Is automatic docking feasible?

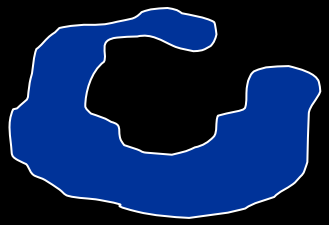
John J. Irwin
UCSF Pharmaceutical Chemistry

Acknowledgements

- Shoichet Lab
- Michael Mysinger
- Niu Huang
- Francesco Colizzi
- Eddie Cao

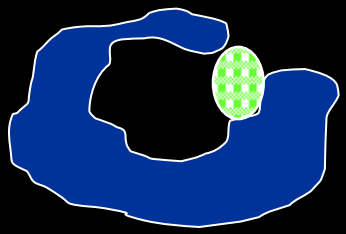
- NIH for funding

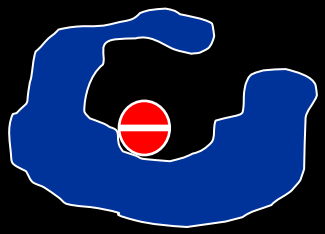
Screening for Novel Inhibitors by Molecular Docking

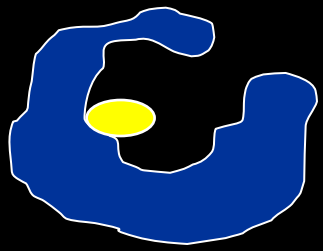


DOCK Blaster

dock

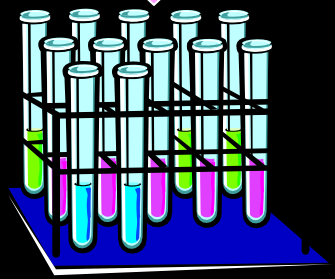






DUD

Test high-scoring molecules



Why is docking difficult?

Binding sites are complicated
Lots of interactions to consider
Everything in competition with water

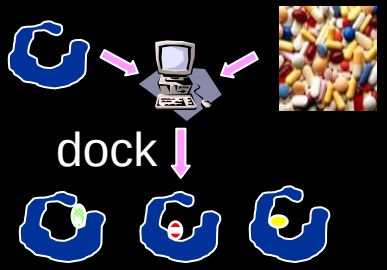
$$K_d = e^{-\Delta G/RT}$$

$$\Delta G_{bind} = \Delta G_{inter} - \Delta G_{solv}$$

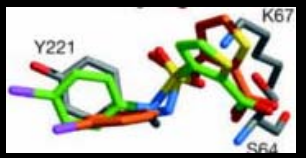
small large large

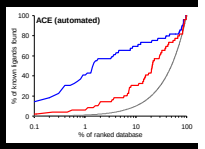
Why is docking difficult to automate?

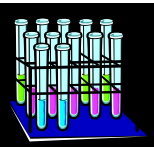
Database preparation



Running docking:
Site preparation
Software configuration
Parameter choices
File manipulations



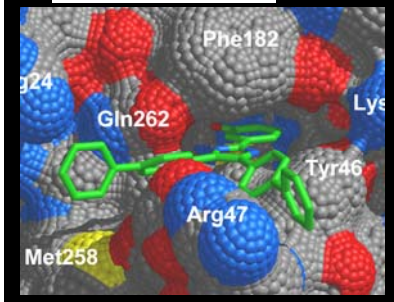




DUD

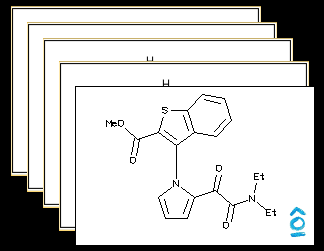
Assessment of results
Using controls when available

chEMBL



Structure

Interpretation of structure



The ZINC Database

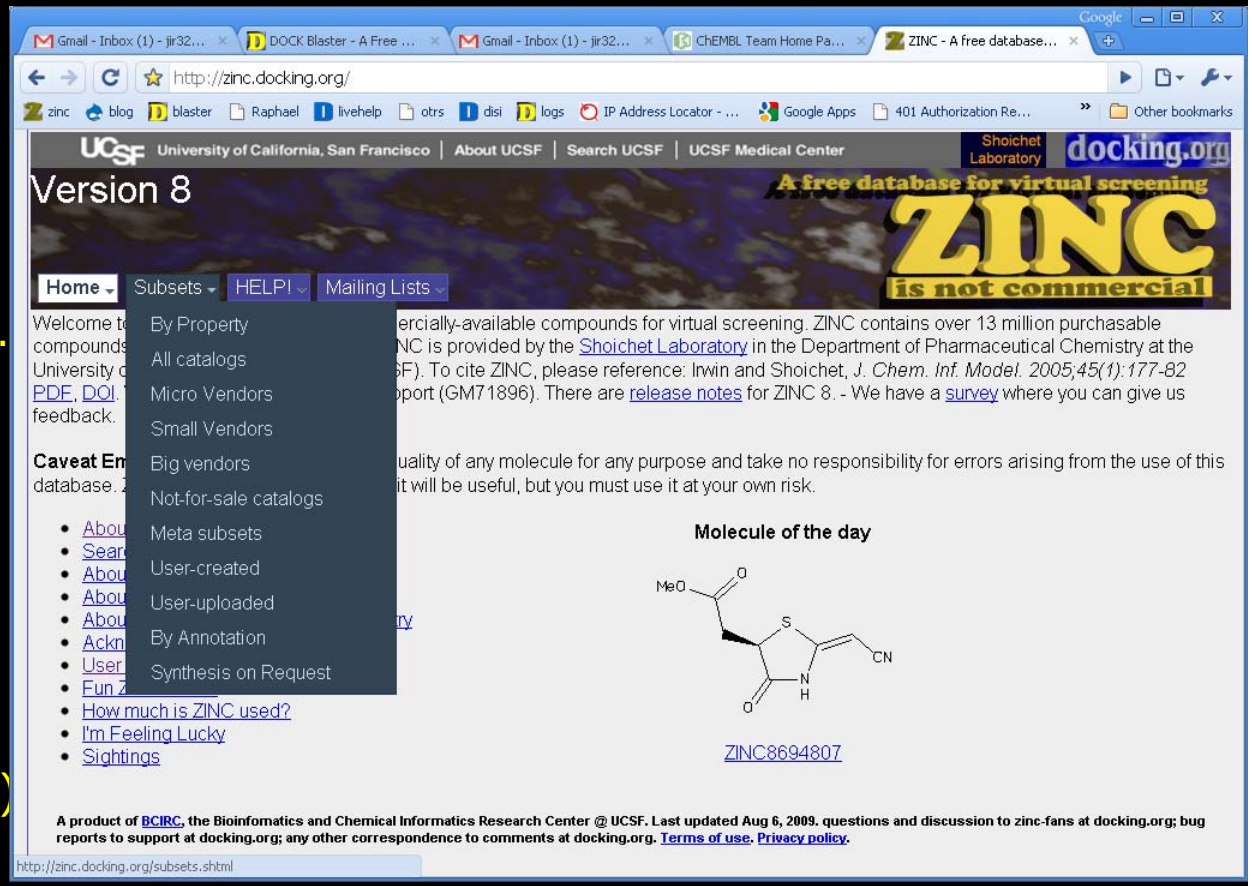
http://zinc.docking.org

21 million compounds
commercially available
structures calculated
multiple conformations
properties (charge, solv, etc…
links to suppliers

Free to the community

Multiple subsets
8.8 M drug-like (Lipinski)
3.4 M lead-like (Oprea…)
450 K fragment-like (Astex, …)

