M1 ISDD - BI (Feb 2018)

PROTEIN DOCKING

Lesson 4

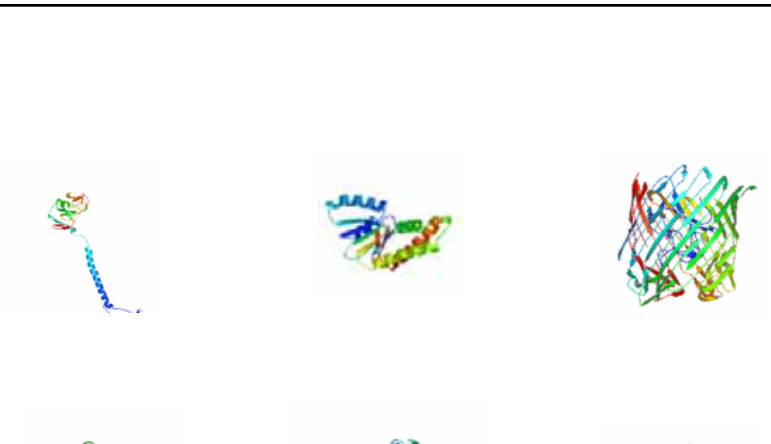
Flexible docking methods

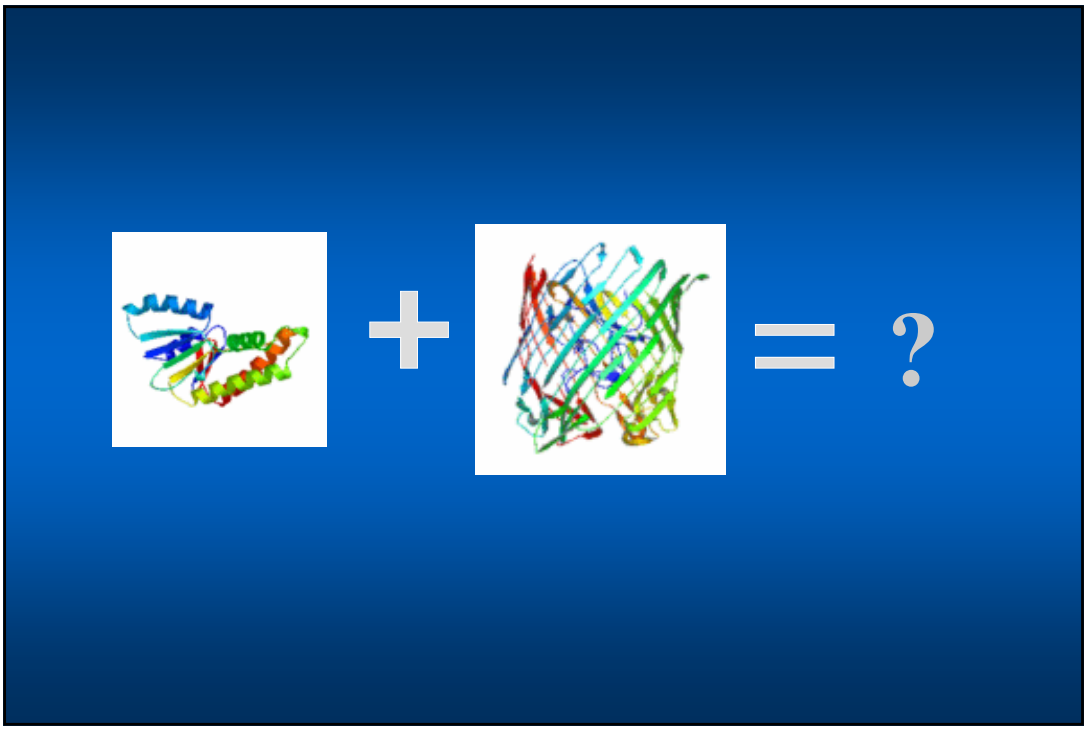
Juan Fernández-Recio

CSIC, BSC

juanf@bsc.es

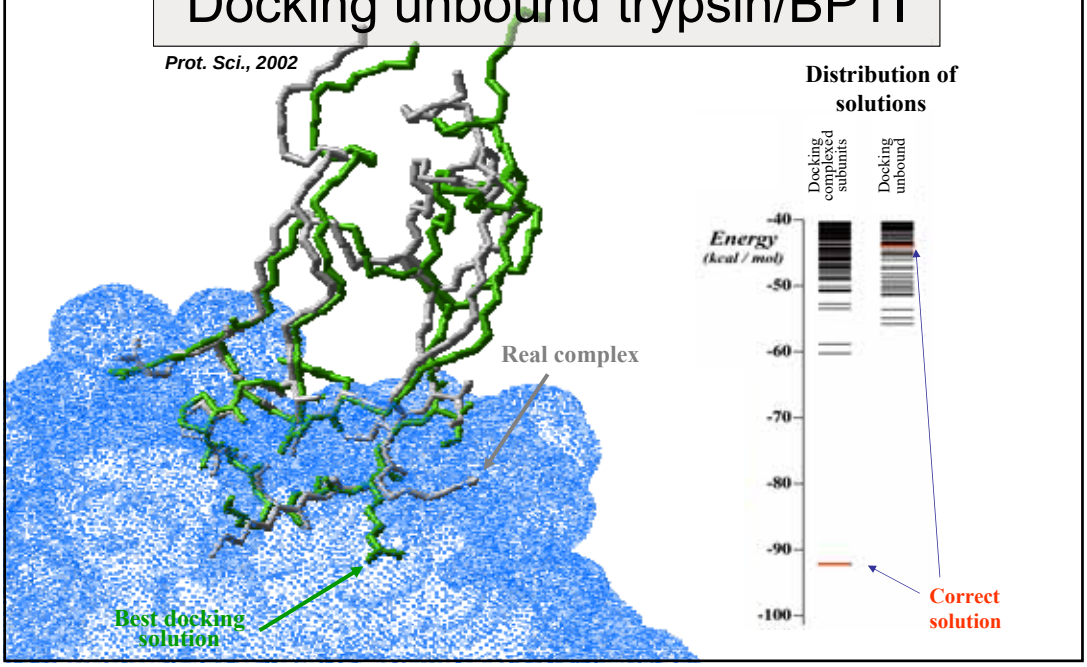
- Introduction
- Rigid-body + Refinement
- Flexible docking search
- Precomputed conformational ensembles
- Conclusions



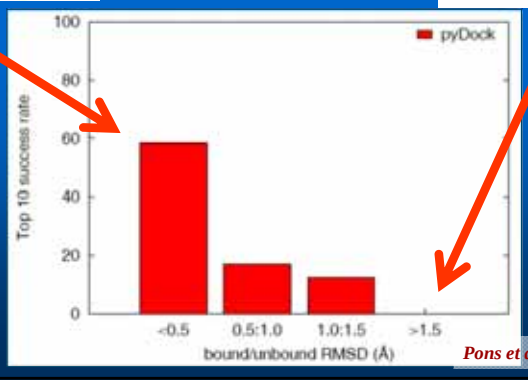


Docking unbound trypsin/BPTI

Prot. Sci., 2002



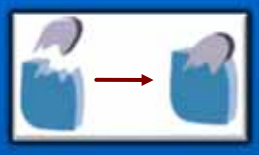
Rigid-body limitations: flexible cases



Pons et al. (2010) Proteins 78, 95-108

Protein-protein binding mechanism

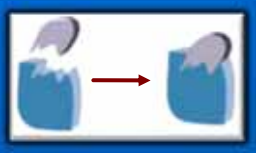
***1894**: LOCK AND KEY*: E. FISCHER



MOLECULES STERIC COMPLEMENTARY ESSENTIAL FOR AFFINITY AND SPECIFICITY IN THE BINDING.

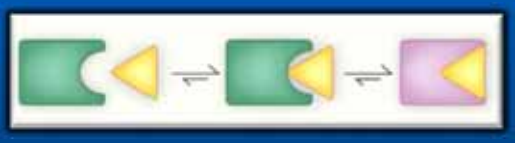
Protein-protein binding mechanism

***1894**: LOCK AND KEY*: E. FISCHER



MOLECULES STERIC COMPLEMENTARY ESSENTIAL FOR AFFINITY AND SPECIFICITY IN THE BINDING.

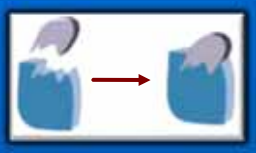
***1958**: INDUCED FIT:* D. E. KOSHLAND



INITIAL CONFORMATIONAL CHANGES INDUCE THE FINAL BINDING.

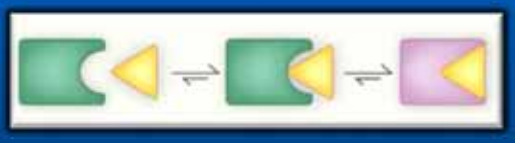
Protein-protein binding mechanism

***1894**: LOCK AND KEY*: E. FISCHER



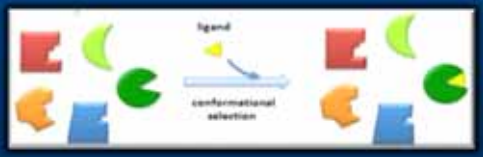
MOLECULES STERIC COMPLEMENTARY ESSENTIAL FOR AFFINITY AND SPECIFICITY IN THE BINDING.

***1958**: INDUCED FIT:* D. E. KOSHLAND

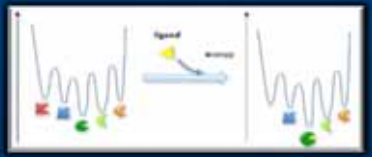


INITIAL CONFORMATIONAL CHANGES INDUCE THE FINAL BINDING.

