

# Is automatic docking feasible?

John J. Irwin

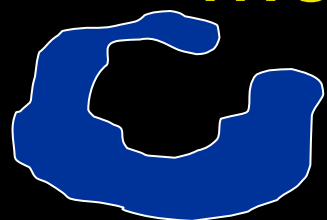
UCSF Pharmaceutical Chemistry

# Acknowledgements

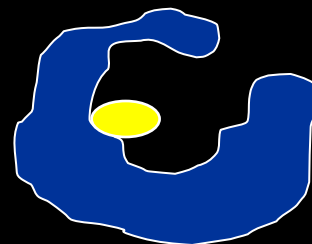
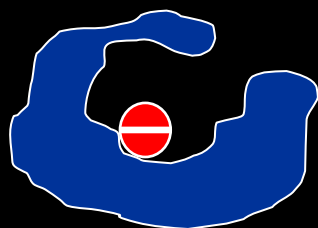
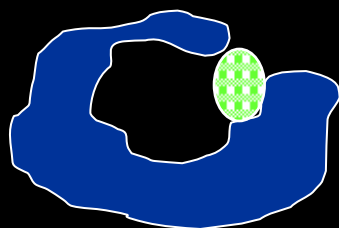
- Shoichet Lab
  - Michael Mysinger
  - Niu Huang
  - Francesco Colizzi
  - Eddie Cao
- 
- NIH for funding



# Screening for Novel Inhibitors by Molecular Docking

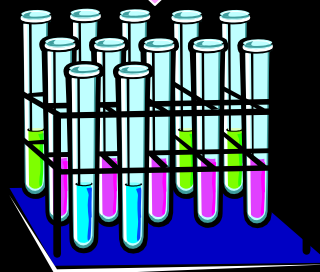


dock



DUD

Test high-scoring  
molecules



# Why is docking difficult?

Binding sites are complicated

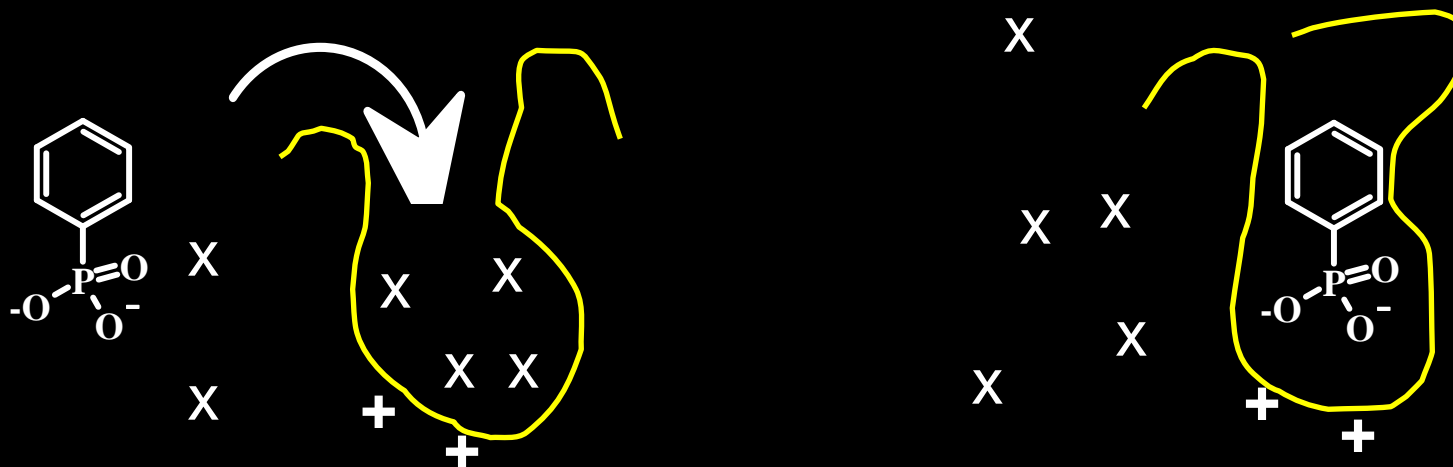
Lots of interactions to consider

Everything in competition with water

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

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

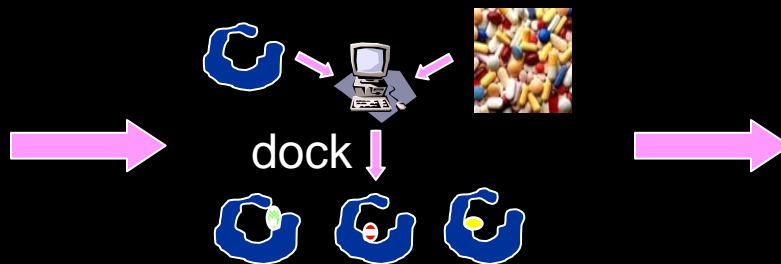
small                      large                      large



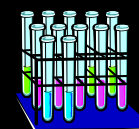
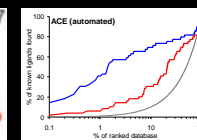
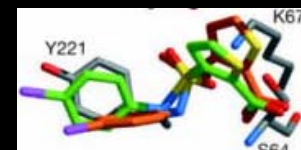
# Why is docking difficult to automate?