Availlable in popular formats
SMILES, SDF, mol2, flexibase



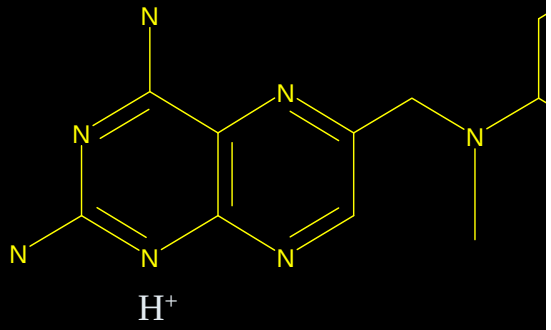
Updated continuously (10,000 new today)
Over 2 million new compounds per year
Over 1 million depletions per year

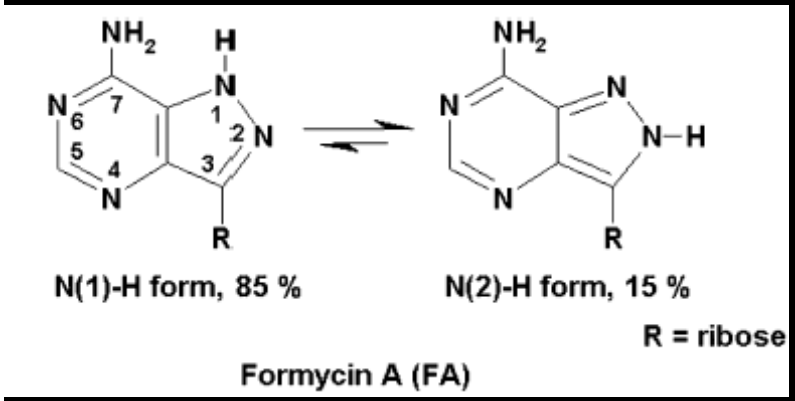
Irwin & Shoichet *JCIM* **2005**

Compound Vendors in ZINC

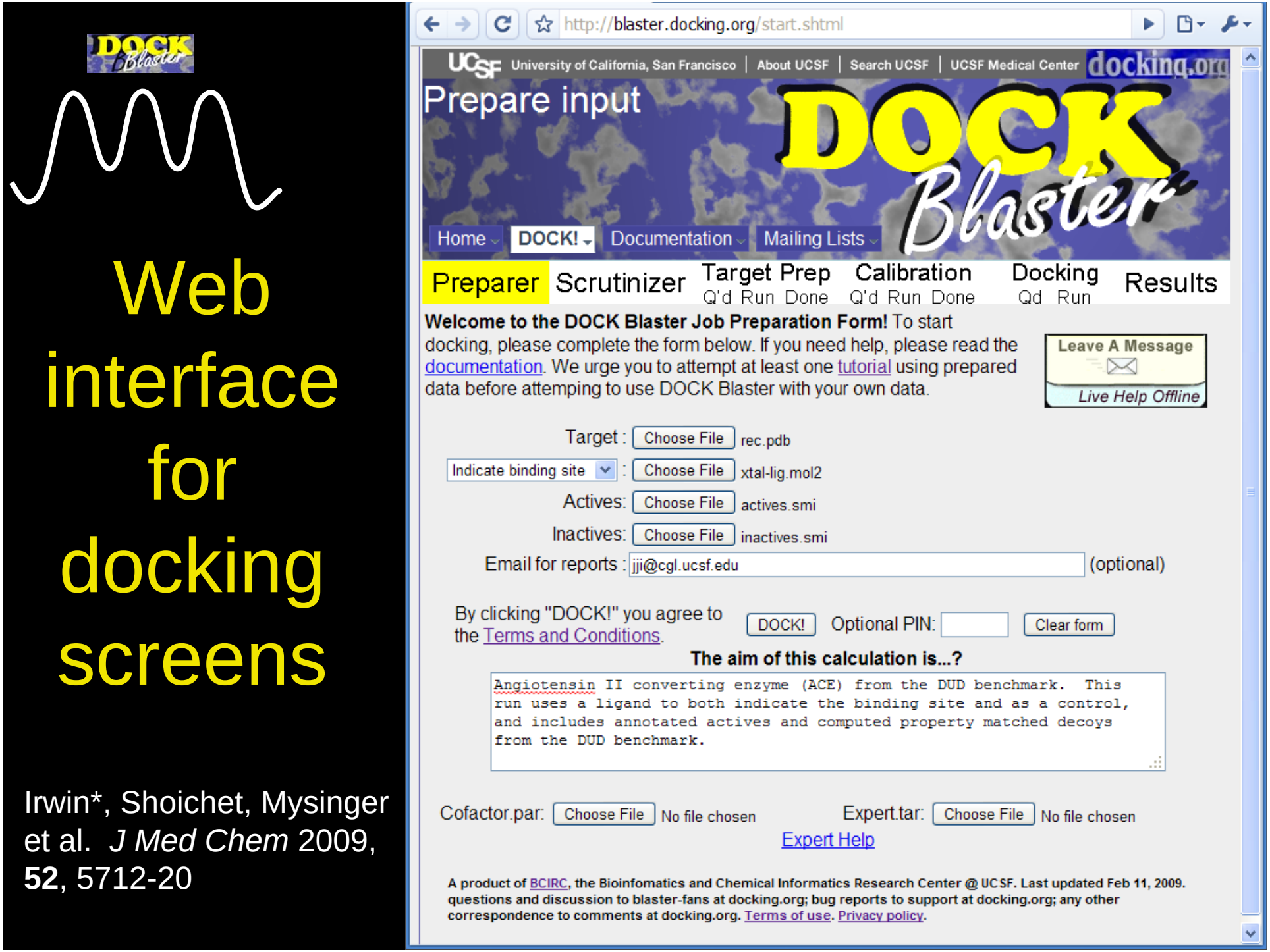
Some Database Preparation Pitfalls

- Nuisance compounds (filter or annotate)
- Stereochemistry
- Protonation / Charge
- Tautomers



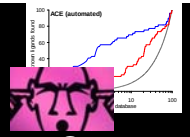


Formycin A, an adenosine bioisostere, augments insulin release



Web interface for docking screens

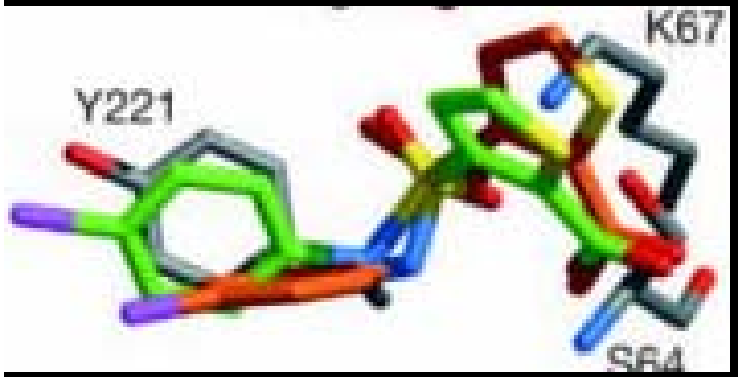
Irwin*, Shoichet, Mysinger et al. *J Med Chem* 2009, **52**, 5712-20

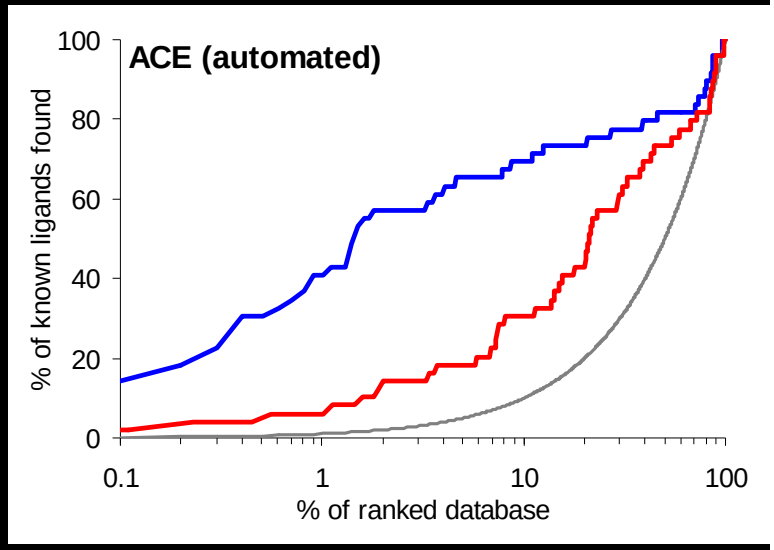


Is docking working?

