



# Three-dimensional modelling of protein complexes

Allegra Via



Consiglio Nazionale delle Ricerche



<https://goo.gl/forms/9CowK2OOTVxVCE7Q2>

# How PPI surfaces can be studied?

## Experimental:

- The most significant contribution to understanding the PPI surface comes from structural biology via **X-ray crystallography** or **NMR** as well as **mutational studies**

## Computational:

- Prediction of interaction/binding sites
- Prediction of protein-protein complexes
- *In silico* mutational studies

# Computational alanine-scanning in protein-protein interfaces

**DrugScore PPI**

*In silico* alanine scanning

Email  optional (see notes below)

PDB-ID  OR  [Sfoglia...](#) (\*)

Chain(s)  (comma separated)

 [Can't read?](#)

Via submit, you have read and accepted the [terms and conditions](#).

© Dennis M. Krüger, 2009

 ROBETTA BETA  
Full-chain Protein Structure Prediction Server

---

Structure Prediction Fragment Libraries Alanine Scanning DNA Interface Scan  
[ Queue ] [ Submit ]  
[ Register / Update ] [ Docs / FAQs ] [ Login ]

**Submit a job to the Computational Interface Alanine Scanning Server**

Please do not submit more than 10 targets at a time

**Required**

[Registered Username:](#) or [Registered Email Address:](#)

Job Name:

**Warning! Complex will be publicly accessible.**

Upload Complex:  [Sfoglia...](#)

<http://robetta.bakerlab.org/alascansubmit.jsp>

<http://cpclab.uni-duesseldorf.de/dsppi/>

**DrugScorePPI** and **Robetta Alascanning** are dedicated to the identification of hot spots at protein-protein interfaces

<http://prism.ccbb.ku.edu.tr/hotpoint/>

The screenshot shows the HotPOINT web interface. At the top left is the logo 'HotPOINT' with a red and orange geometric pattern, followed by 'HOT SPOT PREDICTION SERVER FOR PROTEIN INTERFACES'. To the right is the KOC UNIVERSITY logo. A vertical 'MENU' bar on the left includes links for Home, Documents, Links, and About. Below the menu is a citation: 'Tuncbag N, Gursoy A, Keskin O. Identification of computational hotspots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy. *Bioinformatics*, 2009 Jun 15;25(12):1513-20. [\[Link\]](#)'. A note below says: 'Below you can try our prediction algorithm by entering the four letter PDB code of a protein or uploading your own structure file that is in the PDB format with the chain identifiers. Please do not submit PDB files containing only one chain. This will return an error! Hotpoint requires two chain identifiers which corresponds to a protein interface.' There are three input fields: 'Run our prediction algorithm for a particular input protein.', 'Enter the four letter PDB code for automatic download from PDB: ', and 'Or load your structure file from disk:  Stopfile '. Below these is a Jmol visualization window titled 'Interactive Jmol visualization for 1AXDAB'. The visualization shows two protein chains, A and B, represented by blue and yellow sticks. The interface residues for chain A are highlighted in cyan, and the interface residues for chain B are highlighted in yellow. Red spheres represent predicted hot spots at the interface.

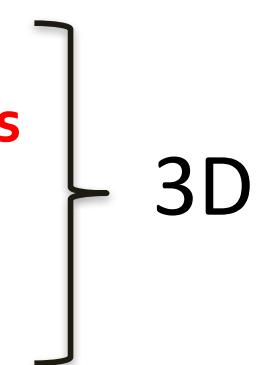
**HotPOINT** calculates solvent accessibilities and total contact potentials. It predicts hot spots at the complex interface

**PIC** recognises various kinds of interactions, such as disulphide bonds, hydrophobic interactions, ionic interactions, hydrogen bonds, aromatic- aromatic interactions, aromatic-sulphur interactions and cation- $\pi$  interactions between proteins in a complex