Database  
preparation



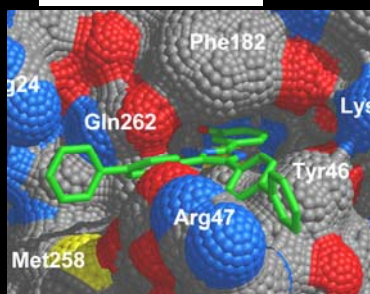
Running docking:  
Site preparation  
Software configuration  
Parameter choices  
File manipulations



DUD

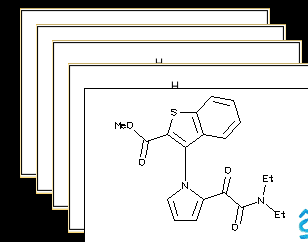
Assessment of results  
Using controls when  
available

chEMBL

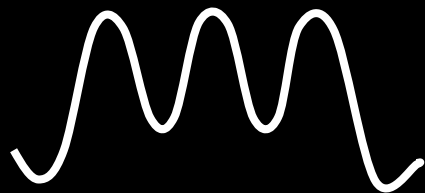


Structure

Interpretation  
of structure



**ZINC**



# 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, ...)

Available in popular formats  
SMILES, SDF, mol2, flexibase

Irwin & Shoichet *JCIM* 2005

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



# Compound Vendors in ZINC



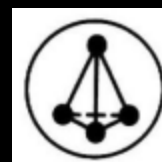
NANOSYN



PHARMEKS



Key Organics

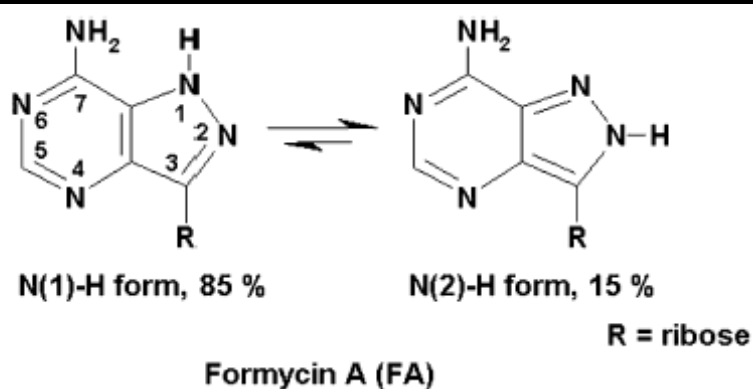
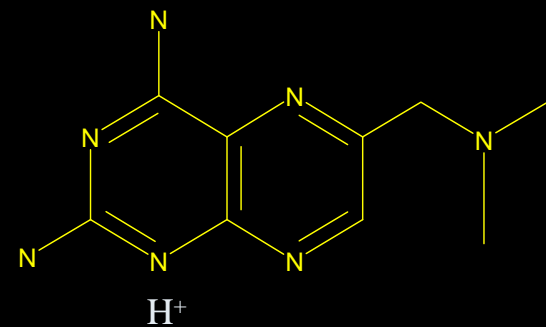


COMBI-BLOCKS Inc.



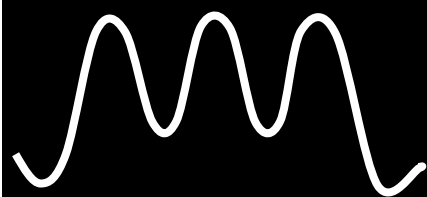
# Some Database Preparation Pitfalls

- 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

← → ↻ ☆ <http://blaster.docking.org/start.shtml> ▶ 📁 🔧

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center [docking.org](http://docking.org)

## Prepare input

# DOCK Blaster

Home ▾ DOCK! ▾ Documentation ▾ Mailing Lists ▾

**Preparer** Scrutinizer Target Prep Calibration Docking Results  
Q'd Run Done Q'd Run Done Q'd Run

Welcome to the DOCK Blaster Job Preparation Form! To start docking, please complete the form below. If you need help, please read the [documentation](#). We urge you to attempt at least one [tutorial](#) using prepared data before attempting to use DOCK Blaster with your own data.

[Leave A Message](#)  
✉  
[Live Help Offline](#)

Target:  rec.pdb  
Indicate binding site: ▾  xtal-lig.mol2  
Actives:  actives.smi  
Inactives:  inactives.smi  
Email for reports:  (optional)

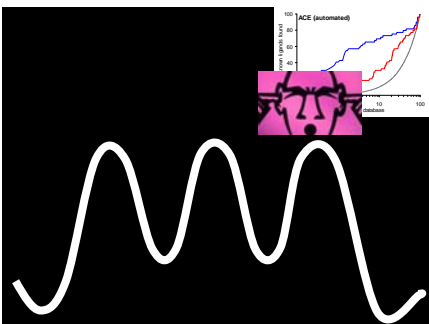
By clicking "DOCK!" you agree to the [Terms and Conditions](#).  Optional PIN:

**The aim of this calculation is...?**

Angiotensin II converting enzyme (ACE) from the DUD benchmark. This run uses a ligand to both indicate the binding site and as a control, and includes annotated actives and computed property matched decoys from the DUD benchmark.

Cofactor.par:  No file chosen Expert.tar:  No file chosen  
[Expert Help](#)

A product of [BCIRC](#), the Bioinformatics and Chemical Informatics Research Center @ UCSF. Last updated Feb 11, 2009. questions and discussion to blaster-fans at docking.org; bug reports to support at docking.org; any other correspondence to comments at docking.org. [Terms of use](#). [Privacy policy](#).