Pose-fidelity

Enrichment / Rank





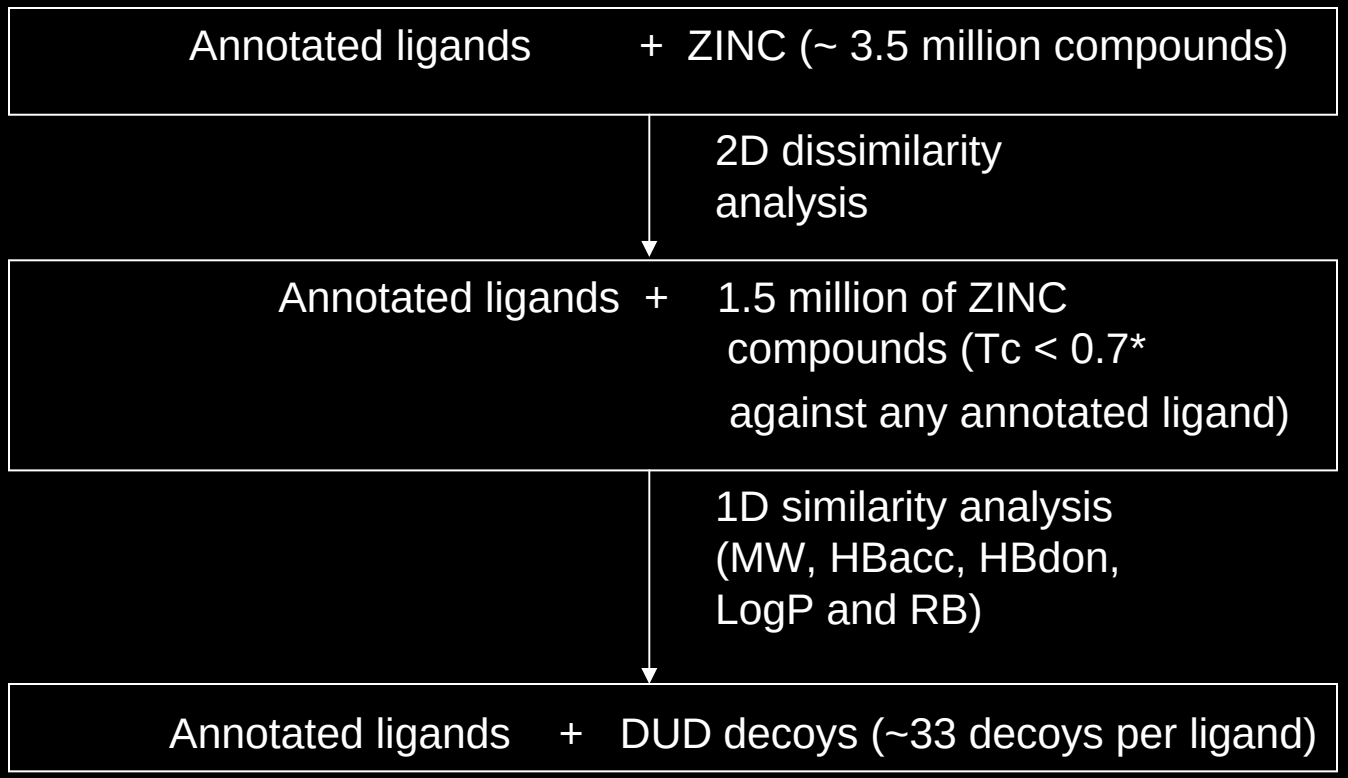
Single ligand metric and multi-ligand metric

Measured by RMSD (A)

Measured by rank vs physically matched decoys

DUD – a Directory of Useful Decoys

How it was Assembled



Huang, Shoichet*, Irwin*, *J. Med. Chem.* ***49***, 6789 – 6801 (2006)

Benchmarking Virtual Screening with DUD

Protein	Number of ligands	Protein	Number of ligands	Protein	Number of ligands	Protein	Number of ligands
$ER_{antagonist}$	40	FGFr1	118	Thrombin	65	PARP	33
$ER_{agonist}$	67	SRC	162	COMT	12	ALR2	26
AR	74	P38 MAP	234	ADA	23	PNP	25
RXRa	20	PDGFrb	156	ACE	49	SAHH	33
PPARg	81	VEGFr2	74	PDE5	50	HIVRT	39
MR	15	CDK2	50	GART	21	AChE	105
GR	78	TK	22	DHFR	201	InhA	85
PR	27	Trypsin	43	AmpC	21	HMGR	35
HSP90	23	fXa	142	GPB	52	COX-1	25
EGFr	416	HIVPR	53	NeuA	49	COX-2	349

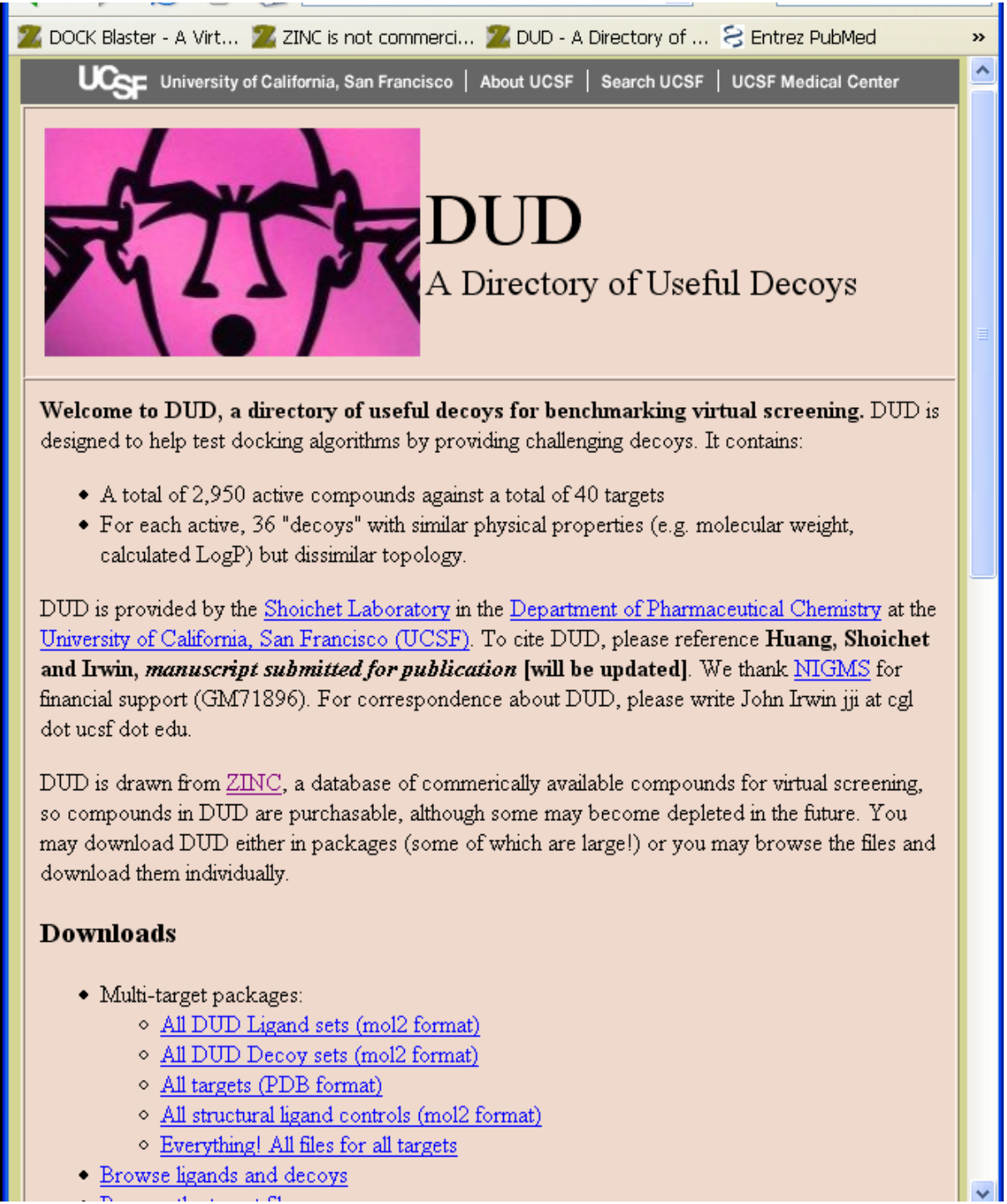
Huang, Shoichet*, Irwin*, *J. Med. Chem.* ***49***, 6789 – 6801 (2006)