<http://crick.mbu.iisc.ernet.in/~PIC/>

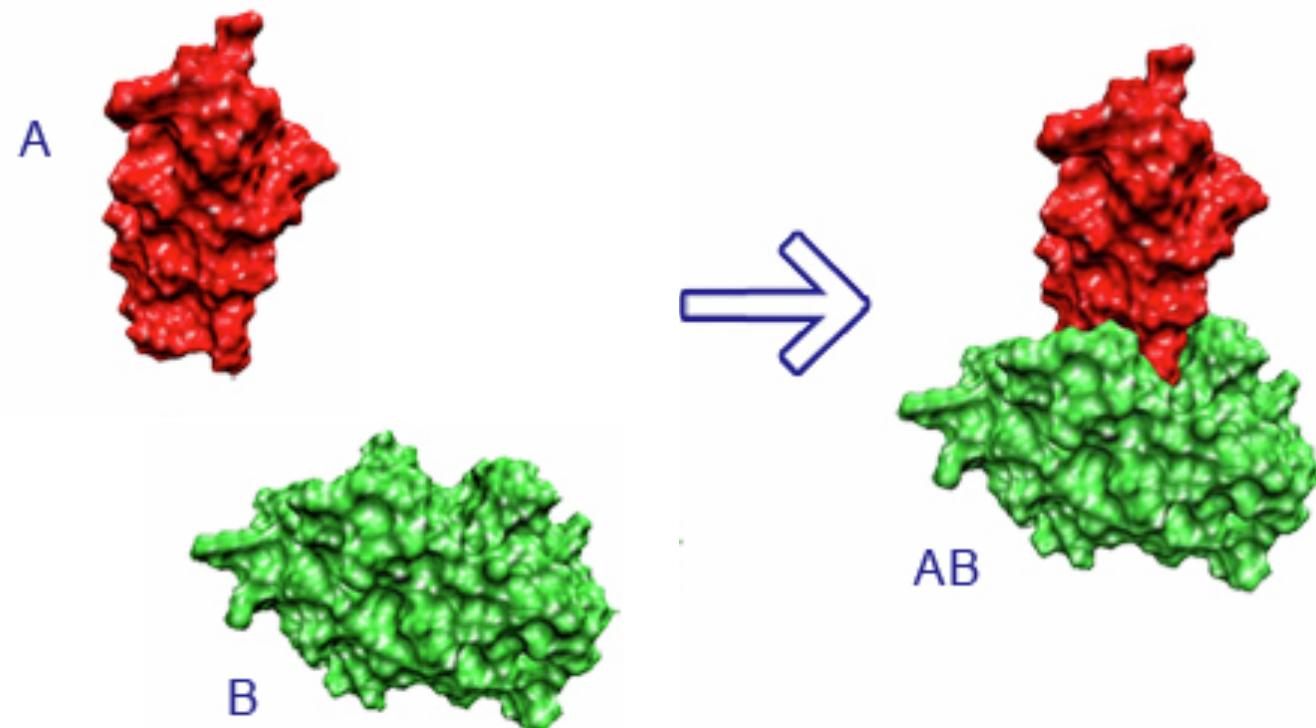
The screenshot shows the PIC web interface. At the top center is the title 'P I C : Protein Interactions Calculator'. Below it is the text 'Molecular Biophysics Unit, Indian Institute of Science, Bangalore.' A navigation bar at the bottom has links for HOME, HELP, CRITERIA, SUBMIT JOB, CONTACT US, and LAB PAGE. A message in the middle says: 'You can also submit your jobs on the [new link](#).'. Below that is a section titled 'INTRODUCTION:' with the following text: 'Protein Interactions Calculator (PIC) is a server which recognizes various kinds of interactions; such as disulphide bonds, hydrophobic interactions, ionic interactions, hydrogen bonds, aromatic- aromatic interactions, aromatic-sulphur interactions and cation -  $n$  interactions within a protein or between proteins in a complex. It also determines the accessible surface area as well as the distance of a residue from the surface of the protein. The input should be in the Protein data bank(.pdb) format. Interactions are calculated based on empirical or semi-empirical set of rules.'