***1999:** CONFORMATIONAL SELECTION:* R. NUSSINOV



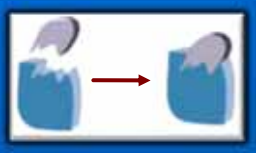
PROTEINS AS CONFORMATIONAL ENSEMBLES



INITIAL INTERACTION, EQUILIBRIUM SHIFT AND FINAL PROTEIN BINDING.

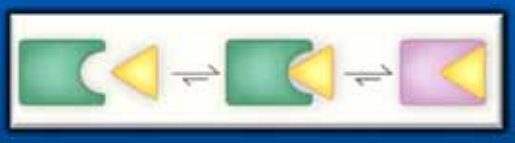
Protein-protein binding mechanism

***1894**: LOCK AND KEY*: E. FISCHER



- RIGID-BODY DOCKING

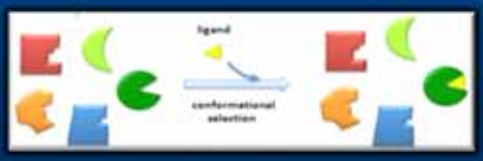
***1958**: INDUCED FIT:* D. E. KOSHLAND

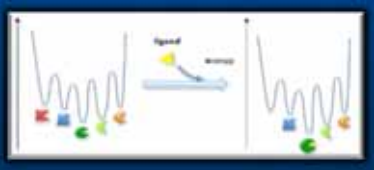


- *small:* RIGID-BODY + REFINEMENT

- *large:* FLEXIBLE DOCKING SEARCH

***1999:** CONFORMATIONAL SELECTION:* R. NUSSINOV



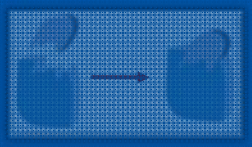


- PRECOMPUTED ENSEMBLES

- Introduction
- **Rigid-body + Refinement**
- Flexible docking search
- Precomputed conformational ensembles
- Conclusions

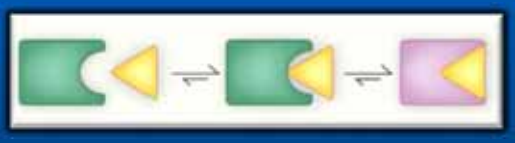
Protein-protein binding mechanism

1894: LOCK AND KEY: E. FISCHER



- RIGID-BODY DOCKING

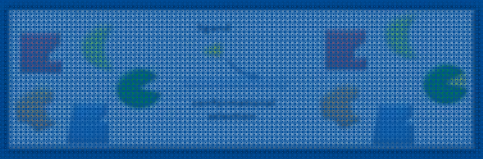
***1958**: INDUCED FIT:* D. E. KOSHLAND

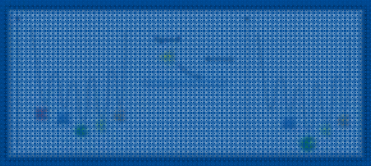


- *small:* RIGID-BODY + REFINEMENT

- *large:* FLEXIBLE DOCKING SEARCH

1999: CONFORMATIONAL SELECTION: S. KUMAR

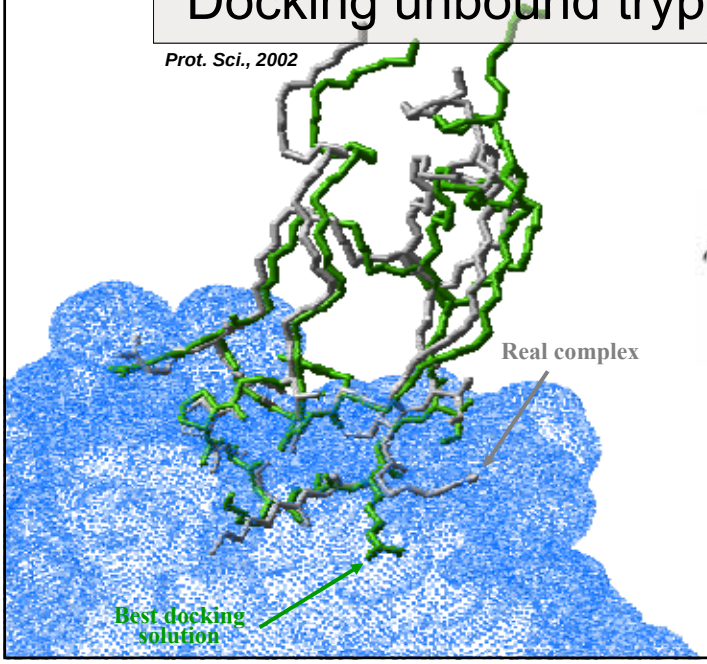




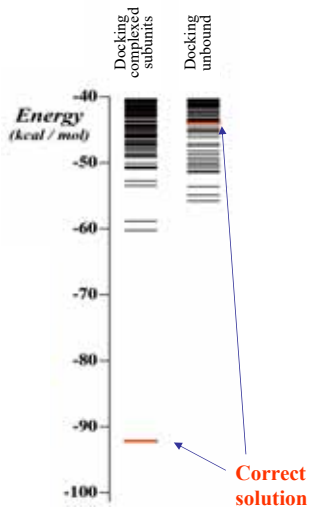
- PRECOMPUTED ENSEMBLES

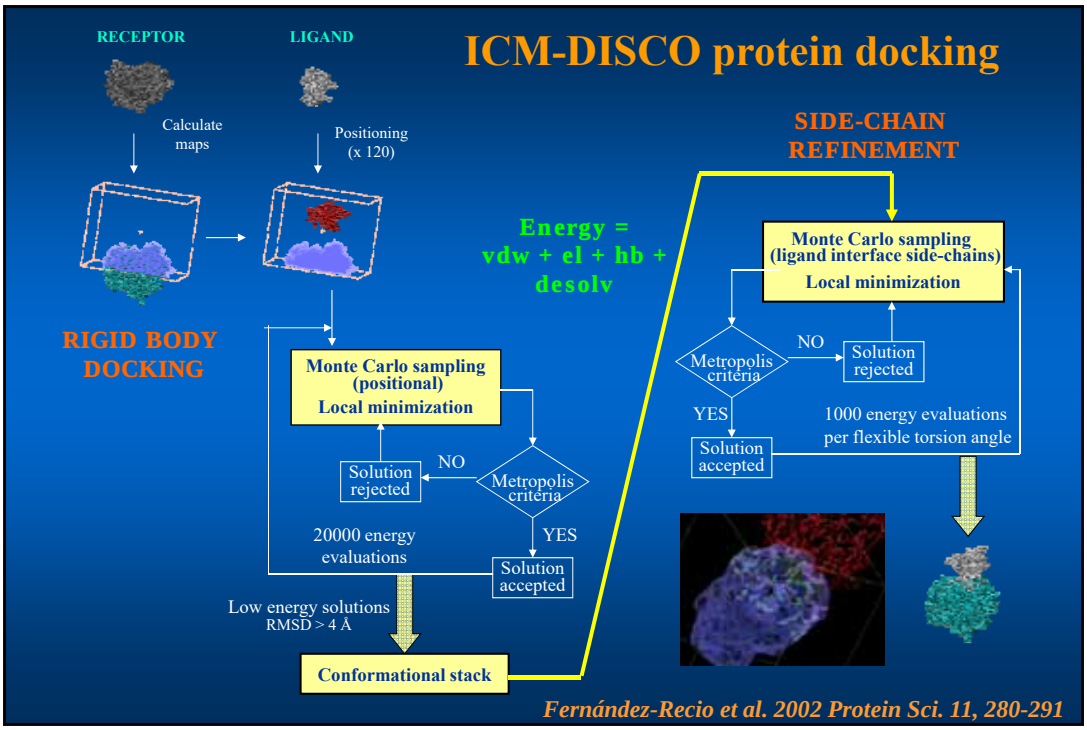
Docking unbound trypsin/BPTI

Prot. Sci., 2002



Distribution of solutions



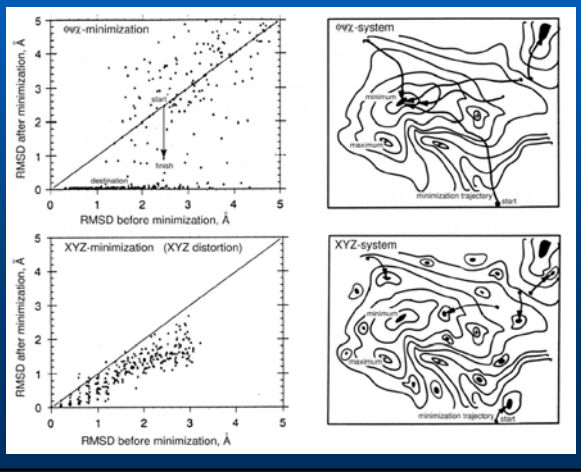


ICM Interface Side-Chain Optimization

Procedure:

Free-energy optimization with pseudo-Brownian moves and Biased Probability Monte Carlo