DUD is free

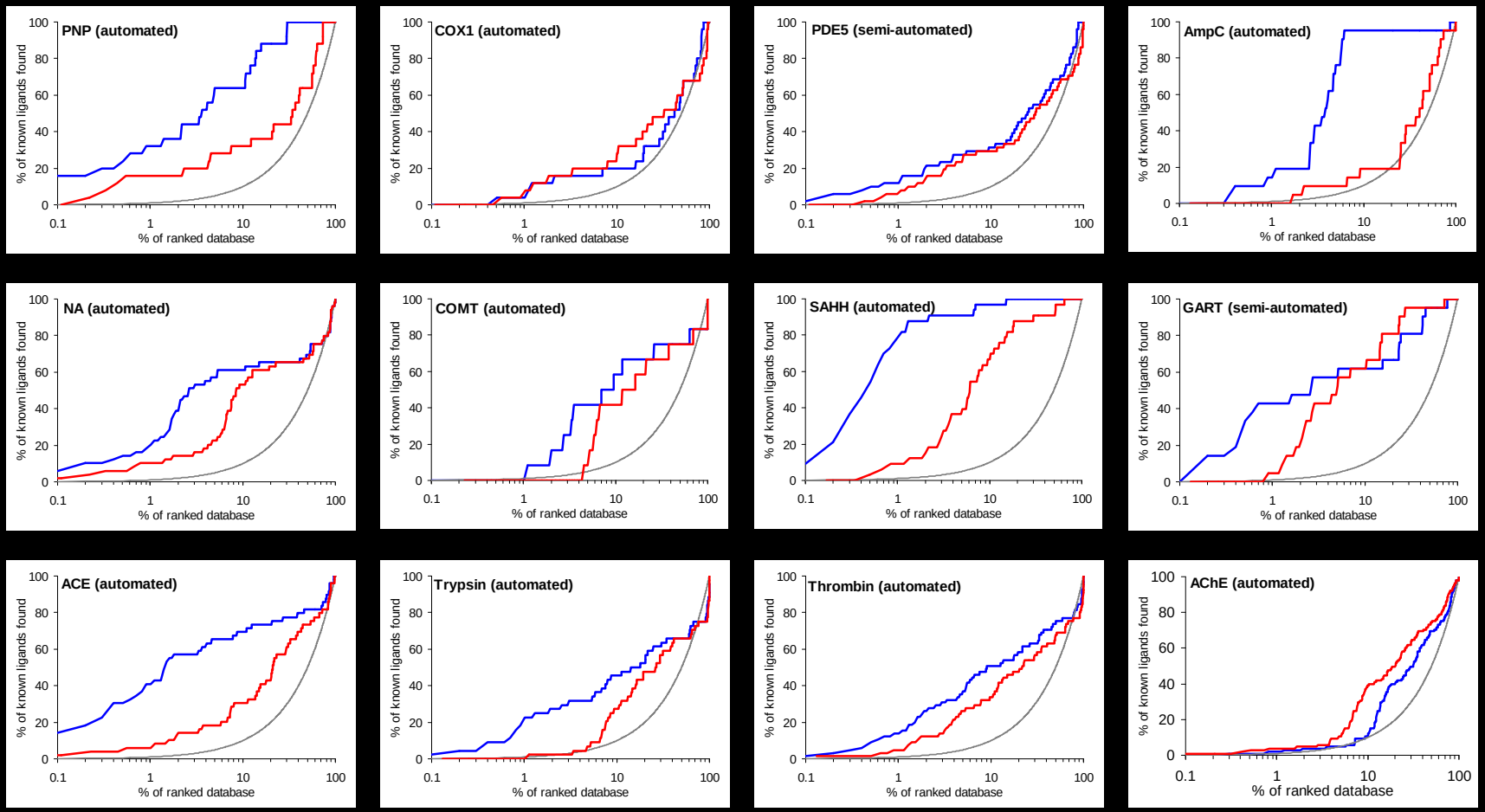
40 targets
2,950 ligands
95,358 decoys

mol2 format
All docking files

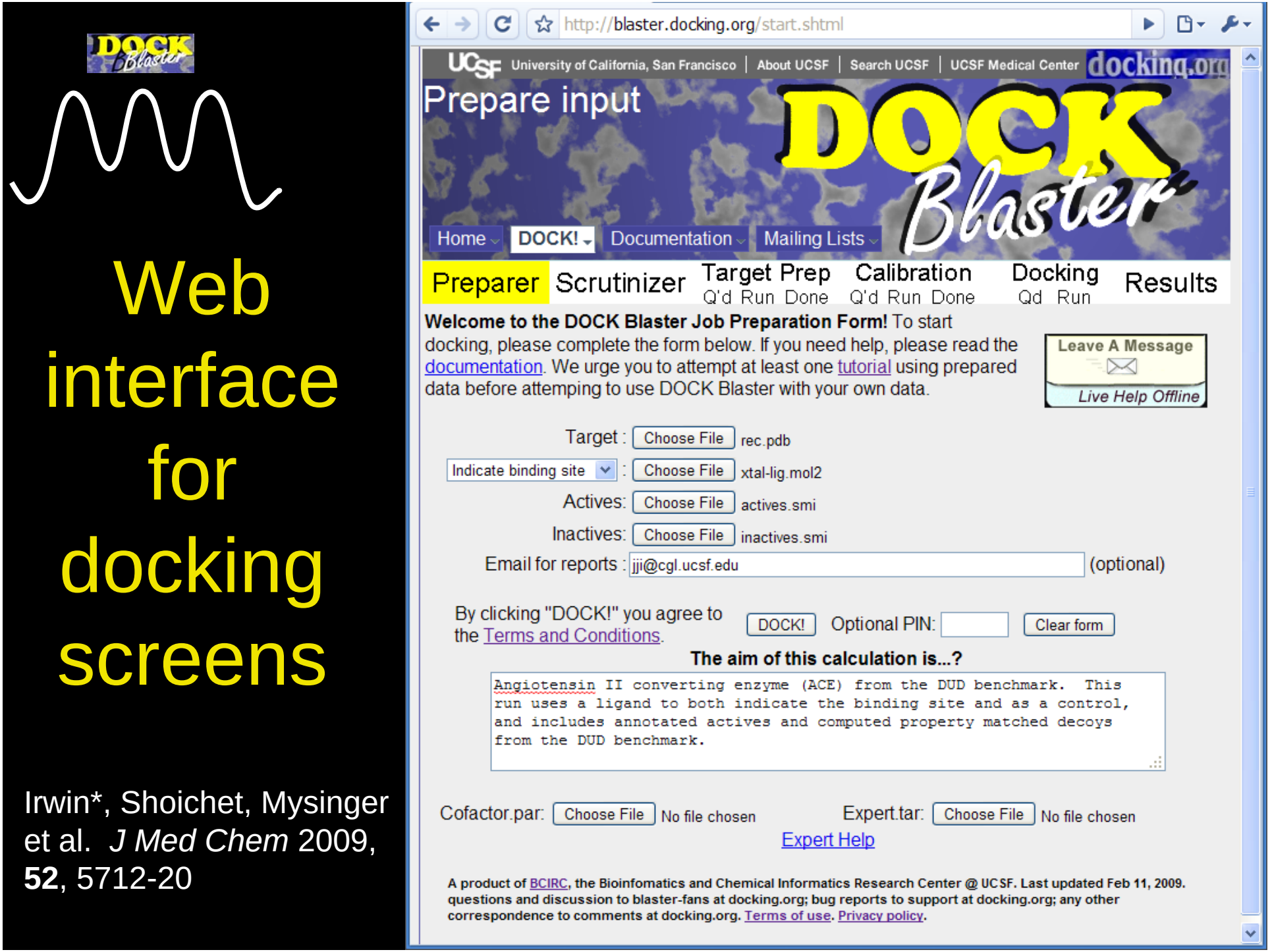