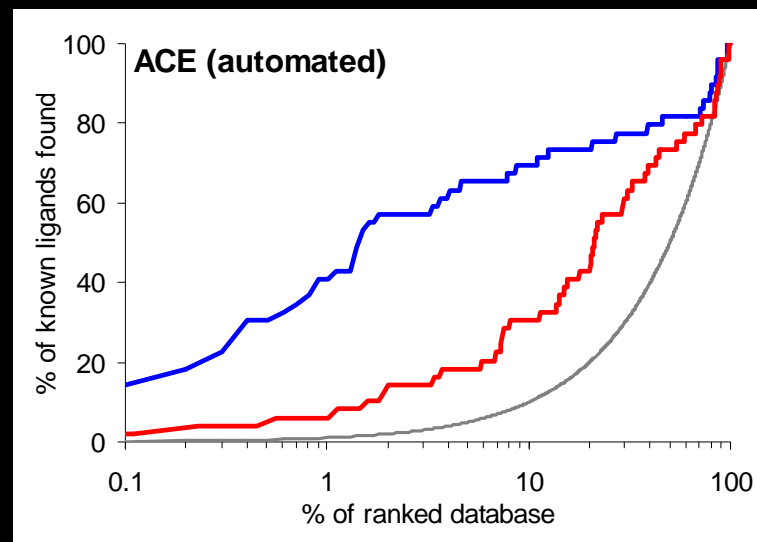


Pose-fidelity

# Is docking working?



Enrichment / Rank



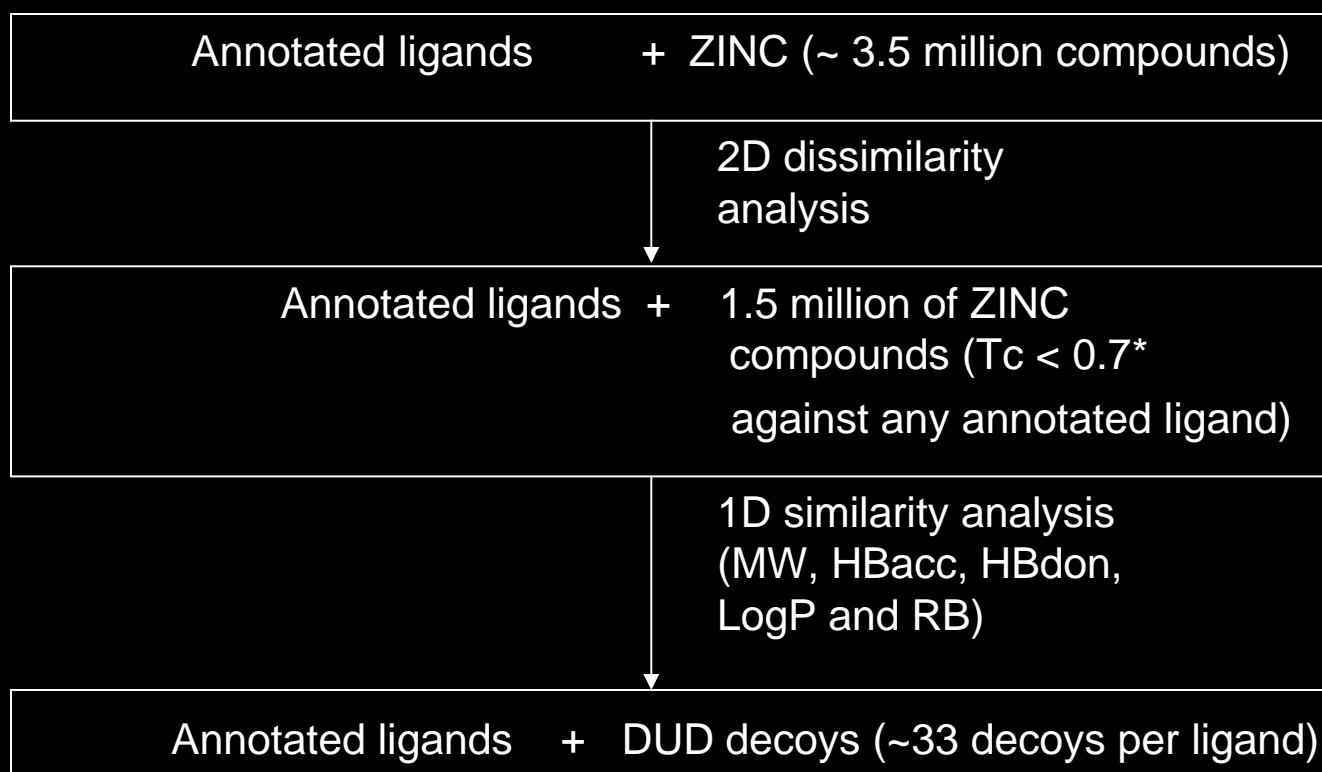
Single ligand metric and multi-ligand metric

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 <sub>antagonist</sub>	40	FGFr1	118	Thrombin	65	PARP	33
ER <sub>agonist</sub>	67	SRC	162	COMT	12	ALR2	26
AR	74	P38 MAP	234	ADA	23	PNP	25
RXR $\alpha$	20	PDGFr $\beta$	156	ACE	49	SAHH	33
PPAR $\gamma$	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

DOCK Blaster - A Virt... ZINC is not commerci... DUD - A Directory of ... Entrez PubMed

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center



## DUD

### A Directory of Useful Decoys

Welcome to DUD, a directory of useful decoys for benchmarking virtual screening. DUD is designed to help test docking algorithms by providing challenging decoys. It contains:

- A total of 2,950 active compounds against a total of 40 targets
- For each active, 36 "decoys" with similar physical properties (e.g. molecular weight, calculated LogP) but dissimilar topology.