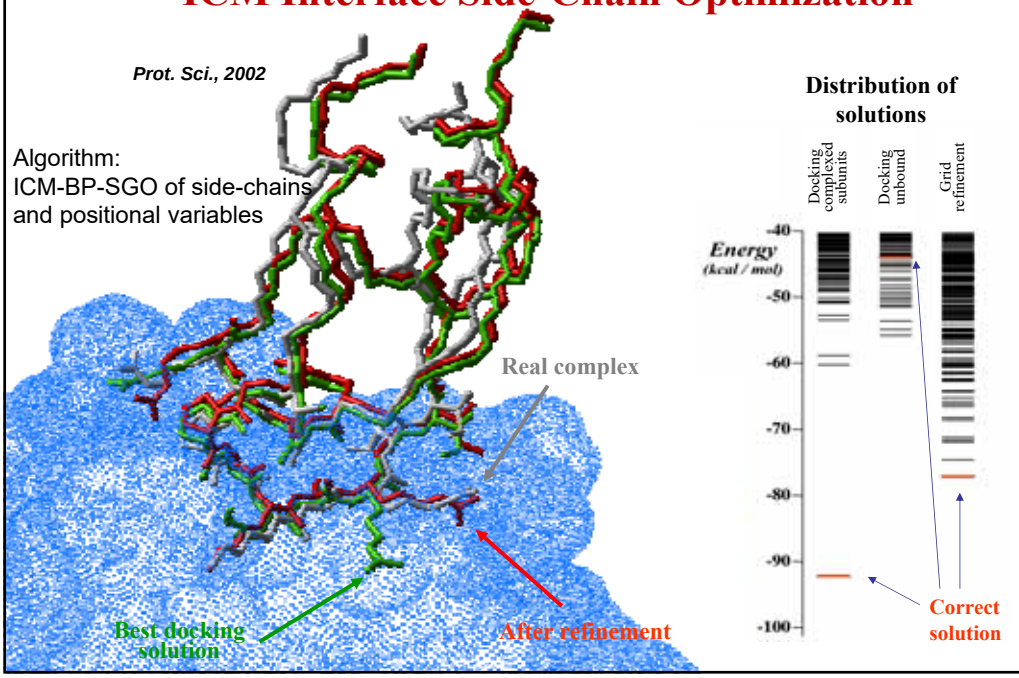
(Abagyan et al. 1994 J.Comp.Chem. 15, 488-506).



ICM Interface Side-Chain Optimization

Prot. Sci., 2002

Algorithm:
ICM-BP-SGO of side-chains
and positional variables

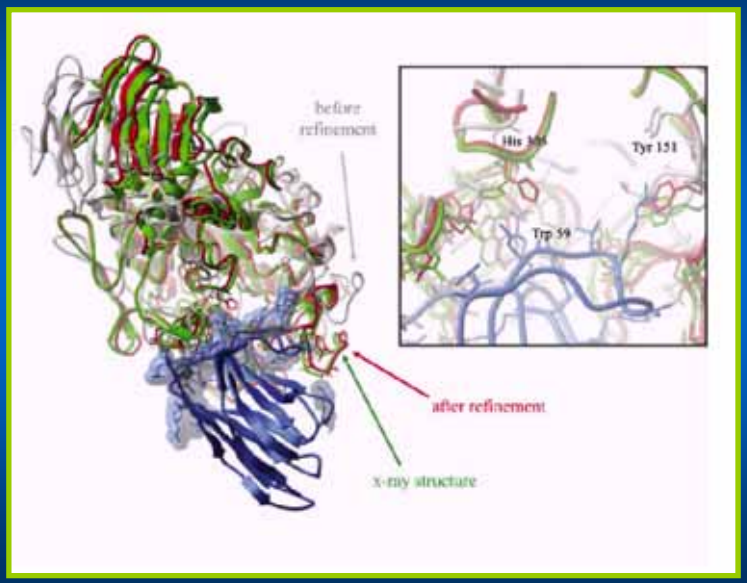


Distribution of solutions

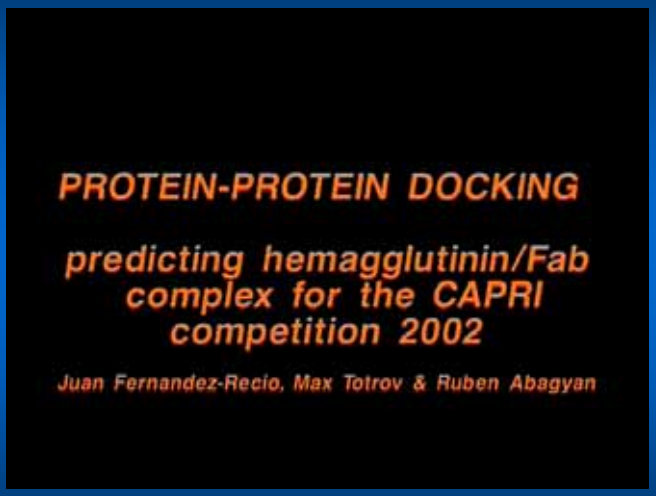
ICM-DISCO Results

α-amylase

VHH Ab



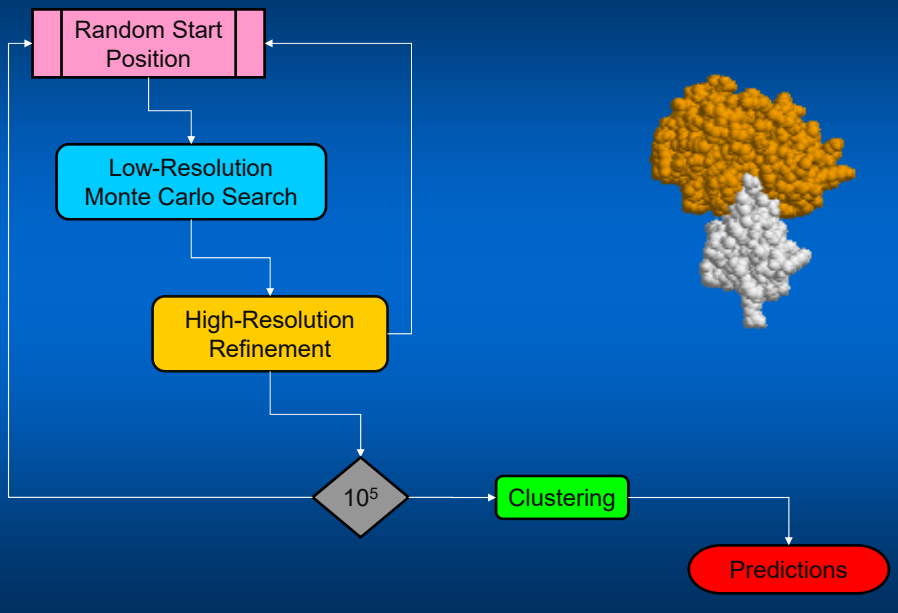
Rigid-Body Docking + Side-Chain Refinement



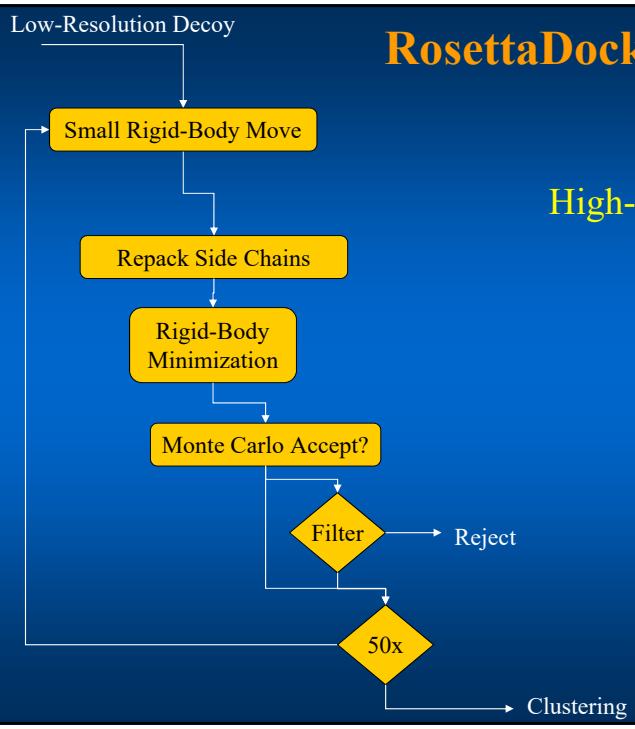
www.molsoft.com

Fernandez-Recio et al. (2003) Proteins 52, 113-117

RosettaDock



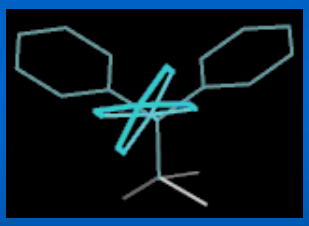
RosettaDock



High-Resolution Refinement

RosettaDock

1) **Random Rigid-body movement**
2) **Side-Chain Packing**
   - Choose side chains from Dunbrack's backbone-dependent rotamer library



Phenylalanine rotamers
(Richardson, 2000)

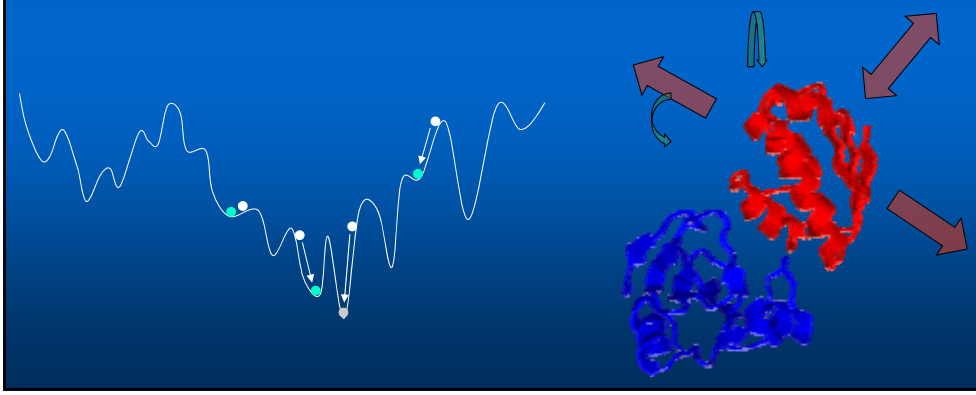
   - Vary χ_1, χ_2, χ_3, χ_4 angles
   - Minimize a full-atom energy function w.r.t. all rotamer combinations