dud.docking.org



“Own decoys” are most challenging



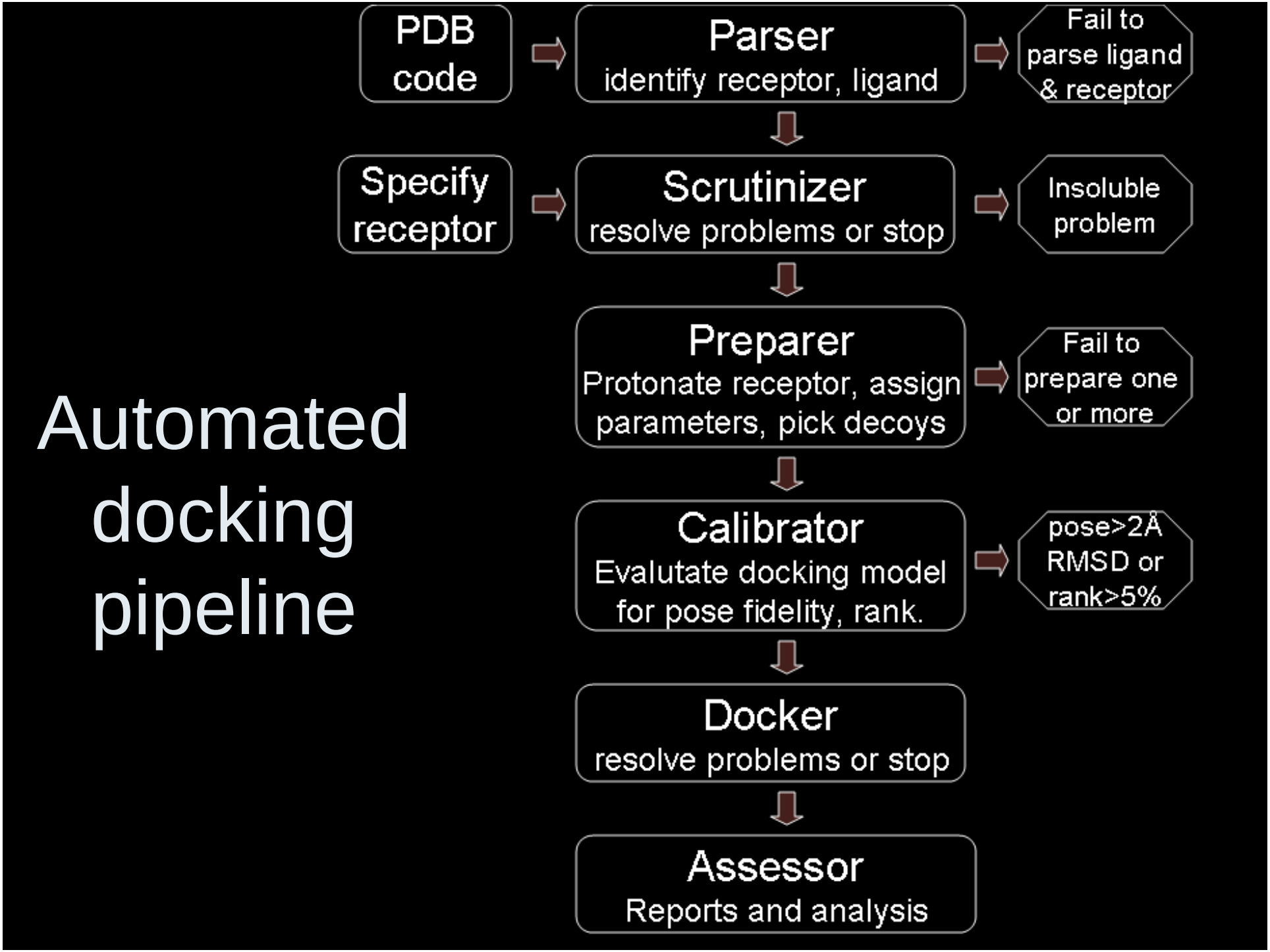
Blue = all DUD, Red = own decoys, grey = random



