	0.34	0.54	0.01	0.50	0.28
D	Sugar binding proteins	30	0.76	0.58	0.45	0.09	0.02	0.00	0.05	0.00	0.29	0.09
E	Aspartic proteinases	38	0.08	0.01	-0.52	0.13	0.02	0.00	0.08	0.01	-0.08	0.00

[Marsden04]

[Marsden04] Study

Conclusion

- All five scoring functions have modest correlation ($r^2 < .32$) with measured K_d values

[Marsden04]

Discussion

?

References

[Eldridge97] M.D. Eldridge, C.W. Murray, T.R. Auton, G.V. Paolini, R.P. Mee, "Empirical scoring functions. I: The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes," J. Comput-Aided Mol. Des., 11, 1997, pp. 425-445.

[Friesner04] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, "Glide: A New Approach for Rapid, Accurate Docking and Scoring.," J. Med. Chem, 47, 2004, pp. 1739-1749.

[Gohlke00] H. Gohlke, M. Hendlich, G. Klebe, "Knowledge-based scoring function to predict protein-ligand interactions," J. Mol. Biol., 295, 2000, pp. 337-356.

[Jones97] G. Jones, P. Willett, R.C. Glen, A.R. Leach, R. Taylor, "Development and Validation of a Genetic Algorithm for Flexible Docking," J. Mol. Biol., 267, 1997, pp. 727-748.

[Kellenberger04] E. Kellenberger, J. Rodrigo, P. Muller, D. Rognan, "Comparative evaluation of eight docking tools for docking and virtual screening accuracy," Proteins, 57, 2, 2004, pp. 225-242.

[Kuntz82] I.D. Kuntz, J.M. Blaney, S.J. Oatley, R. Langridge, T.E. Ferrin, "A geometric approach to macromolecule-ligand interactions," J. Mol. Biol, 161, 1982, pp. 269-288.

[Marsden04] Marsden PM, Puvanendrampillai D, Mitchell JBO and Glen RC., "Predicting protein ligand binding affinities: a low scoring game?", Organic Biomolecular Chemistry, 2, 2004, p. 3267-3273.

[Mitchell99] J.B.O. Mitchell, R. Laskowski, A. Alex, and J.M. Thornton, "BLEEP - potential of mean force describing protein-ligand interactions: II. Calculation of binding energies and comparison with experimental data", J. Comput. Chem., 20, 11, 1999, pp. 1177-1185.

[Morris98] Morris, G. M., Goodsell, D. S., Halliday, R.S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J. "Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function", J. Computational Chemistry, 19, 1998, p. 1639-1662.

[Muegge99] I. Muegge, Y.C. Martin, "A general and fast scoring function for protein-ligand interactions: A simplified potential approach," J. Med. Chem., 42, 1999, pp. 791-804.

[Rarey96] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, "A Fast Flexible Docking Method using an Incremental Construction Algorithm," Journal of Molecular Biology, 261, 3, 1996, pp. 470-489.