(Brian Kuhlman & David Baker, Nature Struct. Biol. 2001)

RosettaDock

3) Full-atom rigid-body minimization

– Use a conjugate-gradient search to find the local score minimum relative to a rigid body translation and rotation



RosettaDock

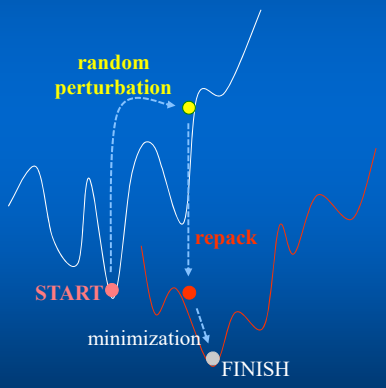
Full-Atom scoring

Score	Form / Source	Discriminatory z-value
Repulsive van der Waals	Modified Lennard-Jones 6-12	73.0
Attractive van der Waals	Lennard-Jones 6-12	45.0
Surface area solvation	Surface area (see Tsai 2003)	28.5
Gaussian solvent-exclusion	Lazaridis & Karplus, 1999	27.2
Rotamer probability	Dunbrack & Cohen, 1997	19.6
Hydrogen bonding	Empirical, Kortemme *et al*. 2003	14.9 & 6.8 (BB/BB)
Residue pair probability	Empirical, Kuhlman & Baker 2000	6.9
Electrostatics	Coulomb model with simple charges	0.4-15.1 (LR rep)

RosettaDock

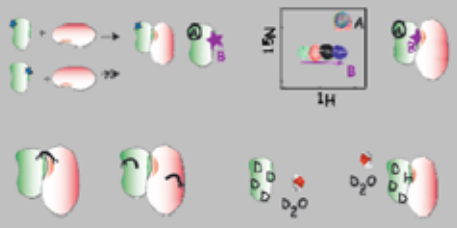
Summary of refinement cicle

(Simultaneous rigid-body displacement and side chain minimization)



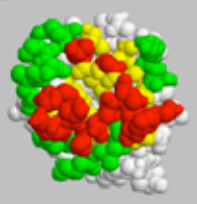
HADDOCK

Combine biochemical or biophysical data and semi-flexible docking to model biomolecular complexes



Interface definition:

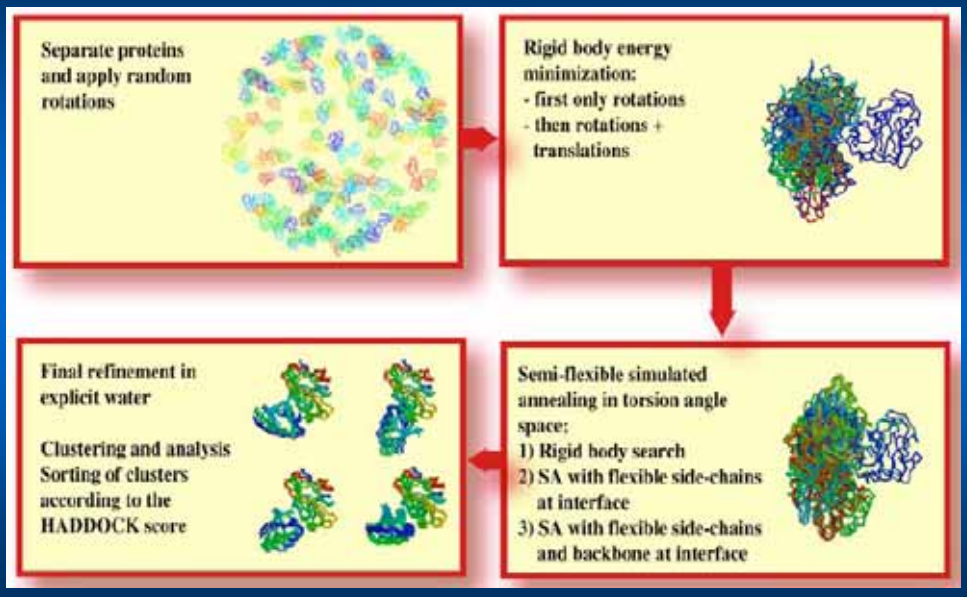
- Active residues: involved in the interaction (e.g. from NMR data, mutagenesis, ...) and high solvent accessibility
- Passive residues: all solvent accessible neighbors



Ambiguous interaction restraints:

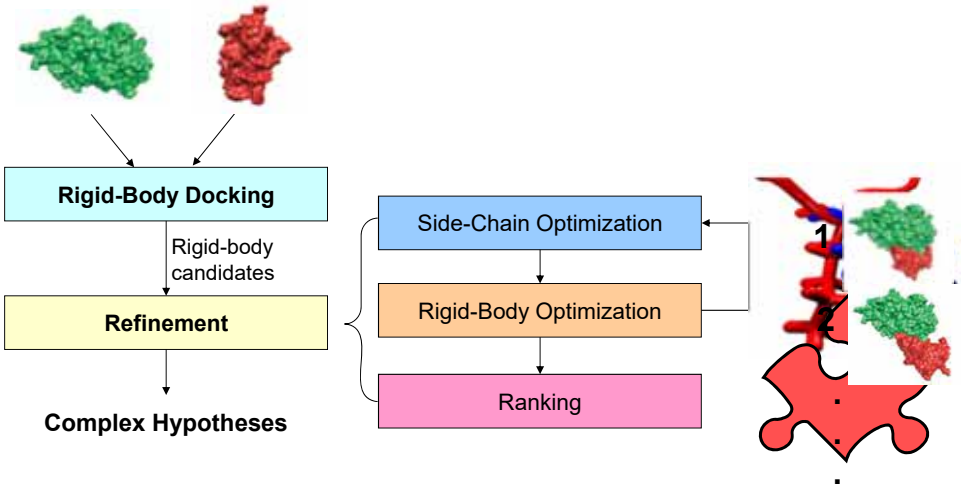
$$d_{IAB}^{eff} = \left(\sum_{m_{IA}=1}^{N_{atoms}} \sum_{k=1}^{N_{atoms}} \sum_{n_k=1}^{N_{atoms}} \frac{1}{d_{m_{IA}k}^{6}} \right)^{\frac{1}{6}}$$

HADDOCK



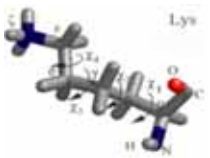
FireDock

Fast Interaction Refinement in Molecular Docking



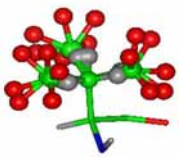
Interface Side-Chain Optimization

1. Selection of movable residues
   2 versions of Interface Side-Chain Optimization:
   - Full (FISCO)
   - Restricted (RISCO)
     only clashing residues are movable

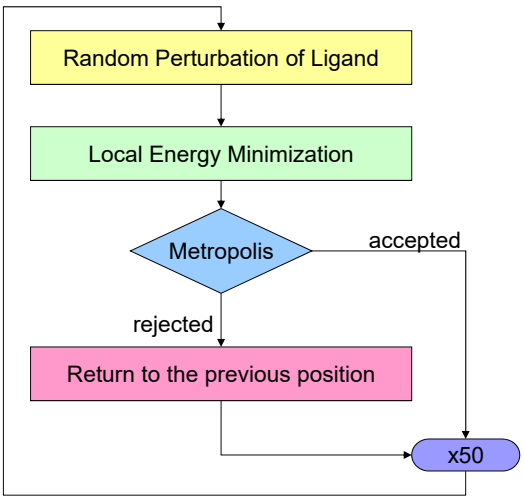


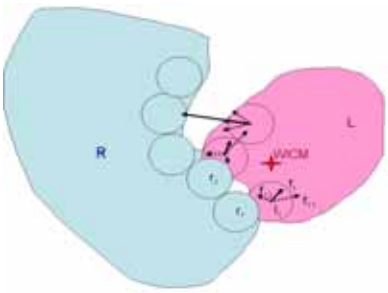
2. Energy calculation for backbone dependent rotamers

3. Solving the optimization using ILP



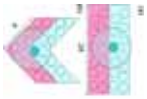
Rigid-Body Optimization



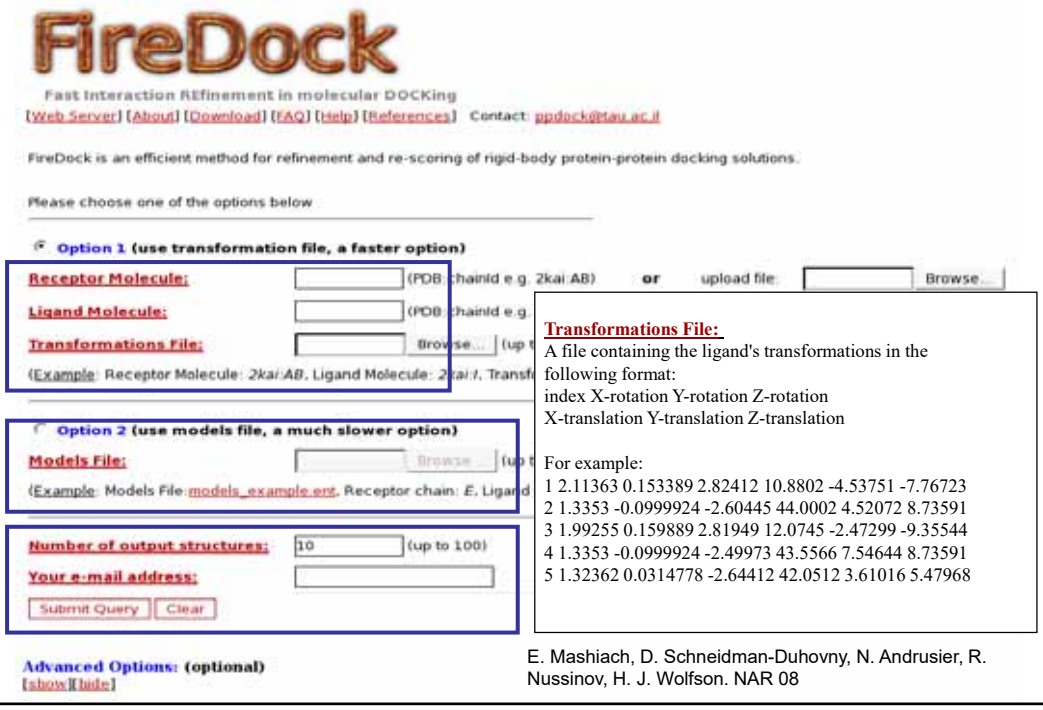


Ranking

- According to binding score:
  - ACE
  - electrostatics
  - π-stacking & aliphatic interactions
  - van der Waals
  - "insideness"