# Prediction of protein-protein interactions

- Phylogenetic profiling
  - Prediction of co-evolved protein pairs based on similar phylogenetic trees
  - Rosetta stone method
  - Association methods
  - Bayesian network modelling
  - Domain-pair exclusion analysis
  - Supervised learning problem
  - Gene fusion
  - **Classification methods**
  - **Inference of interactions from homologous structures**
  - **Identification of structural patterns (hot spots)**
  - **Protein-protein docking**
- 
- 3D

# What is molecular docking?

- Computational schemes that attempt to find the “best” matching between two molecules, a **receptor** and a **ligand**
- The *molecular docking problem* can be defined as follows: **given the atomic coordinates of two molecules, predict their “correct” bound association**



# Protein-protein docking: modelling the structure of protein complexes

- Used to model the quaternary structure of complexes formed by two or more interacting proteins
- It is the “gold standard” for prediction of PPIs
- It makes it possible to figure out if two proteins interact and also how the interaction takes place (“**mode**” of binding)
- It is computationally very challenging and thus very unlikely to be applied for high throughput purposes

# What types of questions can docking address?

- Do proteins A (receptor) and B (ligand) bind *in vivo*?

If they do bind:

- What is the spatial configuration they adopt in their bound state?
- What is the structure of the protein complex (**near-native structure**) in atomic details ?
- How strong or weak is their interaction (which types of interactions are present)?
- What is the orientation that maximises the interaction, minimising the energy of the complex?

If they don't bind:

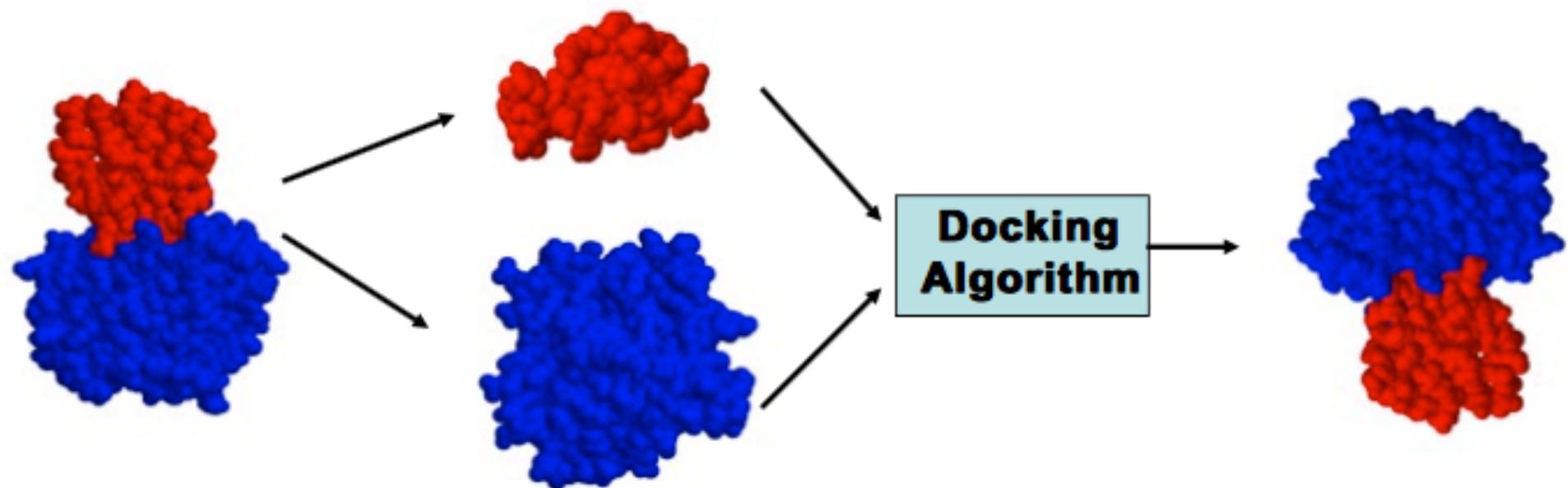
- Would they bind if there was a mutation?