Web interface for docking screens

Irwin*, Shoichet, Mysinger et al. *J Med Chem* 2009, **52**, 5712-20

Automated docking pipeline



Try docking four ways

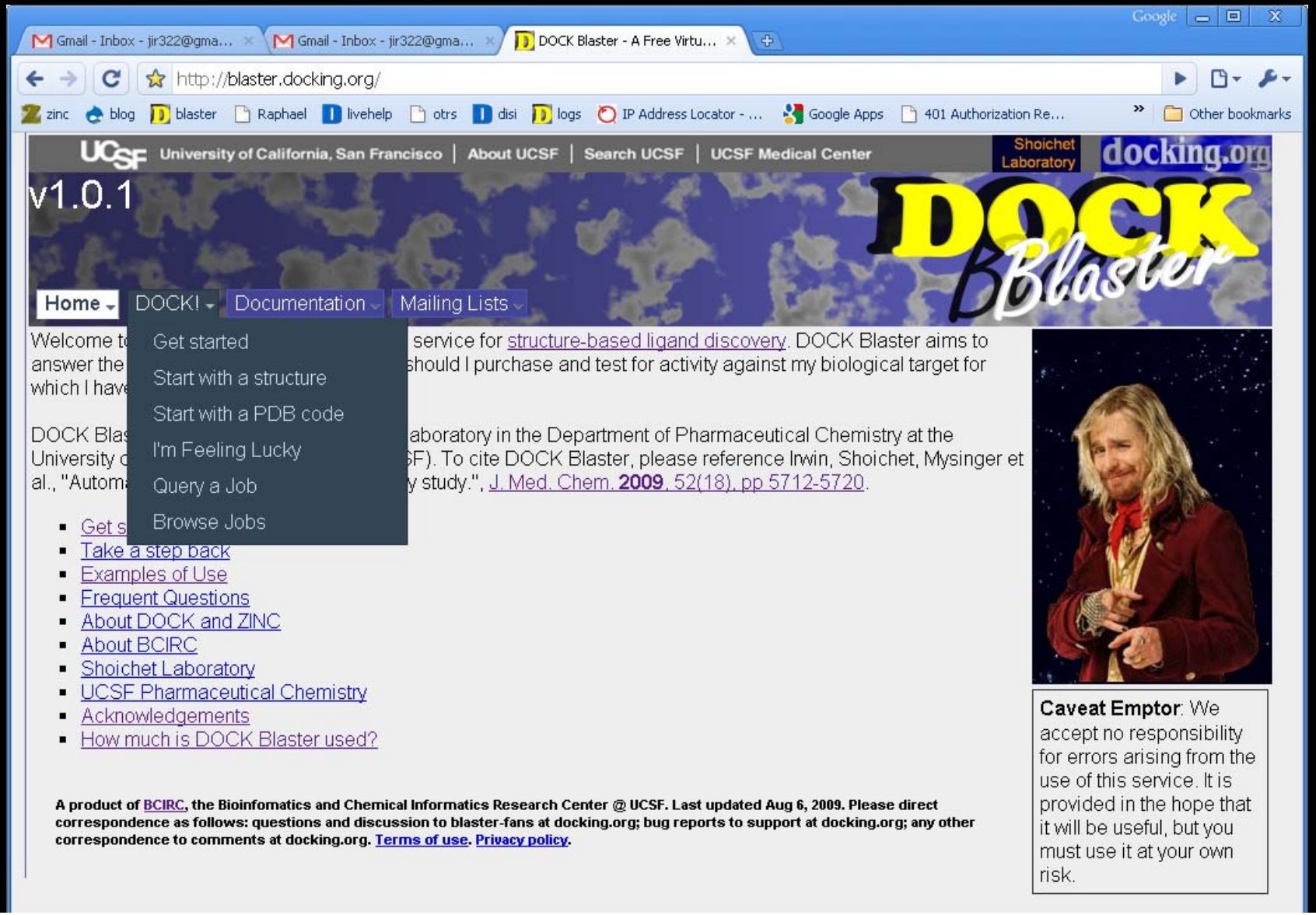
Scoring

Sampling

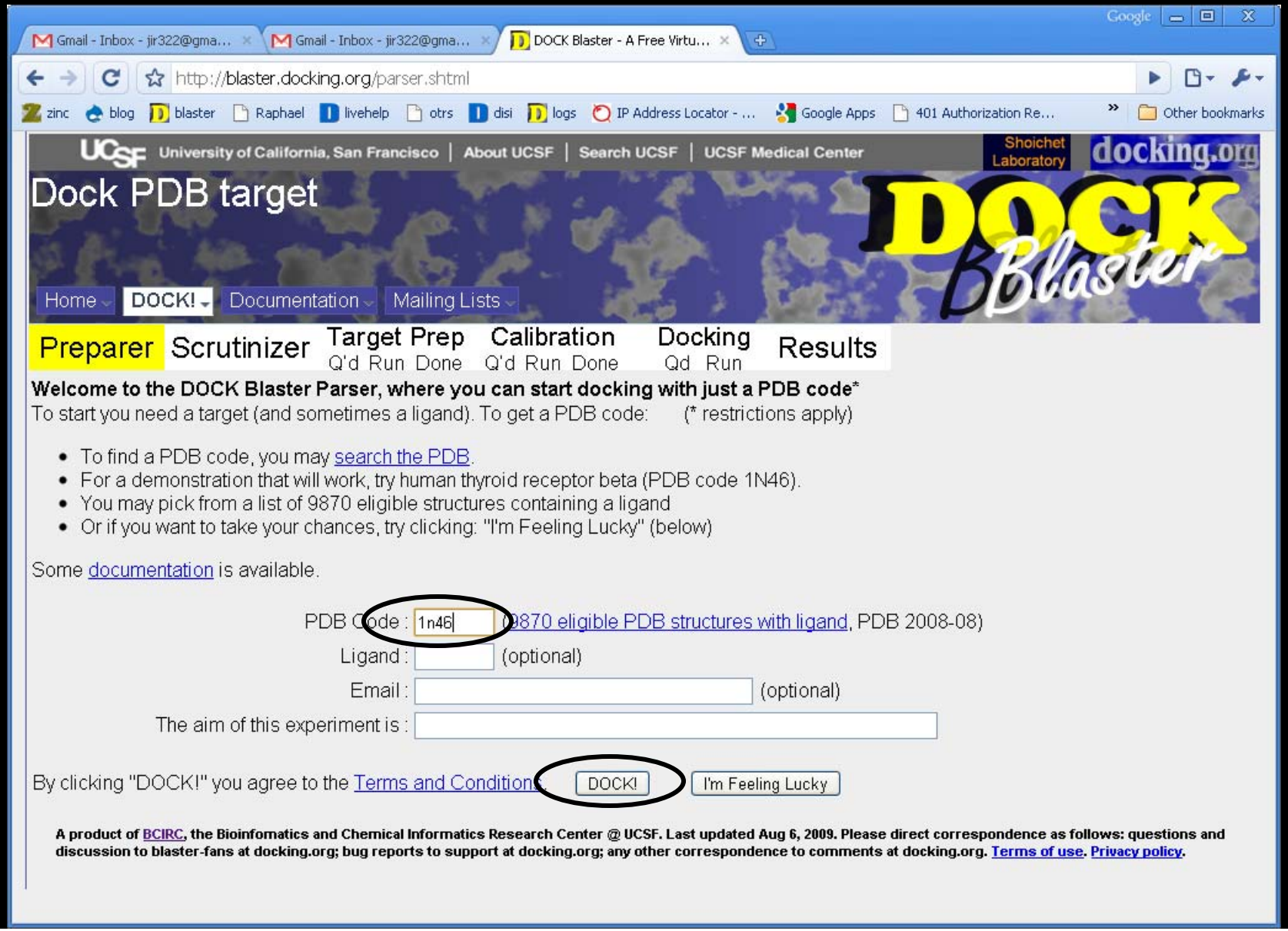
	Polarized	AMBER
Coarse	#1	#2
Fine	#3	#4