$$
\begin{aligned}
E_{EI} &= E_{s\_attrVdW} + 0.95E_{s\_repVdW} + 1.6E_{ACE} + \\
&\quad 0.07E_{attEl} + 0.12E_{repEl} + 0.3E_{l\_repEl} + \\
&\quad 1.32E_{HB+SS} + E_{pipi} + 0.8E_{catpi} + 0.5E_{aliph} + 1.55E_{insideness} \\
E_{AA} &= 1.5E_{s\_attrVdW} + 0.6E_{s\_repVdW} + 1.6E_{ACE} + \\
&\quad 0.21E_{attEl} + 0.21E_{repEl} + 0.46E_{l\_attrEl} + 0.69E_{l\_repEl} + \\
&\quad 1.2E_{HB+SS} + E_{pipi} + 0.7E_{catpi} + 2.5E_{aliph}
\end{aligned}
$$



E. Mashiach, D. Schneidman-Duhovny, N. Andrusier, R. Nussinov, H. J. Wolfson. NAR 08

FireDock

Fast Interaction REfinement in molecular DOCKing

[Web Server] [About] [Download] [FAQ] [Help] [References] Contact: ppdock@tau.ac.il

Advanced Options: **(optional)**
[show][hide]

Complex Type:	Default			
Refinement Level:	Restricted			
Number of RBO Cycles:	50			
Atomic Radius Scale:	0.8			
Bound/Unbound:	Receptor: Unbound		Ligand: Unbound	
Fixed Residues Files:	Receptor:	Browse...	Ligand:	Browse...
Flexible Residues Files:	Receptor:	Browse...	Ligand:	Browse...

Submit Query Clear

FireDock

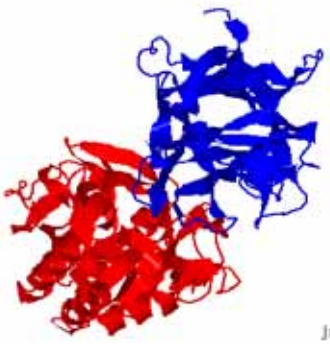
[Web Server] [About] [Download] [FAQ] [Help] [References]

Receptor	Ligand	TransFile	User e-mail
A.pdb	C.pdb	trans.txt	efratmas@gmail.com

Rank	Solution Number	Global Energy ↓	Attractive VdW	Repulsive VdW	ACE	HB	Structure show/hide
1	4	-22.88	-27.16	16.96	-11.83	-4.93	☑
2	64	-21.95	-21.95	7.16	-11.61	-3.60	☐
3	26	-21.50	-22.47	6.91	-11.01	-2.63	☐
4	112	-21.40	-25.74	16.40	-11.18	-3.96	☐
5	27	-20.67	-18.88	6.81	-9.82	-2.82	☐
6	158	-20.00	-26.22	9.17	-10.38	-2.59	☐
7	117	-19.27	-27.51	12.51	-11.22	-2.54	☐
8	131	-17.37	-21.60	8.38	-13.76	-1.44	☐
9	128	-17.01	-25.25	14.27	-9.50	-3.64	☐
10	97	-16.46	-17.62	5.80	-10.33	-4.85	☐
11	118	-16.40	-24.34	15.17	-9.13	-3.45	☐
12	154	-15.38	-23.17	16.41	-12.05	-3.80	☐
13	46	-15.22	-28.68	15.75	-10.95	-3.17	☐
14	36	-15.21	-19.25	8.58	-8.62	-4.05	☐
15	38	-14.84	-17.85	2.36	-8.15	-5.32	☐
16	156	-13.58	-22.35	15.07	-9.46	-3.92	☐
17	139	-12.25	-17.74	7.82	-11.66	-2.97	☐
18	120	-10.88	-25.93	20.57	-13.12	-1.73	☐
19	52	-9.29	-19.63	6.51	-8.62	-3.60	☐
20	111	-9.15	-21.51	10.50	-7.28	-2.76	☐

show next 20 »»

☐ show all/hide all



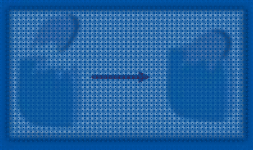
Jmol

download solutions table download best structures

- Introduction
- Rigid-body + Refinement
- **Flexible docking search**
- Precomputed conformational ensembles
- Conclusions

Protein-protein binding mechanism

1894: LOCK AND KEY: E. FISCHER

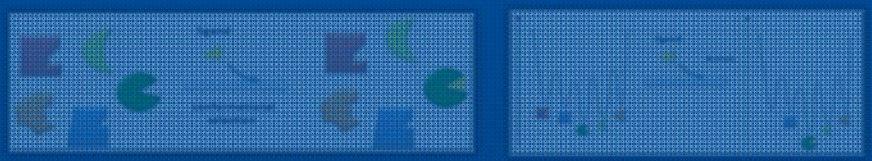


- RIGID-BODY DOCKING

***1958**: INDUCED FIT:* D. E. KOSHLAND

- *small:* RIGID-BODY + REFINEMENT
- ***large:*** **FLEXIBLE DOCKING SEARCH**

1999: CONFORMATIONAL SELECTION: S. KUMAR



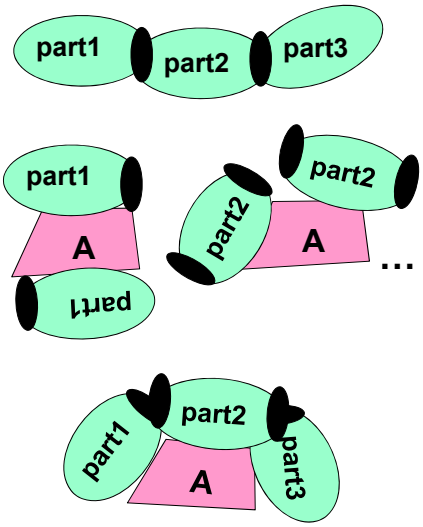
- PRECOMPUTED ENSEMBLES

FlexDock

Detection of Hinges and Rigid Parts in the Flexible Molecule

Rigid Parts Docking via Geometric Hashing

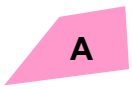
Assembly of partial dockings into a flexible result

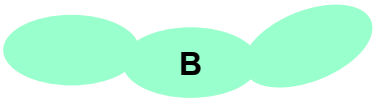


FlexDock Input

Rigid molecule (A)

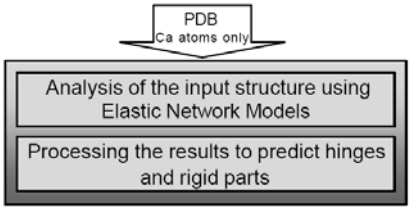
Flexible protein (B)





Can be a protein, DNA, RNA or small ligand molecule.

HINGEprot bioinfo3d.cs.tau.ac.il/HingeProt



Example: Calmodulin

Open conformation (PDB: 4cln)

Closed conformation (PDB: 1prw)

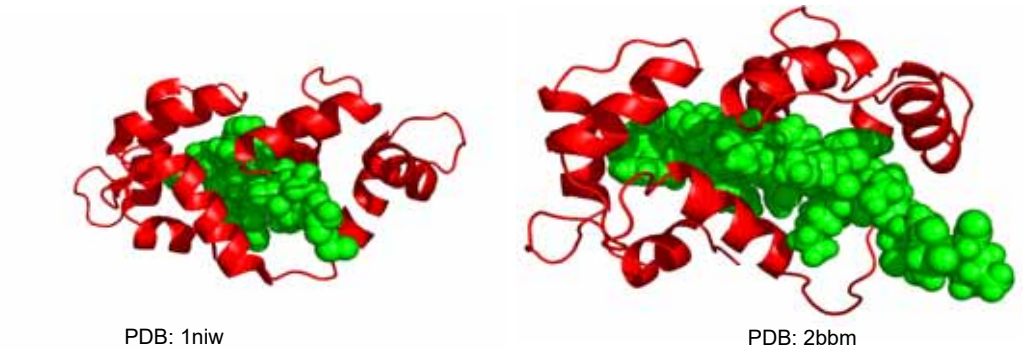
Example: Calmodulin

Open conformation (PDB: 4cln)

Closed conformation (PDB: 1prw)

- The hinge motion involves splitting of one long helix.
- The total rotation of one domain relative to the other is upwards of 150 degrees.
- The RMSD between two structures: >15Å