# Docking prerequisites

- The molecular structure (spatial coordinates) of the interacting partners
  - has been determined experimentally
  - or
  - can be "safely" predicted
- Computing power

# "Bound" docking

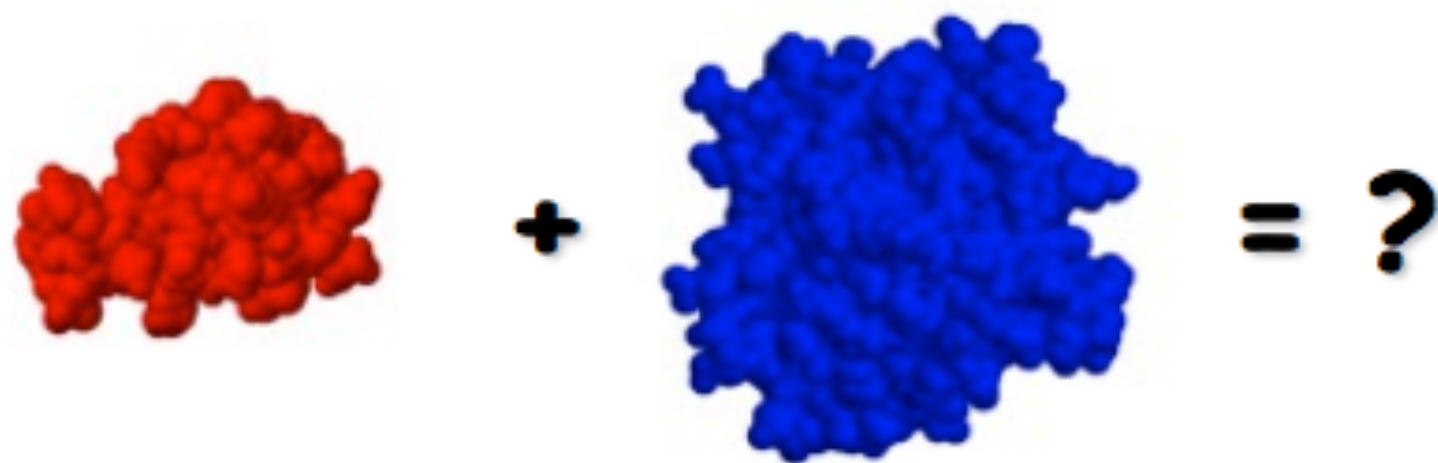
- Schemes that attempt to reconstruct a complex using the bound structures of the receptor and the ligand
- After artificial separation of the receptor and the ligand, the goal is to reconstruct the native complex



- No conformational changes are involved
- **Used as first test to validate the algorithm**

# "Unbound" or "predictive" docking

- Schemes that attempt to reconstruct a complex using the unbound structures of the receptor and the ligand
- An "unbound" structure maybe a **native** structure, a **pseudo-native** structure, or a **modelled** structure
- **Native**: free in solution, in its uncomplexed state
- **Pseudo-native**: structure complexed with a molecule different from the one used for the docking