Thus four docking runs with four different parameters

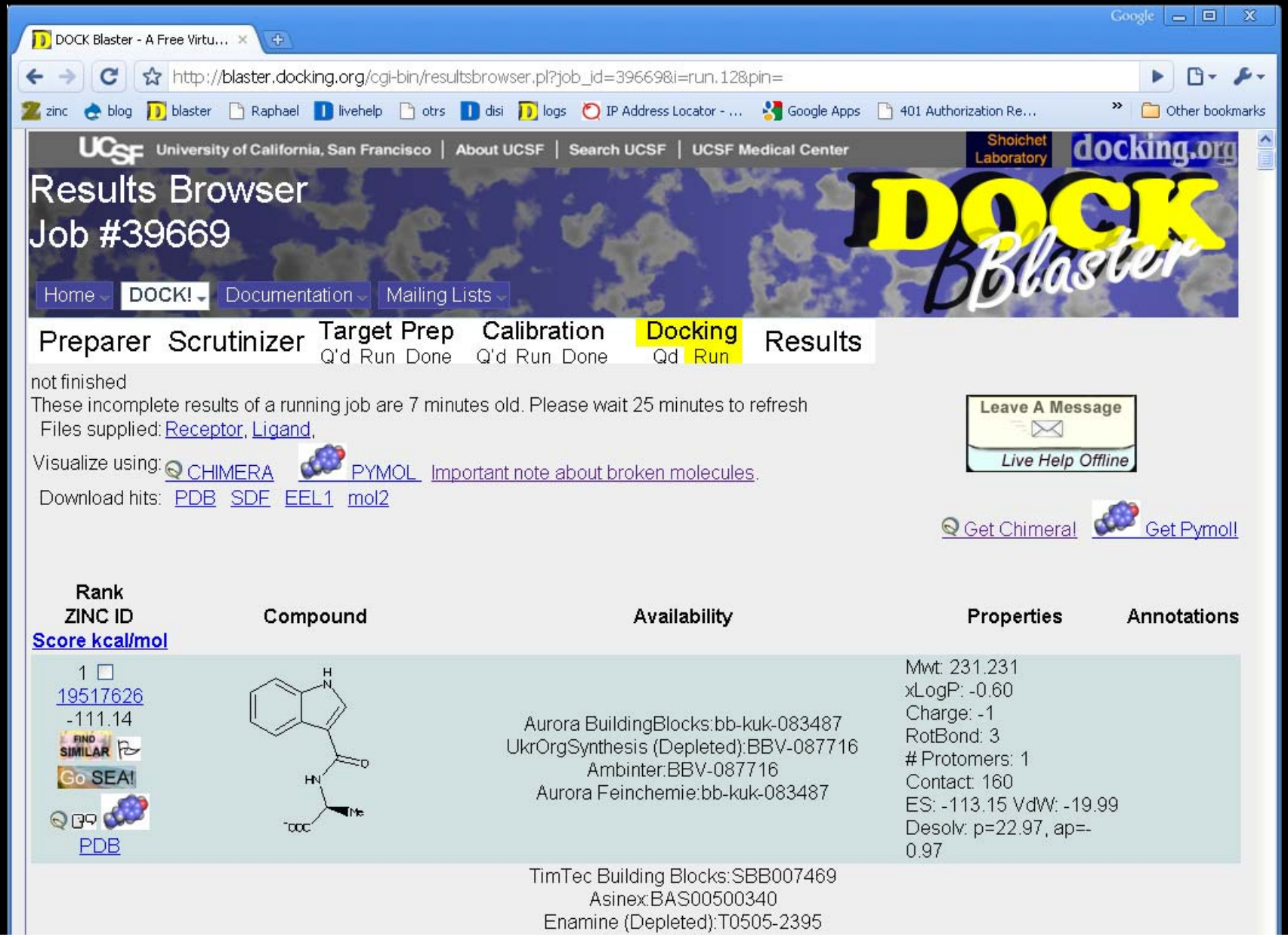
Starting point: http://blaster.docking.org
Choose: Start with a PDB code



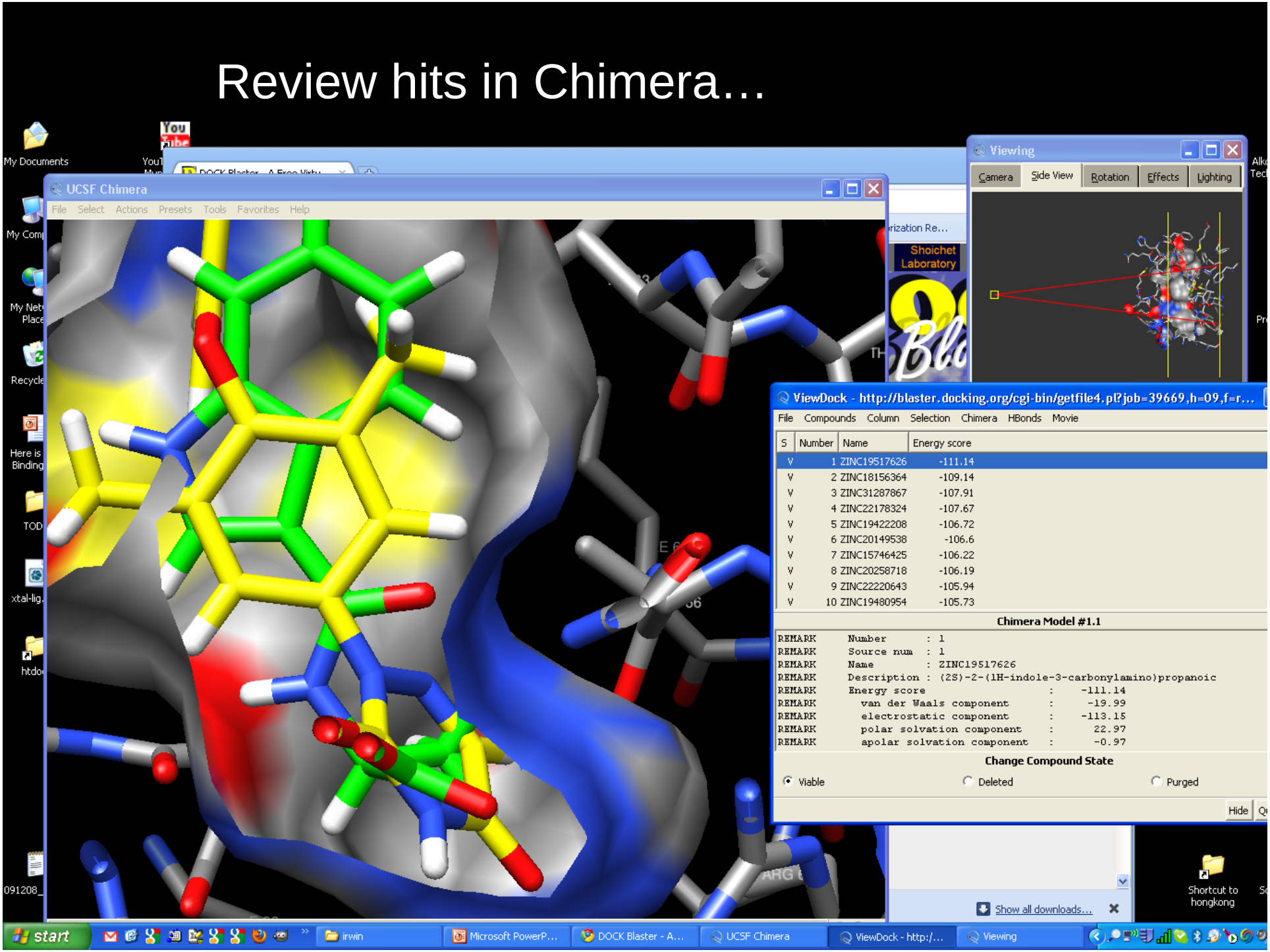
Pick a PDB Code for docking. Click DOCK!