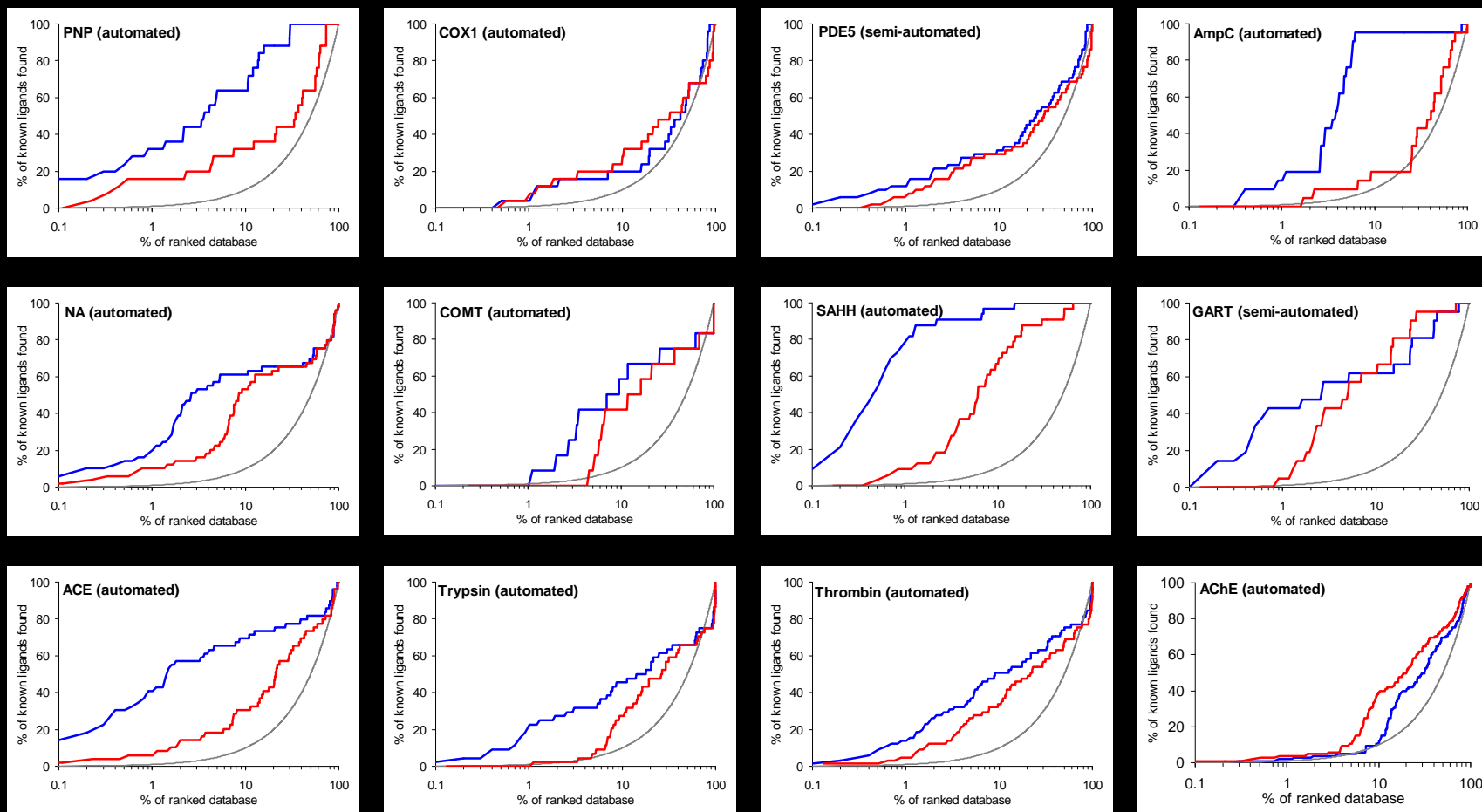
DUD is provided by the [Shoichet Laboratory](#) in the [Department of Pharmaceutical Chemistry](#) at the [University of California, San Francisco \(UCSF\)](#). To cite DUD, please reference **Huang, Shoichet and Irwin, manuscript submitted for publication [will be updated]**. We thank [NIGMS](#) for financial support (GM71896). For correspondence about DUD, please write John Irwin jji at cgl dot ucsf dot edu.

DUD is drawn from [ZINC](#), a database of commercially available compounds for virtual screening, so compounds in DUD are purchasable, although some may become depleted in the future. You may download DUD either in packages (some of which are large!) or you may browse the files and download them individually.

#### Downloads

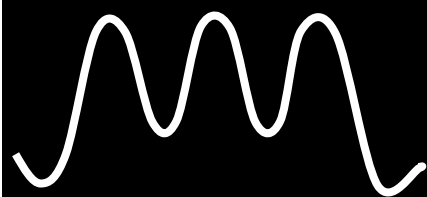
- Multi-target packages:
  - [All DUD Ligand sets \(mol2 format\)](#)
  - [All DUD Decoy sets \(mol2 format\)](#)
  - [All targets \(PDB format\)](#)
  - [All structural ligand controls \(mol2 format\)](#)
  - [Everything! All files for all targets](#)
- [Browse ligands and decoys](#)

# “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

← → ↻ ☆ <http://blaster.docking.org/start.shtml> ▶ 📁 🔧

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center [docking.org](http://docking.org)

## Prepare input

# DOCK Blaster

Home ▾ DOCK! ▾ Documentation ▾ Mailing Lists ▾

**Preparer** Scrutinizer Target Prep Calibration Docking Results  
Q'd Run Done Q'd Run Done Q'd Run

Welcome to the DOCK Blaster Job Preparation Form! To start docking, please complete the form below. If you need help, please read the [documentation](#). We urge you to attempt at least one [tutorial](#) using prepared data before attempting to use DOCK Blaster with your own data.

[Leave A Message](#)  
✉  
[Live Help Offline](#)

Target:  rec.pdb  
Indicate binding site: ▾  xtal-lig.mol2  
Actives:  actives.smi  
Inactives:  inactives.smi  
Email for reports:  (optional)