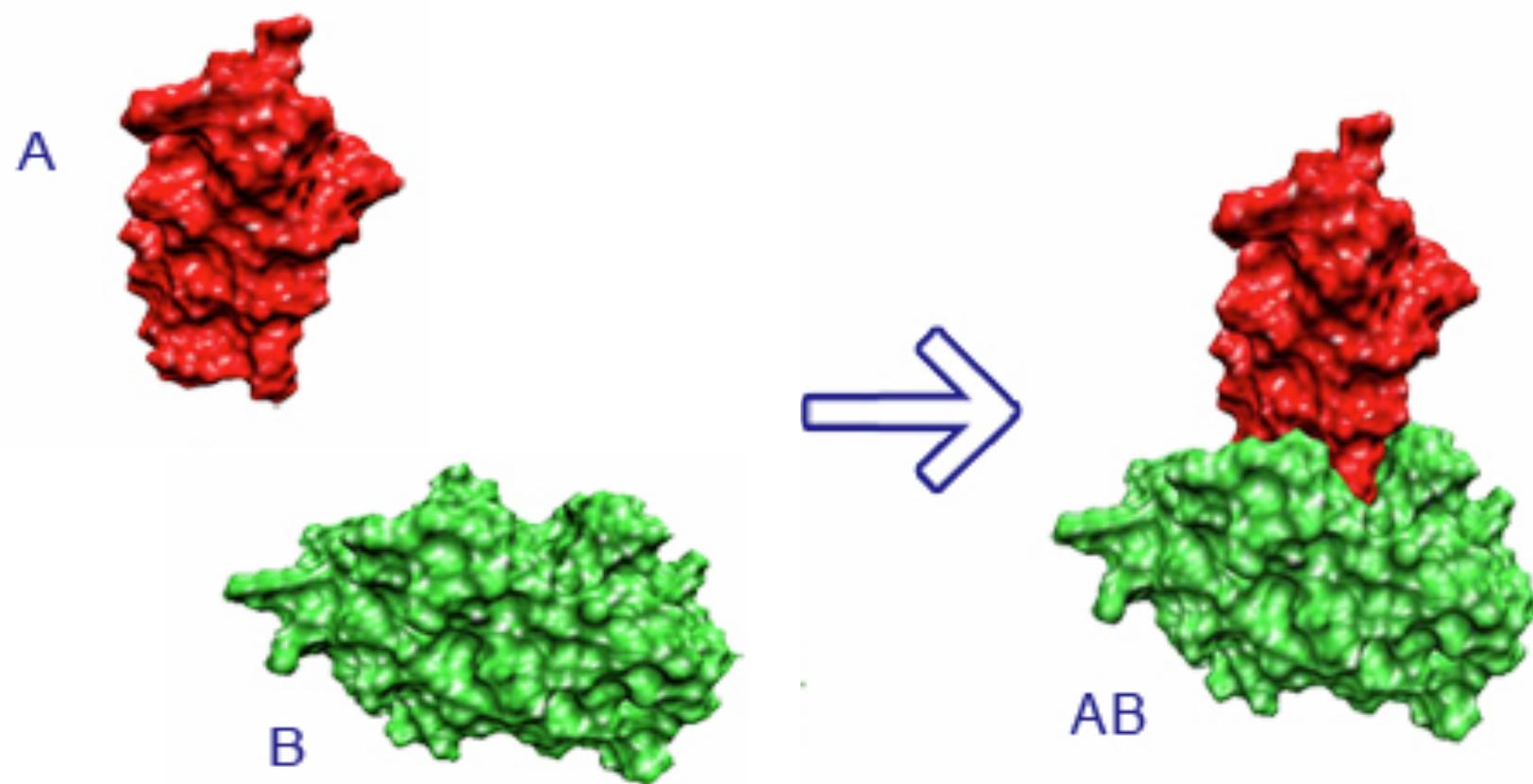


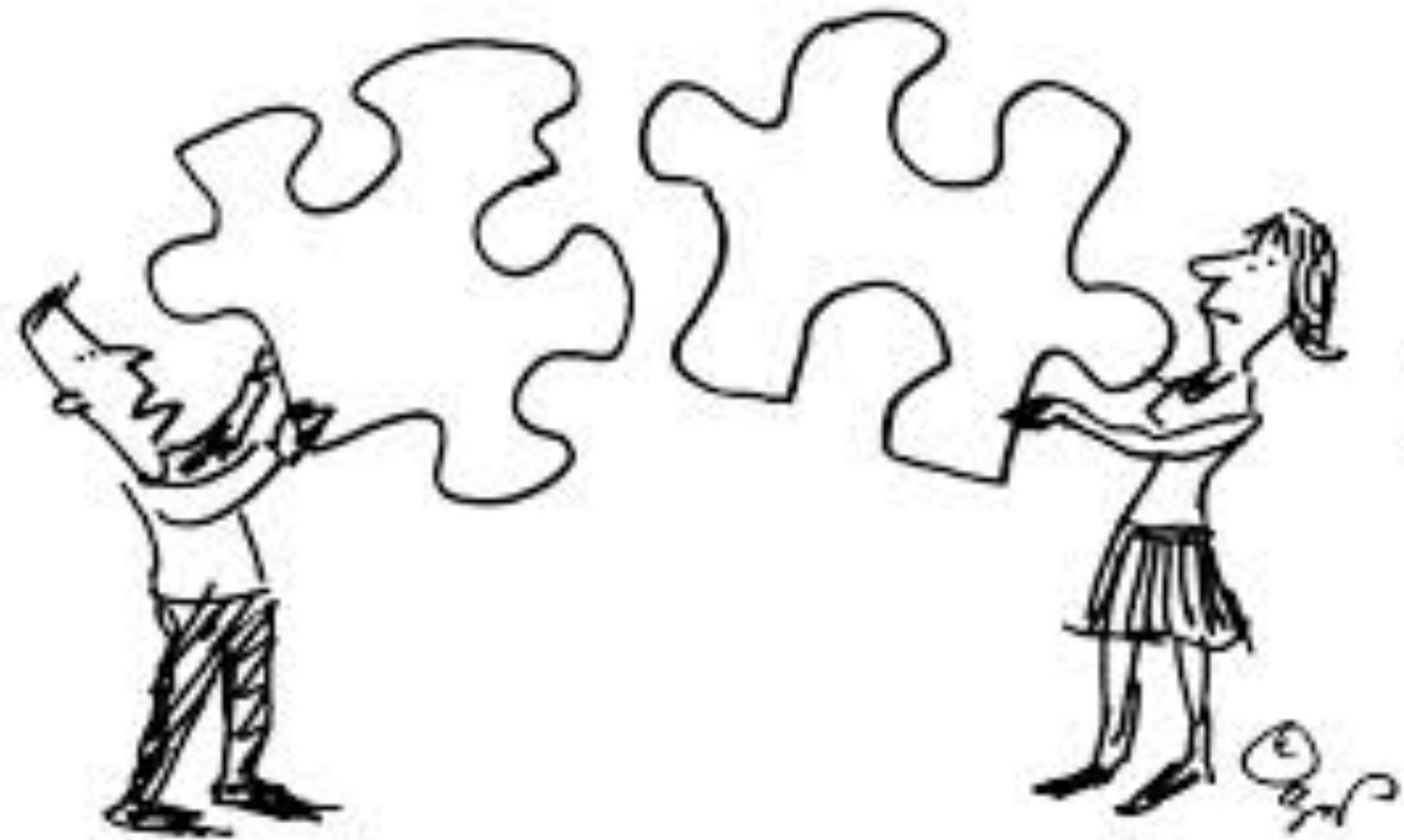
Halperin et al. Proteins, 2002

# Unbound docking

- It is far more complex than bound docking
- Problems: conformational changes (side-chains and backbone movements), experimental errors in the structures, reliability of models.