Calmodulin Bound Conformations

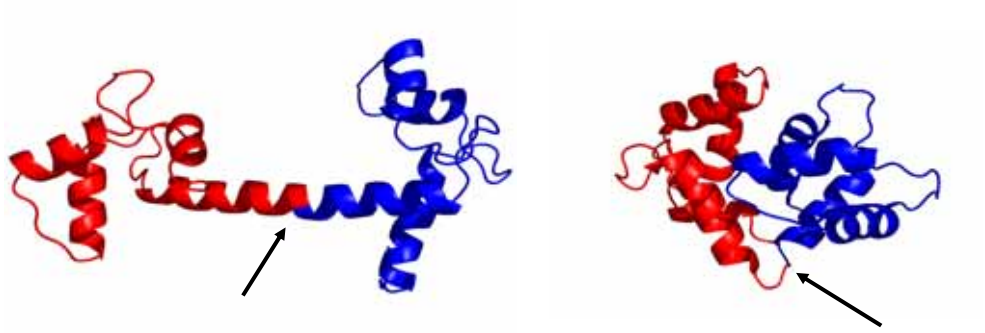


PDB: 1niw

PDB: 2bbm

- Upon binding calmodulin adopts conformation between open and closed conformations that fits its ligand

Calmodulin Partition by HingeProt

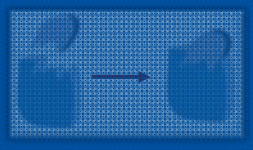


- HingeProt partitions both open and closed structures into two rigid parts (first slowest mode) with 77 as hinge residue

- Introduction
- Rigid-body + Refinement
- Flexible docking search
- **Precomputed conformational ensembles**
- Conclusions

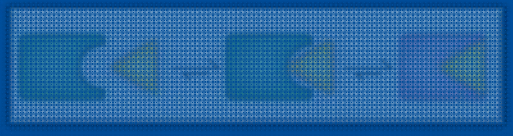
Protein-protein binding mechanism

1894: LOCK AND KEY: E. FISCHER



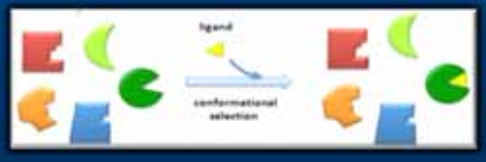
- RIGID-BODY DOCKING

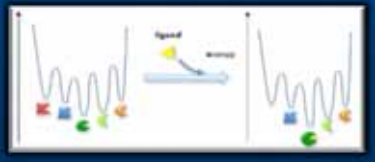
1958: INDUCED FIT: D. E. KOSHLAND



- *small:* RIGID-BODY + REFINEMENT
- *large:* FLEXIBLE DOCKING SEARCH

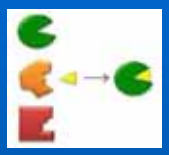
1999: *CONFORMATIONAL SELECTION:* S. KUMAR

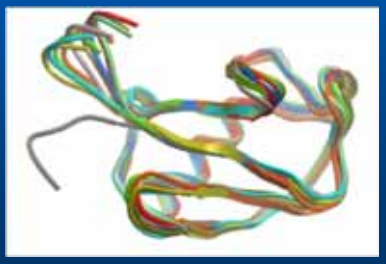


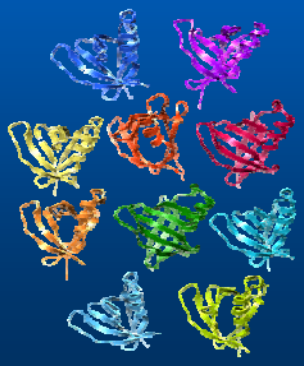


- **PRECOMPUTED ENSEMBLES**

Precomputed ensembles + docking





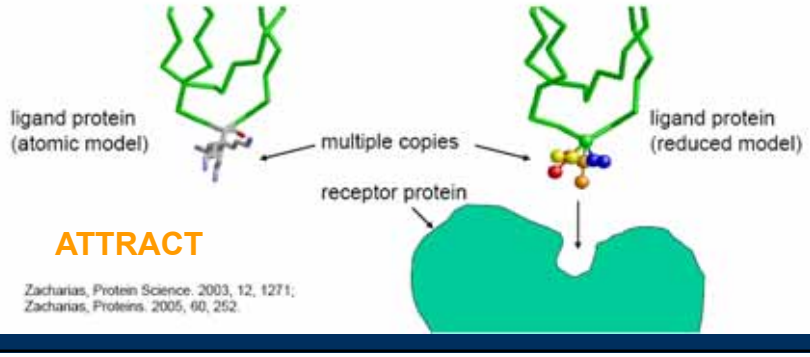


Precomputed ensembles + docking

Docking with multiple side chain copies

- Surface side chains are re-presented by several sterically allowed rotamer copies.
- Selection of most favorable copies during docking using a meanfield or switching approach.

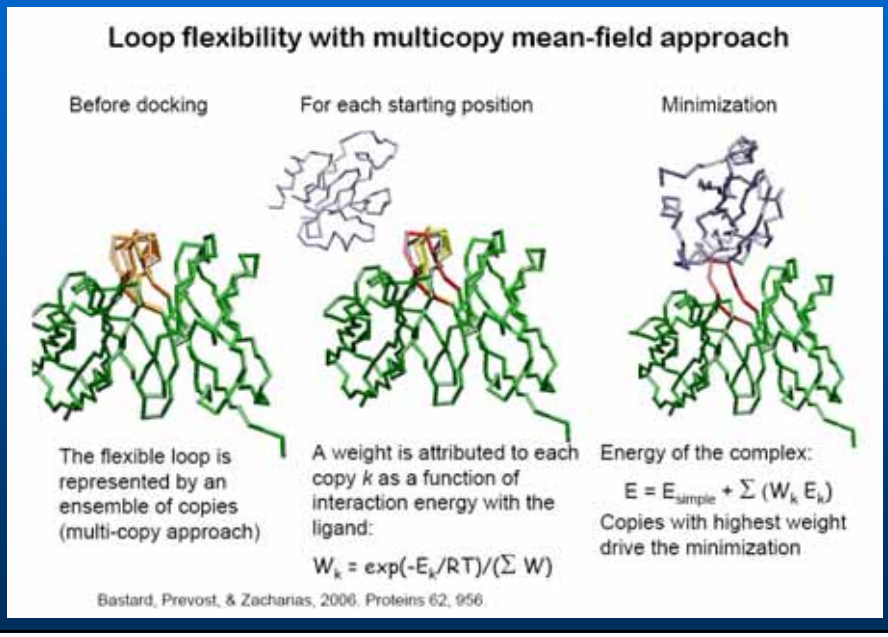
	- copies		+ copies	
Complex	Rank	$Lrmsd_A$	Rank	$Lrmsd_A$
Trypsin-BPTI	>120	7.1	11	1.4
Kallikrein-BPTI	>160	3.4	31	2.4
Chymo-BPTI	>150	3.4	29	3.2
SubtiN/Chy-Inh	>1100	3.8	59	3.0



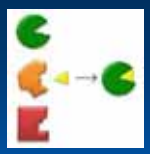
ATTRACT

Zacharias, Protein Science. 2003, 12, 1271;
Zacharias, Proteins. 2005, 60, 252.

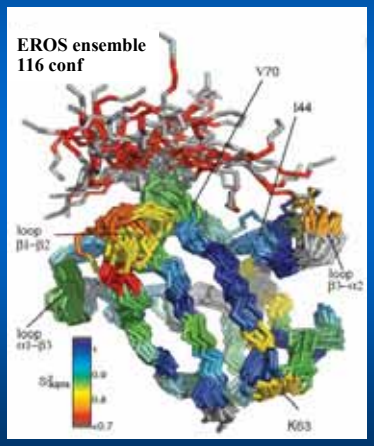
Precomputed ensembles + docking



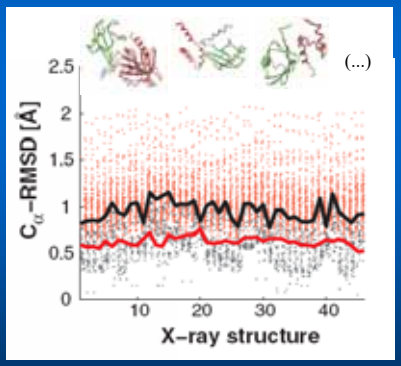
Multiple conformers: ubiquitine ensemble



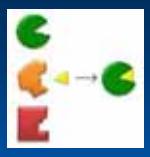
Conformational selection mechanism
(Lange et al. Science 2008, 320, 1471)



46 ubiquitin complexes



< 0.8 Å RMSD

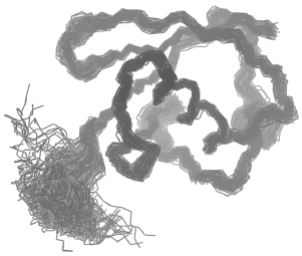


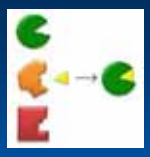
Multiple conformers: ubiquitine ensemble

RDC-based ensembles
(Fenwick et al. JACS 2011, 133, 10366)

PDB 1UBQ

ERNST ensemble
320 conf

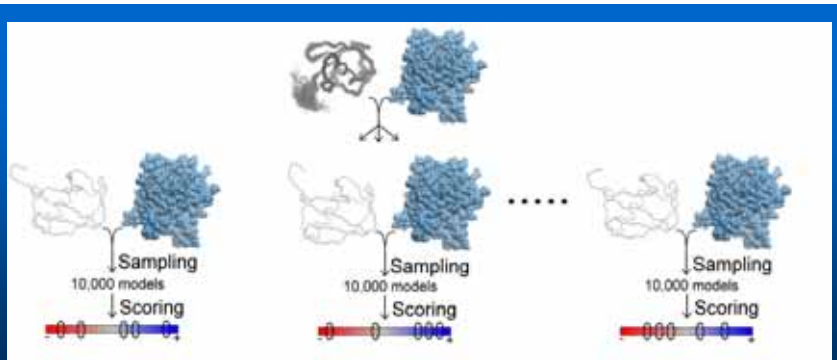




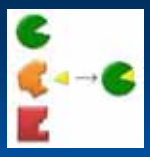
Multiple conformers: ubiquitine ensemble