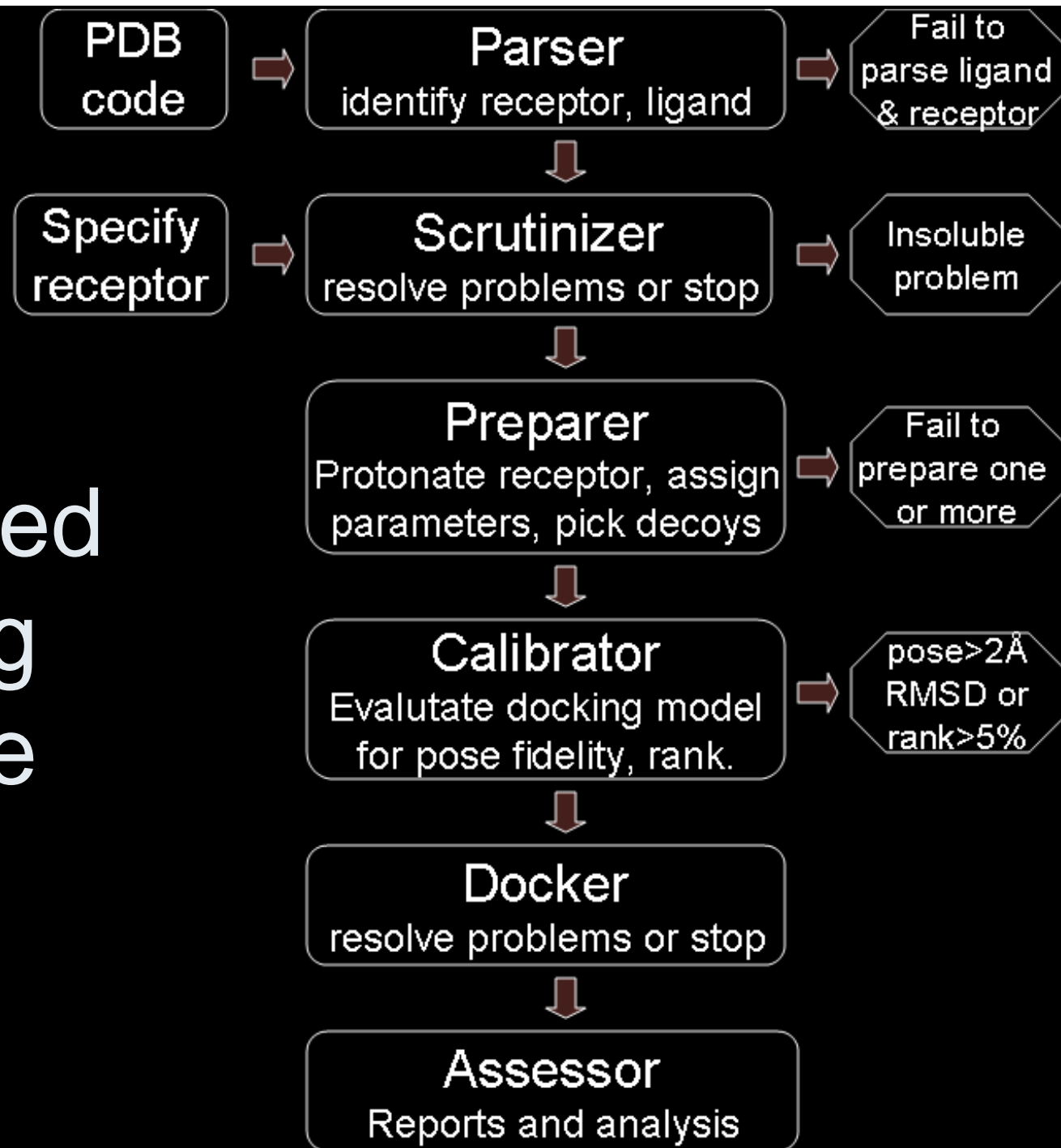
By clicking "DOCK!" you agree to the [Terms and Conditions](#).  Optional PIN:

**The aim of this calculation is...?**

Cofactor.par:  No file chosen Expert.tar:  No file chosen  
[Expert Help](#)

A product of [BCIRC](#), the Bioinformatics and Chemical Informatics Research Center @ UCSF. Last updated Feb 11, 2009. questions and discussion to blaster-fans at docking.org; bug reports to support at docking.org; any other correspondence to comments at docking.org. [Terms of use](#). [Privacy policy](#).

# Automated docking pipeline



# Try docking four ways

*Scoring*

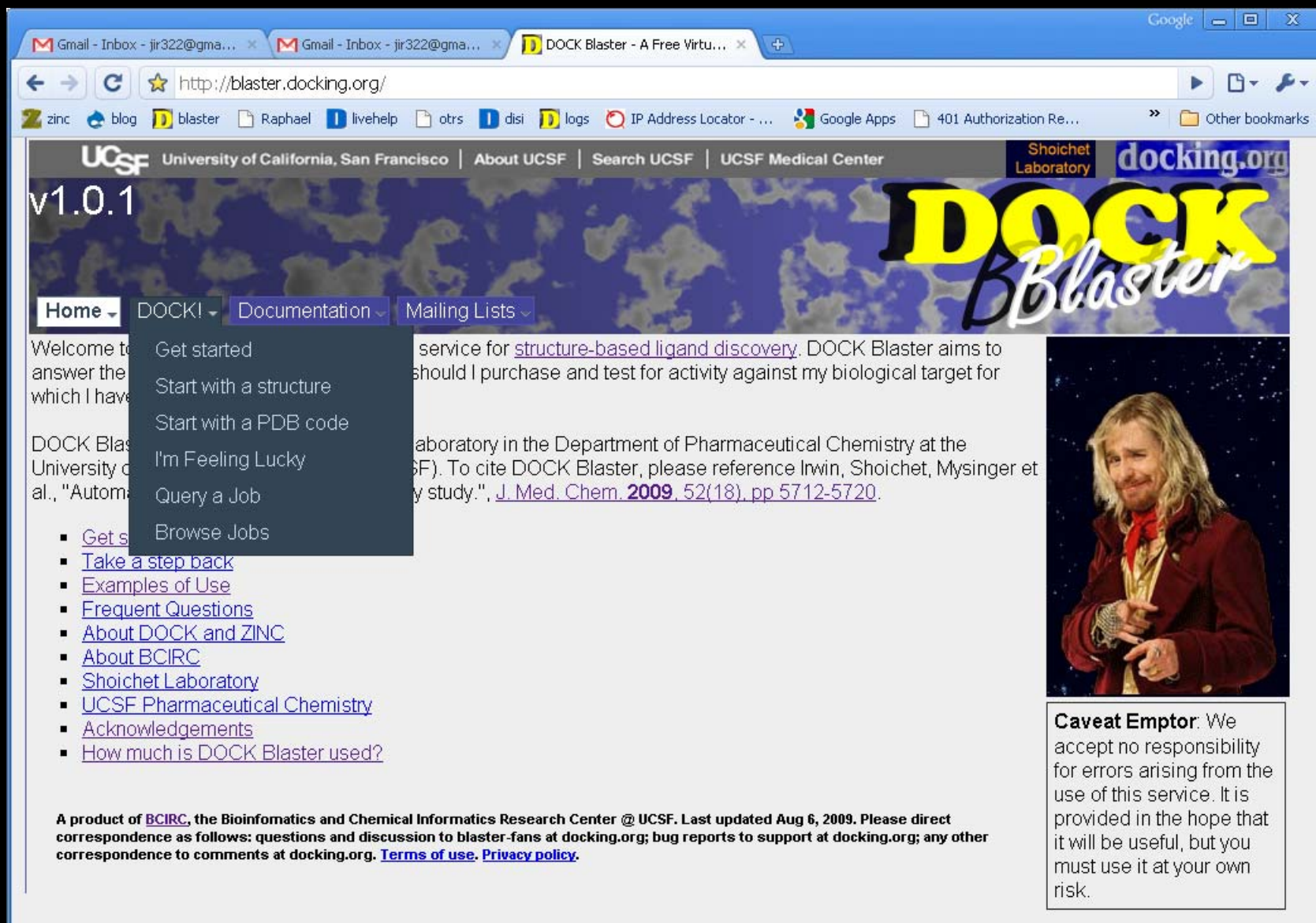
*Sampling*

	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!