# The three key ingredients in docking





# The three key ingredients in docking

Representation of the system

Conformational space search

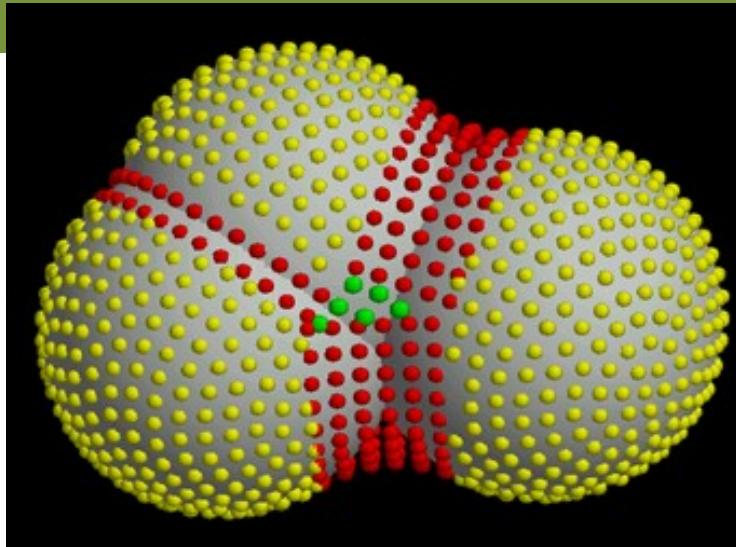
Ranking of potential solutions

# Representation of the system

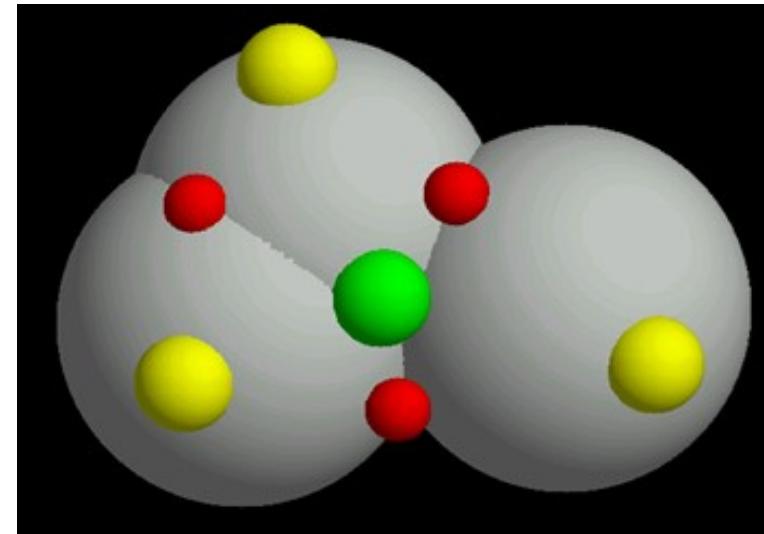
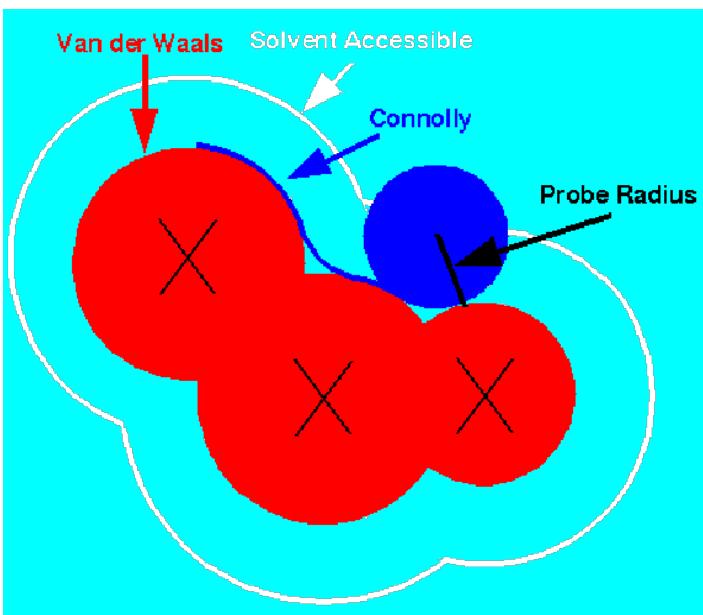
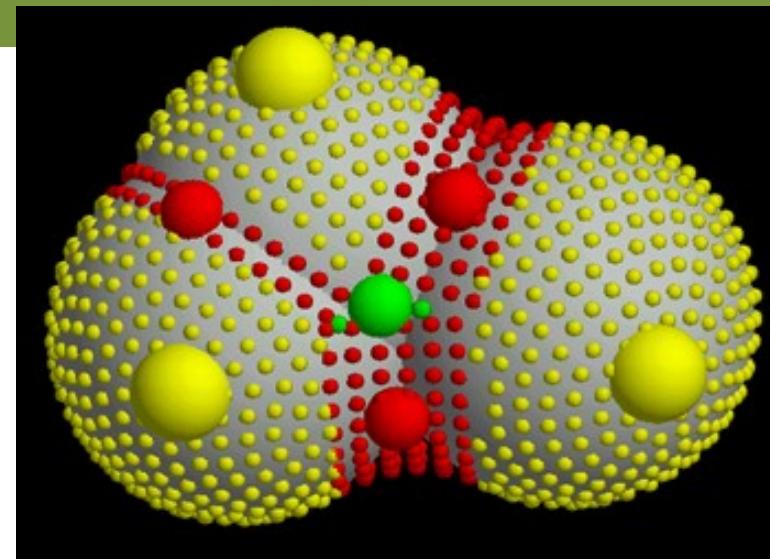
- Docking essentially simulates the interaction of the protein surface
- How do we define a protein surface?
  - Mathematical models (e.g. geometrical shape descriptors, a grid)
  - Static or dynamic treatment of the protein frame (rigid vs flexible)
- **The choice of the system (surface) representation decides the types of conformational search algorithms, and the ways to rank potential solutions**