RDC-based ensembles
(Fenwick et al. JACS 2011, 133, 10366)

11 complexes



x 320

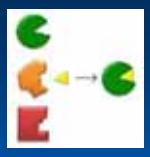


Multiple conformers: ubiquitine ensemble

RDC-based ensembles
(Fenwick et al. JACS 2011, 133, 10366)

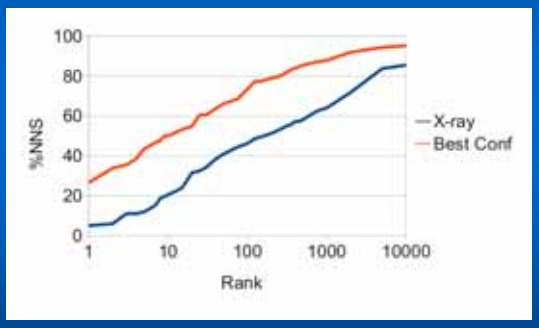
Case	Bound ubiquitin			Unbound ubiquitin		RDC Ensemble	
	single	10 rot	native incl	single	10 rot	Global	Best
1CMX	1	1	1	8	1	4	1
1NBF	1	1	1	3	21	9	1
1OTR	55	136	117	435	170	5	1
1P3Q	80	123	13	21	27	3	1
1S1Q	85	157	12	1031	4816	140	6
1WR1	167	541	541	337	2342	111	3
1XD3	1	1	1	1	2	4	1
1YD8	7	6	6	219	46	5	1
2AYO	532	3	3	74	54	70	1
2C7M	2	5	5	78	2	1	1
2G45	5	3	1	3654	40	117	2

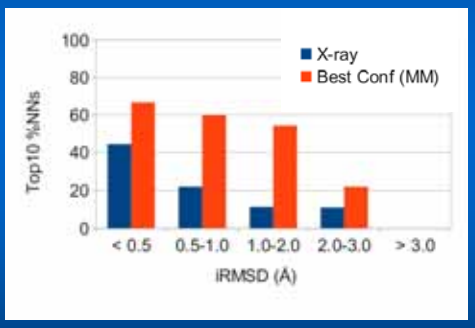
Pons et al. JCTC 2013



Multiple conformers: large-scale test

Protein docking benchmark 3.0 – 124 cases



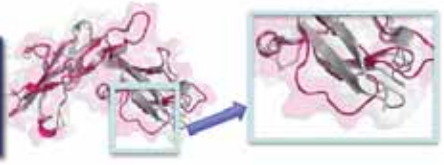


Pallara et al. (2016) JCTC

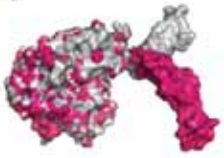
- Introduction
- Rigid-body + Refinement
- Flexible docking search
- Precomputed conformational ensembles
- **Conclusions**

Dealing with conformational changes in docking

- Local changes: (small) loop reorientations and structure changes
  - ✓ Ensemble Docking
  - ✓ Soft + flexible Docking



- Global changes: large scale domain motions (hinge, shear)



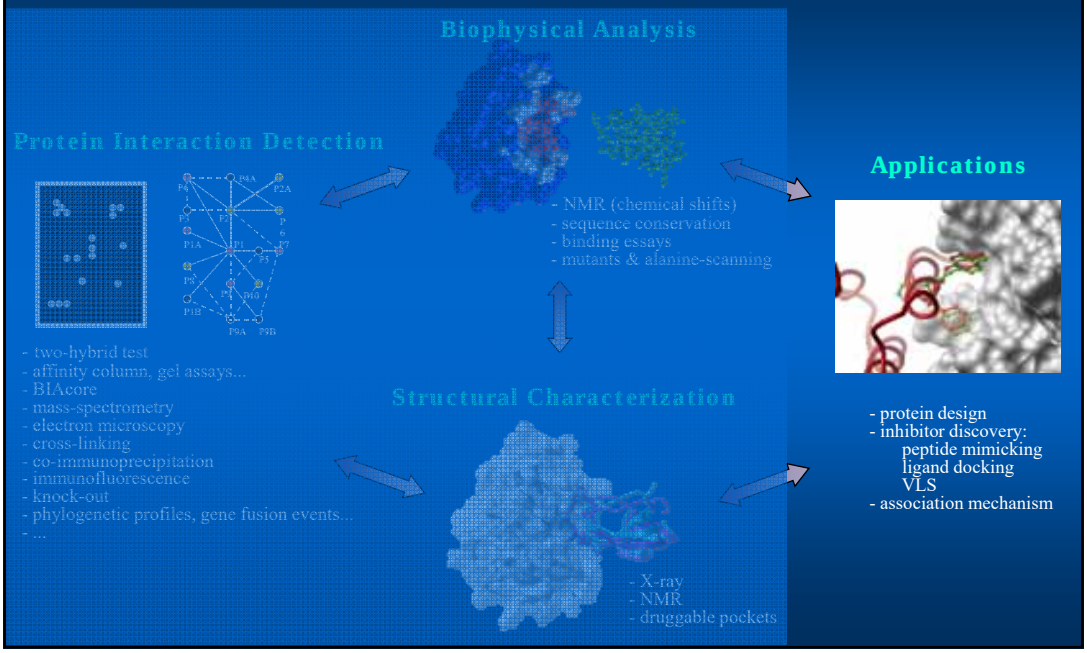
  - ✓ Divide and Conquer
  - ✓ Multi-Body Docking
- Binding-induced folding events.... ???

Flexible docking: conclusions

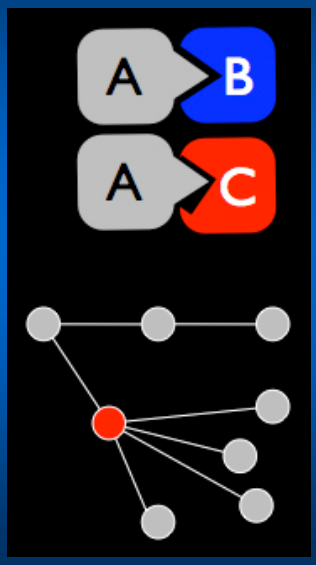
- *Flexibility is the main challenge*
- *Very costly computationally*
- *Refinement useful in small-conformational changes*
- *Flexibility during docking not yet practical*
- *Precomputed ensembles can work in medium-flexible cases*

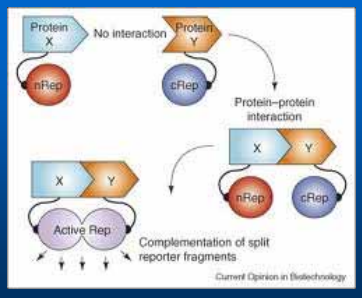
- Importance of protein interactions in biomedicine
- Study of protein-protein interactions
- Computational approaches
- **Applications**

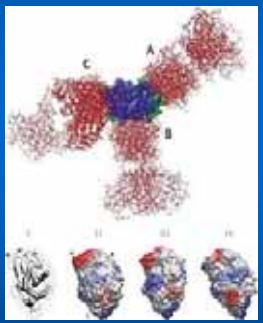
Study of Protein-Protein Interactions



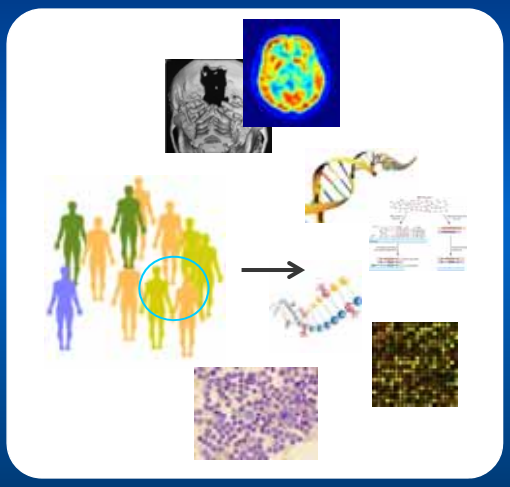
Applications in Biotechnology: Engineering of protein-protein interfaces







Applications in Biomedicine: Prediction and interpretation of pathological mutations



Population Pathology → Stratified Medicine

thousands of newly annotated mutations

BIG DATA

ANALYSIS

Applications in Biomedicine: Prediction and interpretation of pathological mutations

ANALYSIS

Identify Pathology *(precision diagnosis)*

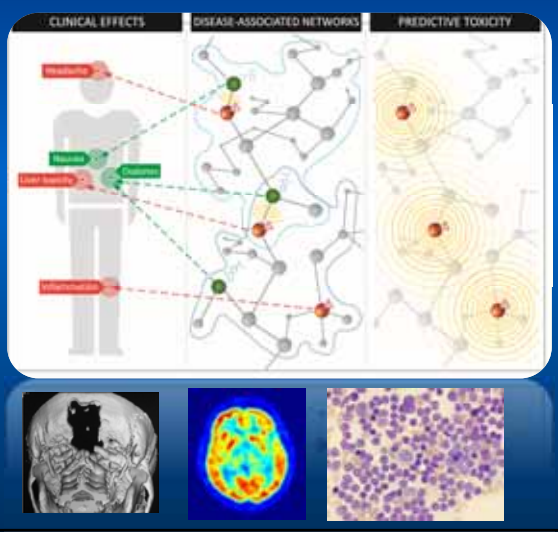
Improve treatment, drug response *(personalized treatment)*