Browser tabs: Gmail - Inbox - jir322@gma..., DOCK Blaster - A Free Virtu...

Address bar: http://blaster.docking.org/parser.shtml

Navigation: Home | **DOCK!** | Documentation | Mailing Lists

## Dock PDB target

# DOCK Blaster

Preparer	Scrutinizer	Target Prep	Calibration	Docking	Results
		Q'd Run Done	Q'd Run Done	Q'd Run	

Welcome to the DOCK Blaster Parser, where you can start docking with just a PDB code\*  
To start you need a target (and sometimes a ligand). To get a PDB code: (\* restrictions apply)

- To find a PDB code, you may [search the PDB](#).
- For a demonstration that will work, try human thyroid receptor beta (PDB code 1N46).
- You may pick from a list of 9870 eligible structures containing a ligand
- Or if you want to take your chances, try clicking: "I'm Feeling Lucky" (below)

Some [documentation](#) is available.

PDB Code:  ([9870 eligible PDB structures with ligand](#), PDB 2008-08)

Ligand:  (optional)

Email:  (optional)

The aim of this experiment is:

By clicking "DOCK!" you agree to the [Terms and Conditions](#).

A product of [BCIRC](#), the Bioinformatics and Chemical Informatics Research Center @ UCSF. Last updated Aug 6, 2009. Please direct correspondence as follows: questions and discussion to [blaster-fans at docking.org](#); bug reports to [support at docking.org](#); any other correspondence to [comments at docking.org](#). [Terms of use](#). [Privacy policy](#).

# Review docking hits. Click on "CHIMERA"

DOCK Blaster - A Free Virtu... x

http://blaster.docking.org/cgi-bin/resultsbrowser.pl?job\_id=39669&i=run.128pin=

zinc blog blaster Raphael livehelp otrs disi logs IP Address Locator - ... Google Apps 401 Authorization Re... Other bookmarks

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center

Shoichet Laboratory docking.org

## Results Browser

### Job #39669

Home DOCK! Documentation Mailing Lists

Preparer Scrutinizer Target Prep Calibration **Docking** Results

Q'd Run Done Q'd Run Done Qd Run

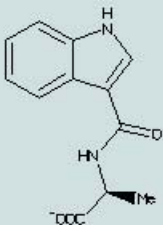
not finished  
These incomplete results of a running job are 7 minutes old. Please wait 25 minutes to refresh  
Files supplied: [Receptor](#), [Ligand](#).

Visualize using: [CHIMERA](#) [PYMOL](#) [Important note about broken molecules](#).

Download hits: [PDB](#) [SDF](#) [EEL1](#) [mol2](#)

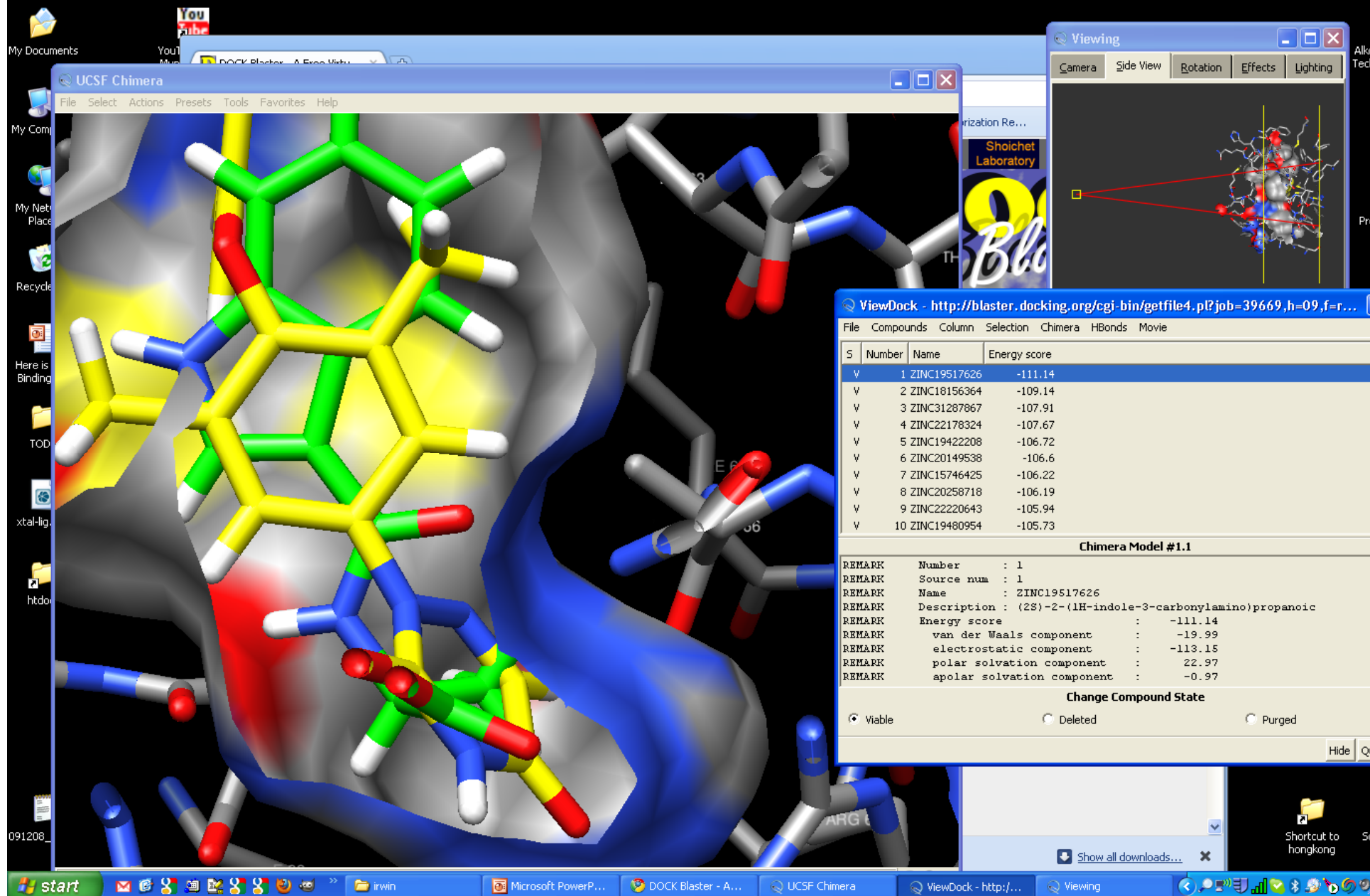
[Leave A Message](#)  
[Live Help Offline](#)