Review docking hits. Click on “CHIMERA”

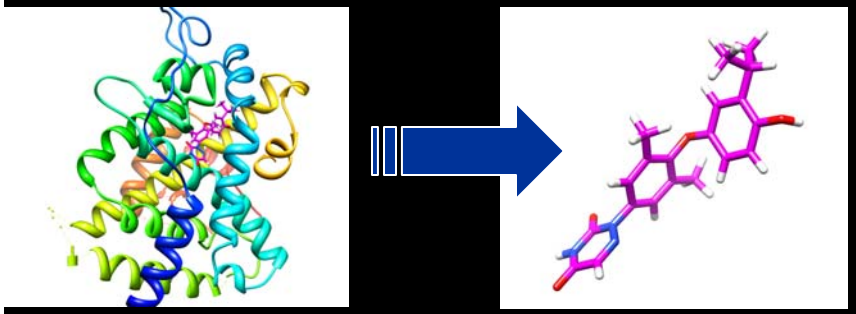


Review hits in Chimera…

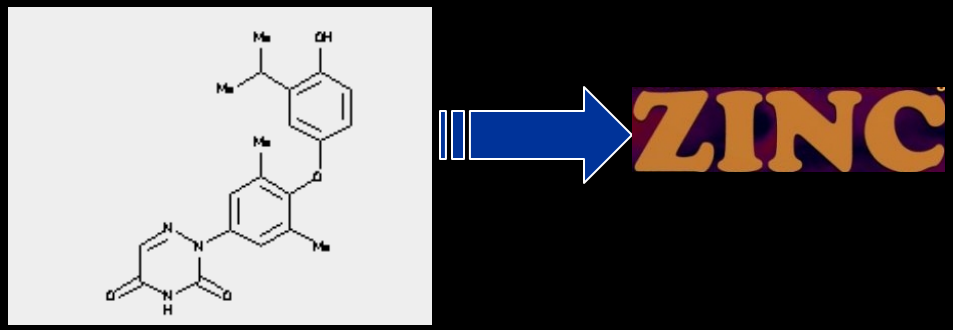


#1. Self-assessment

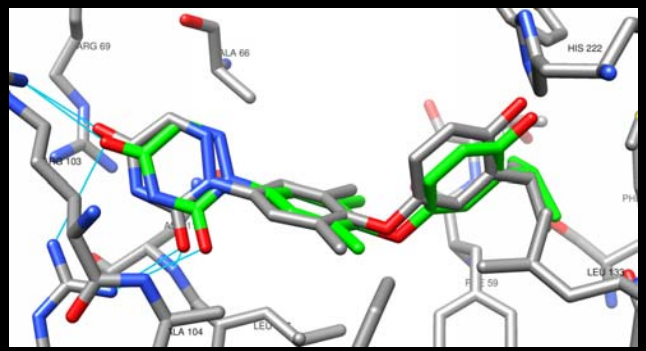
1. Remove ligand from receptor



2. Rebuild ligand without bias



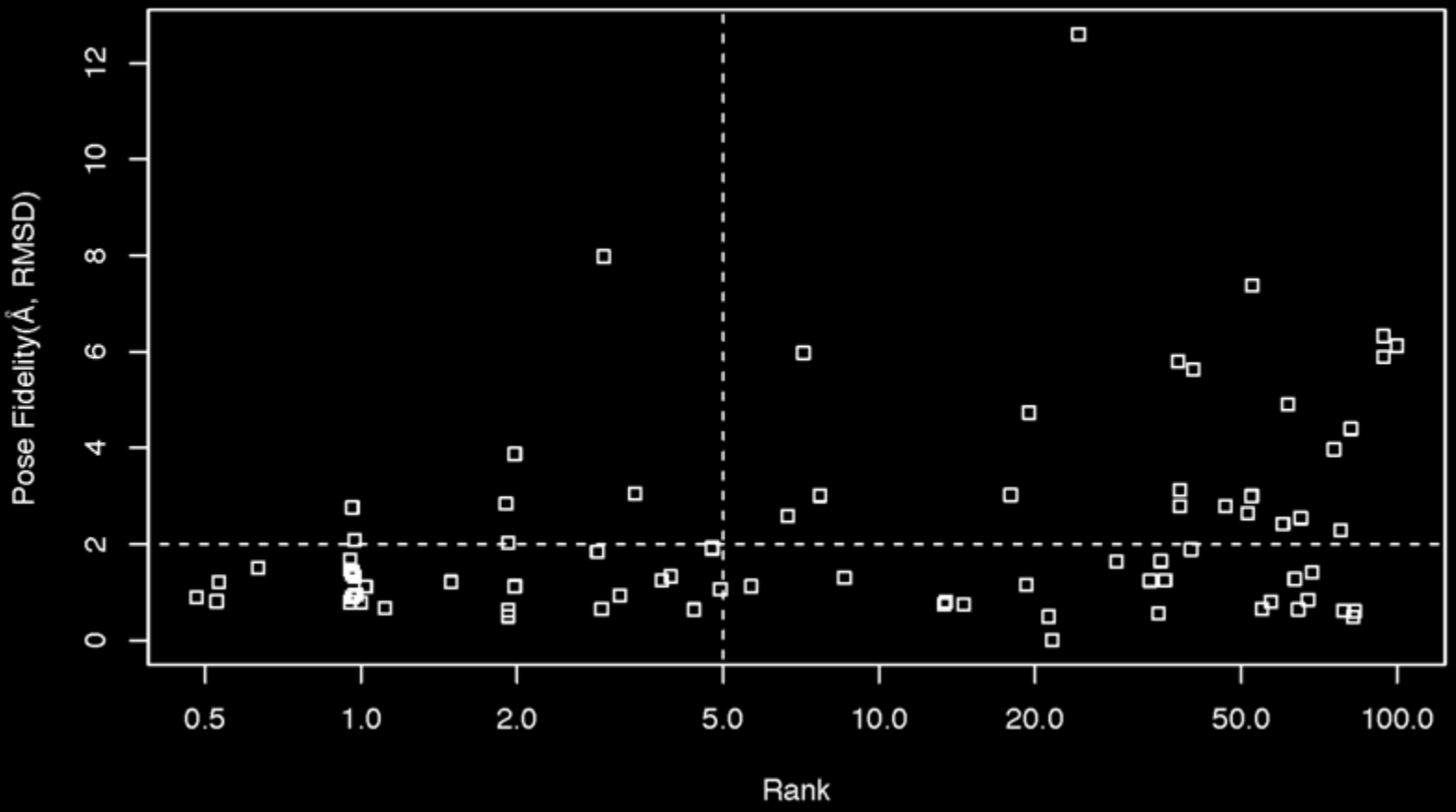
3. Dock ligand and 100 phyisically-matched decoys using 4 parameter sets



4. Evaluate pose-fidelity, enrichment

		Scoring	
Sampling		*Polarized*	*Normal*
	Coarser	3.61 Å / 1%	1.32 Å / 9%
	Finer	1.32 Å / 2 %	2.02 Å / 3%

#2. Pose fidelity does not predict enrichment



Astex-85 benchmark

Experiment: re-dock crystallographic ligand and 100 property-matched decoys

#3. Large benchmark

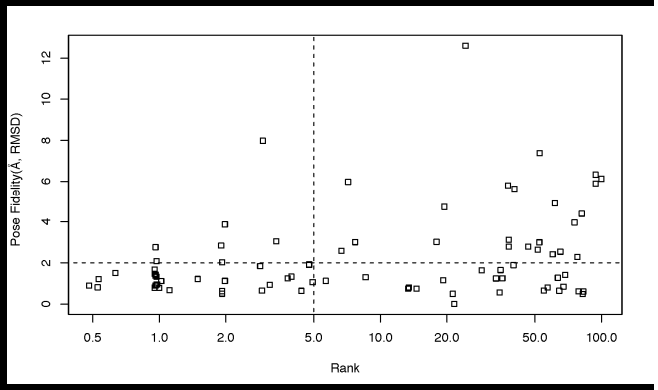
Description	**Astex-85**	**Gold-114**	**DUD-38**	**PDB-9050**
Ligand docked / scored	82	94	36	7,750
Good pose achieved	51	58	23	3,020
Good pose and rank	29	27	15	1,398

Fully automatic docking starting from PDB code (and ligand specification as required)

Four ideas

1. Automatic self-assessment
2. Pose fidelity does not predict enrichment
3. Large benchmark
4. Public service

		Scoring	
		Polarized	*Normal*
Sampling	*Coarser*	3.61 Å / 1%	1.32 Å / 9%
	Finer	1.32 Å / 2 %	2.02 Å / 3%



DOCK Blaster is free to use

http://blaster.docking.org/