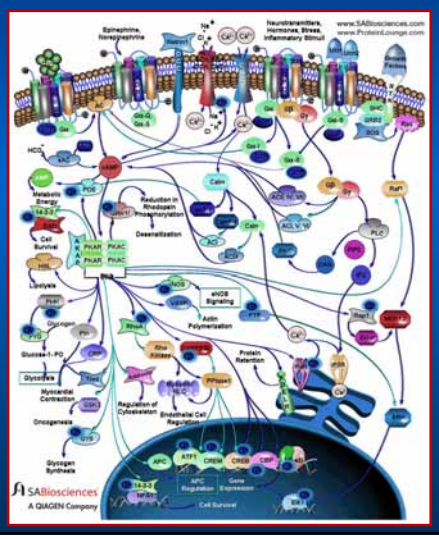
Annotated mutations → Understand cause of disease
New mutations → Predict effect on health/disease

We need to understand the functional mechanisms of pathological mutations at molecular level

Applications in Biomedicine: Prediction and interpretation of pathological mutations

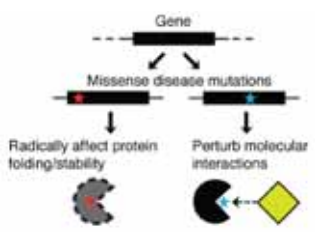
Effect of pathological mutations on PPIs and networks

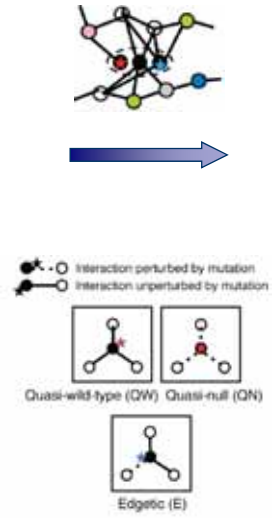




Applications in Biomedicine: Prediction and interpretation of pathological mutations

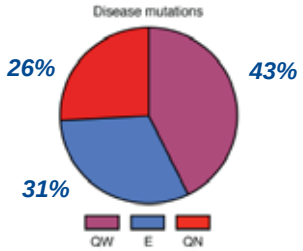
Effect of pathological mutations on PPIs and networks





n=197 mutations

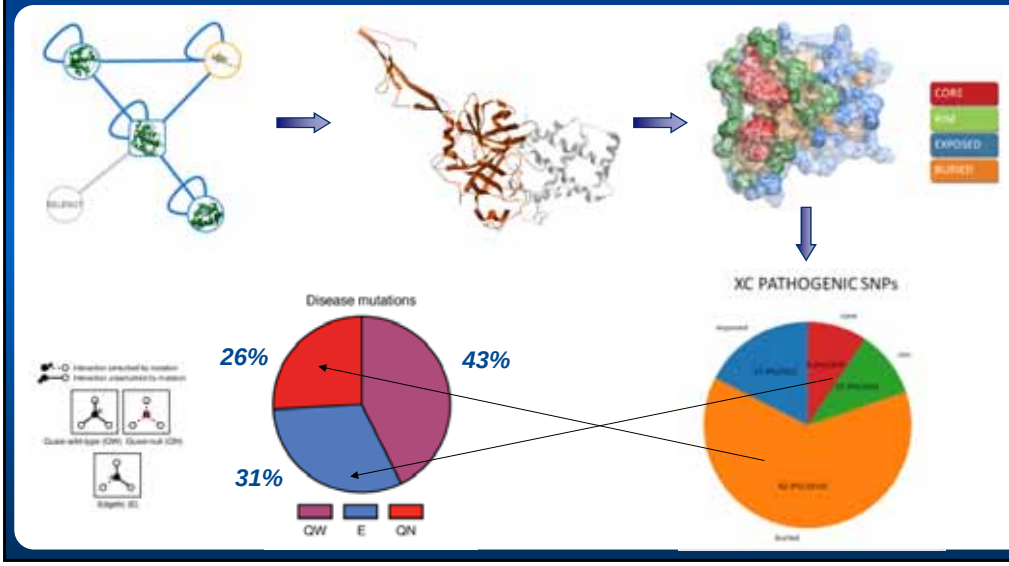
89 proteins with > 1 interaction



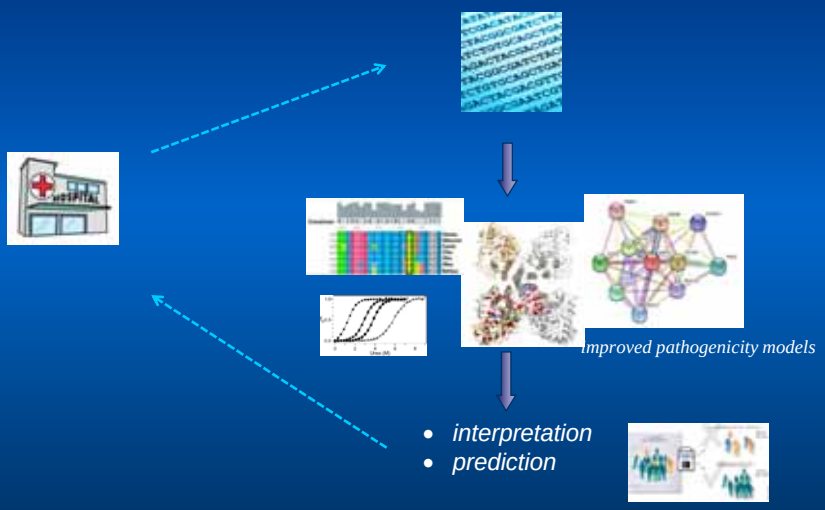
Sahni et al. (2015) Cell 161, 647

Applications in Biomedicine: Prediction and interpretation of pathological mutations

Effect of mutations on PPIs - structural and computational analysis



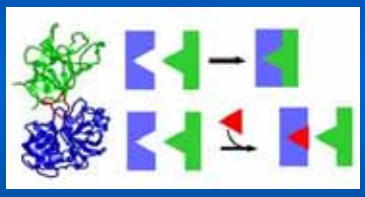
Applications in Biomedicine: Prediction and interpretation of pathological mutations

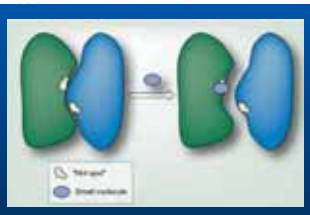


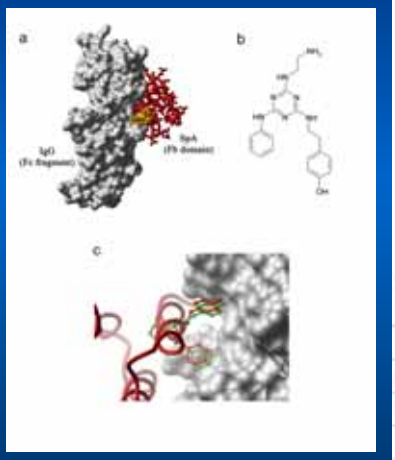
Applications in Biomedicine:
Drug discovery targeting protein-protein interactions

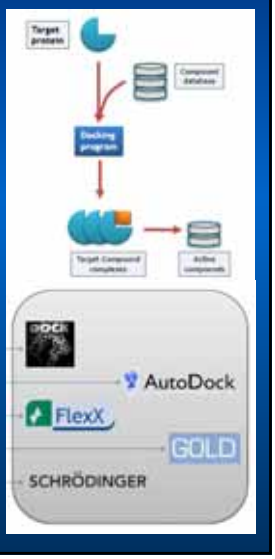
Structure-based drug design, PPI inhibitors

Inhibition of protein-protein interactions





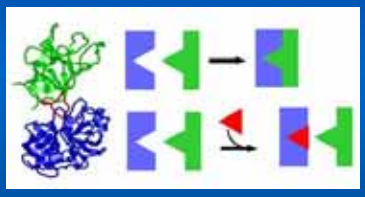


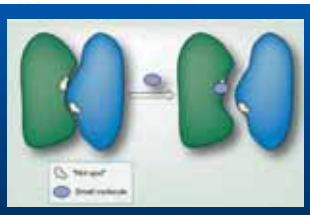


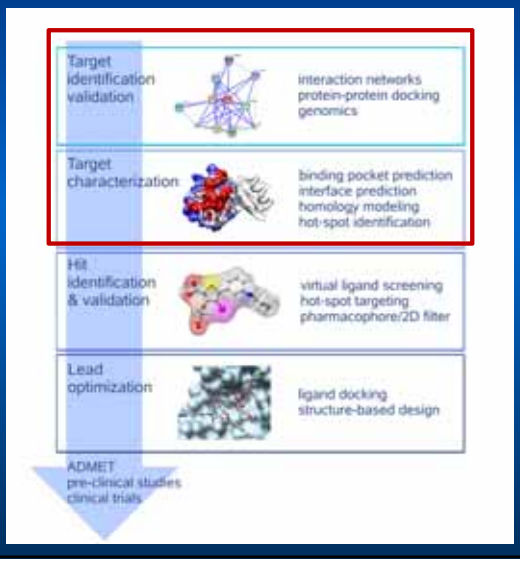
Applications in Biomedicine:
Drug discovery targeting protein-protein interactions

Structure-based drug design, PPI inhibitors

Inhibition of protein-protein interactions







Applications in Biomedicine: Understanding SNPs involving protein-protein interactions

Pathological mutations, population variation in drug response

SNPs in PPIs, pathological mutations, specificity in homologous interactions

Understanding drug response

speficitity, drug response

homologous interactions

SNPs