[Get Chimera!](#) [Get Pymol!](#)

Rank	Compound	Availability	Properties	Annotations
<b>ZINC ID</b> <b>Score kcal/mol</b>				
1 <input type="checkbox"/> <a href="#">19517626</a> -111.14 <a href="#">FIND SIMILAR</a> <a href="#">Go SEA!</a> <a href="#">PDB</a>		Aurora BuildingBlocks:bb-kuk-083487 UkrOrgSynthesis (Depleted):BBV-087716 Ambinter:BBV-087716 Aurora Feinchemie:bb-kuk-083487	Mwt: 231.231 xLogP: -0.60 Charge: -1 RotBond: 3 # Protomers: 1 Contact: 160 ES: -113.15 VdW: -19.99 Desolv: p=22.97, ap=-0.97	
		TimTec Building Blocks:SBB007469 Asinex:BAS00500340 Enamine (Depleted):T0505-2395		

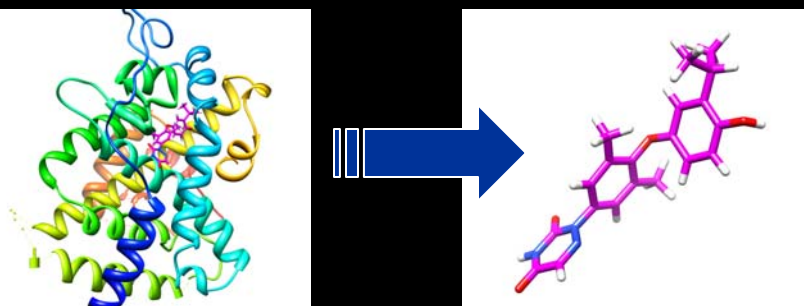


# Review hits in Chimera...

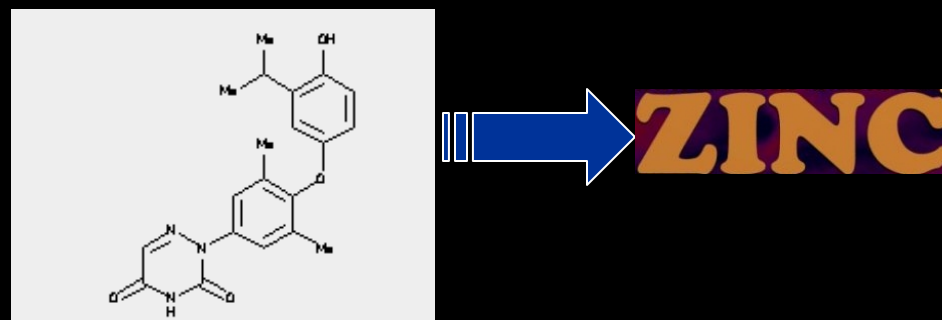


# #1. Self-assessment

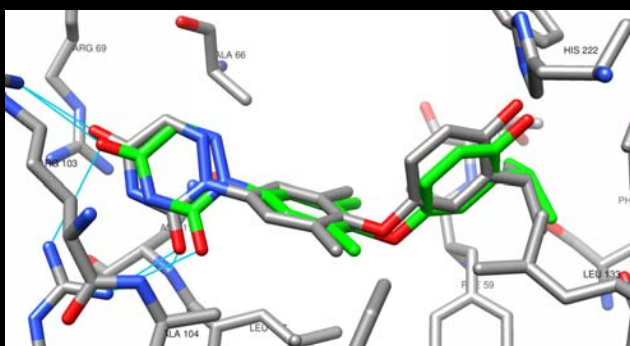
1. Remove ligand from receptor



2. Rebuild ligand without bias



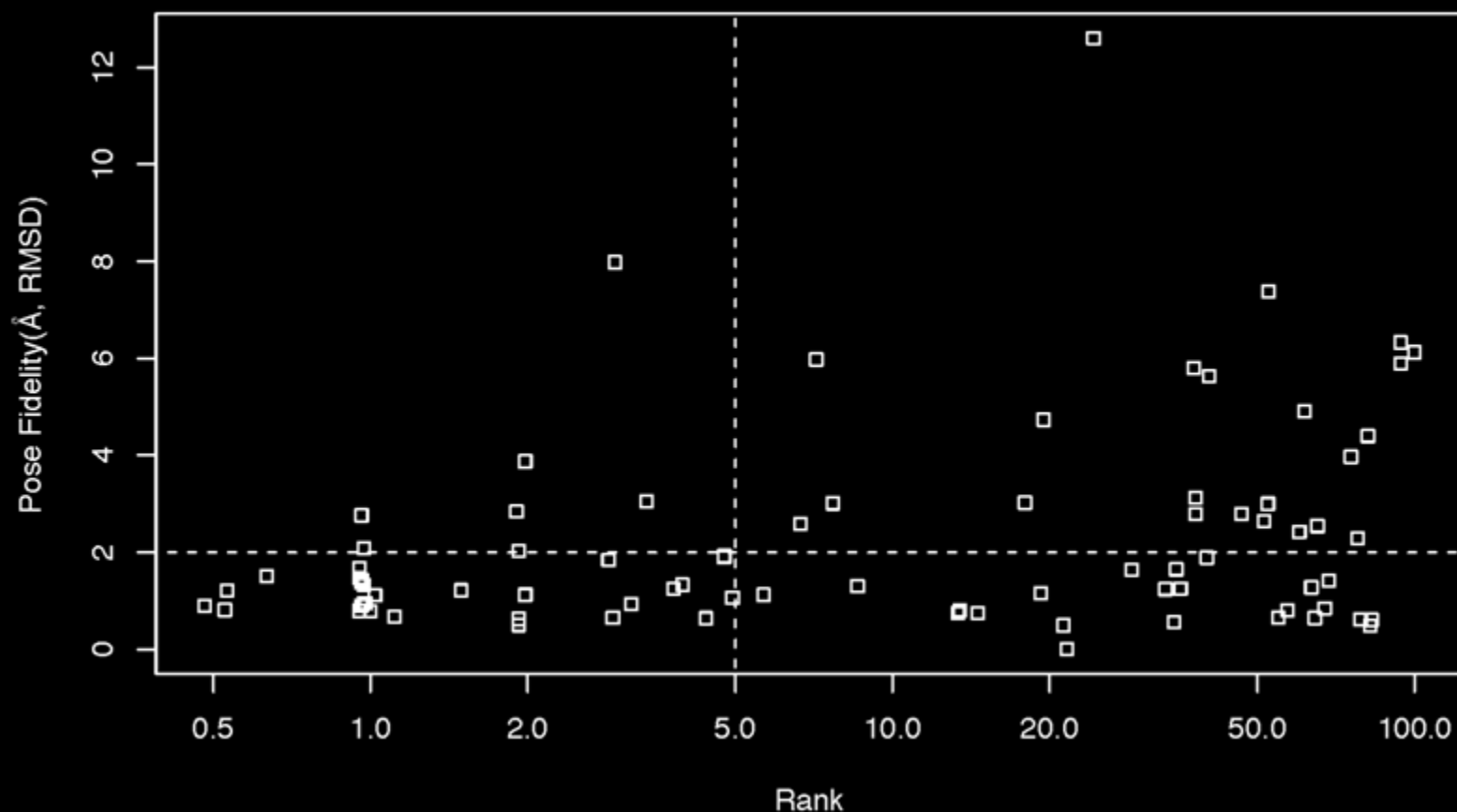
3. Dock ligand and 100 physically-matched decoys using 4 parameter sets



4. Evaluate pose-fidelity, enrichment

		Scoring	
Sampling		<i>Polarized</i>	<i>Normal</i>
	<i>Coarser</i>	3.61 Å / 1%	1.32 Å / 9%
	<i>Finer</i>	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

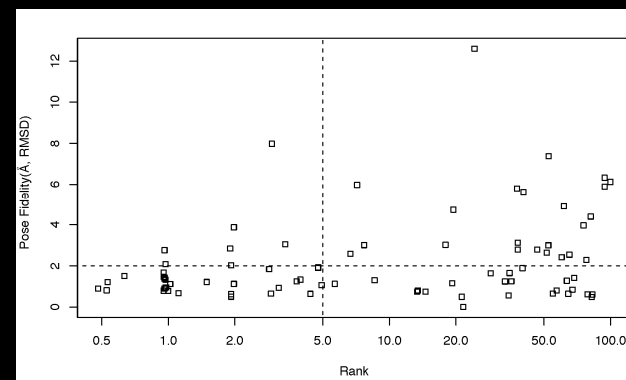
Description	<b>Astex- 85</b>	<b>Gold- 114</b>	<b>DUD- 38</b>	<b>PDB- 9050</b>
Ligand docked / scored	82	94	36	7,750
Good pose achieved	51	58	23	3,020
Good pose and rank	29	27	15	1,398

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	
Sampling		<i>Polarized</i>	<i>Normal</i>
	<i>Coarser</i>	3.61 Å / 1%	1.32 Å / 9%
	<i>Finer</i>	1.32 Å / 2 %	2.02 Å / 3%



http://blaster.docking.org/

# DOCK Blaster is free to use

The screenshot shows the DOCK Blaster website. At the top, there is a navigation bar with links for UCSF, University of California, San Francisco, About UCSF, Search UCSF, UCSF Medical Center, and docking.org. The main header features the text "v1.0β" and "DOCK Blaster" in large, stylized yellow and white fonts. Below the header, there is a navigation menu with options: Home, DOCK!, Documentation, and Mailing Lists. The DOCK! menu is expanded, showing a list of actions: Get started, Start with a structure, Start with a PDB code, I'm feeling lucky!, Query a Job, and Browse Jobs. The main content area includes a welcome message, a description of the service, and a list of links for further information. On the right side, there is a photograph of a man with long blonde hair, wearing a red jacket and a red scarf, pointing towards the camera.

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center [docking.org](http://docking.org)

v1.0β

# DOCK Blaster

Home ▾ DOCK! ▾ Documentation ▾ Mailing Lists ▾

Welcome to the DOCK Blaster service for [structure-based ligand docking](#). The service is designed to answer the following question: What is the best molecule to bind to a structure?

DOCK Blaster is a web-based service for [structure-based ligand docking](#). The service is designed to answer the following question: What is the best molecule to bind to a structure?

laboratory in the Department of Pharmaceutical Chemistry, University of California, San Francisco (UCSF). Please reference Irwin et al., *J. Med. Chem.* 46: 2793-2802 (2003).

The normal place to start is [here](#). To skip ahead, try the [tutorials](#) which illustrate by example some common applications. To back up to the bigger picture, try [this](#). Or try one of these links:

- [Get started](#)
- [DOCK and ZINC](#)
- [BCIRC](#)
- [Shoichet Laboratory](#)
- [UCSF Pharmaceutical Chemistry](#)
- [Acknowledgements](#)
- [How much is DOCK Blaster used?](#)