# Surface representation

Dense surface (Connolly)

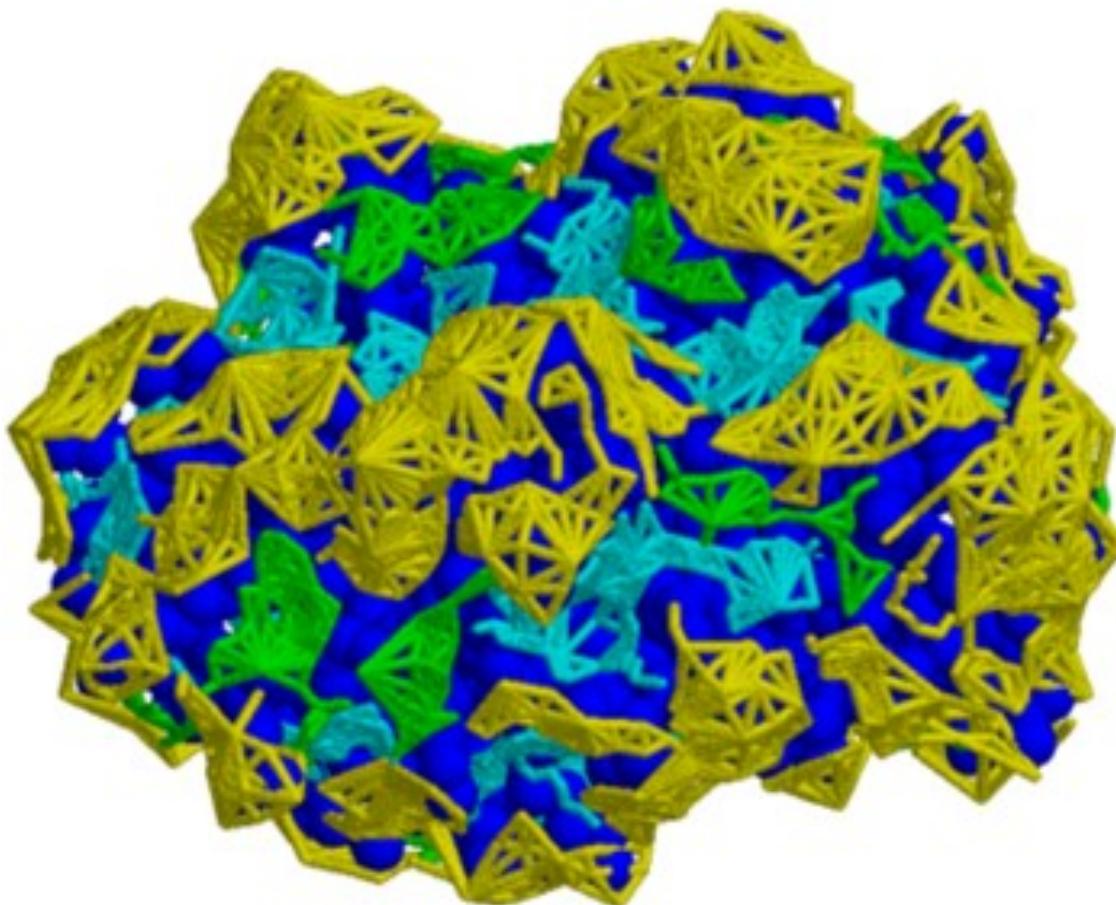


Sparse surface (Shuo Lin et al.)



# Patch detection

- Divide the surface into connected, non-intersecting, equal sized patches of critical points with similar curvature



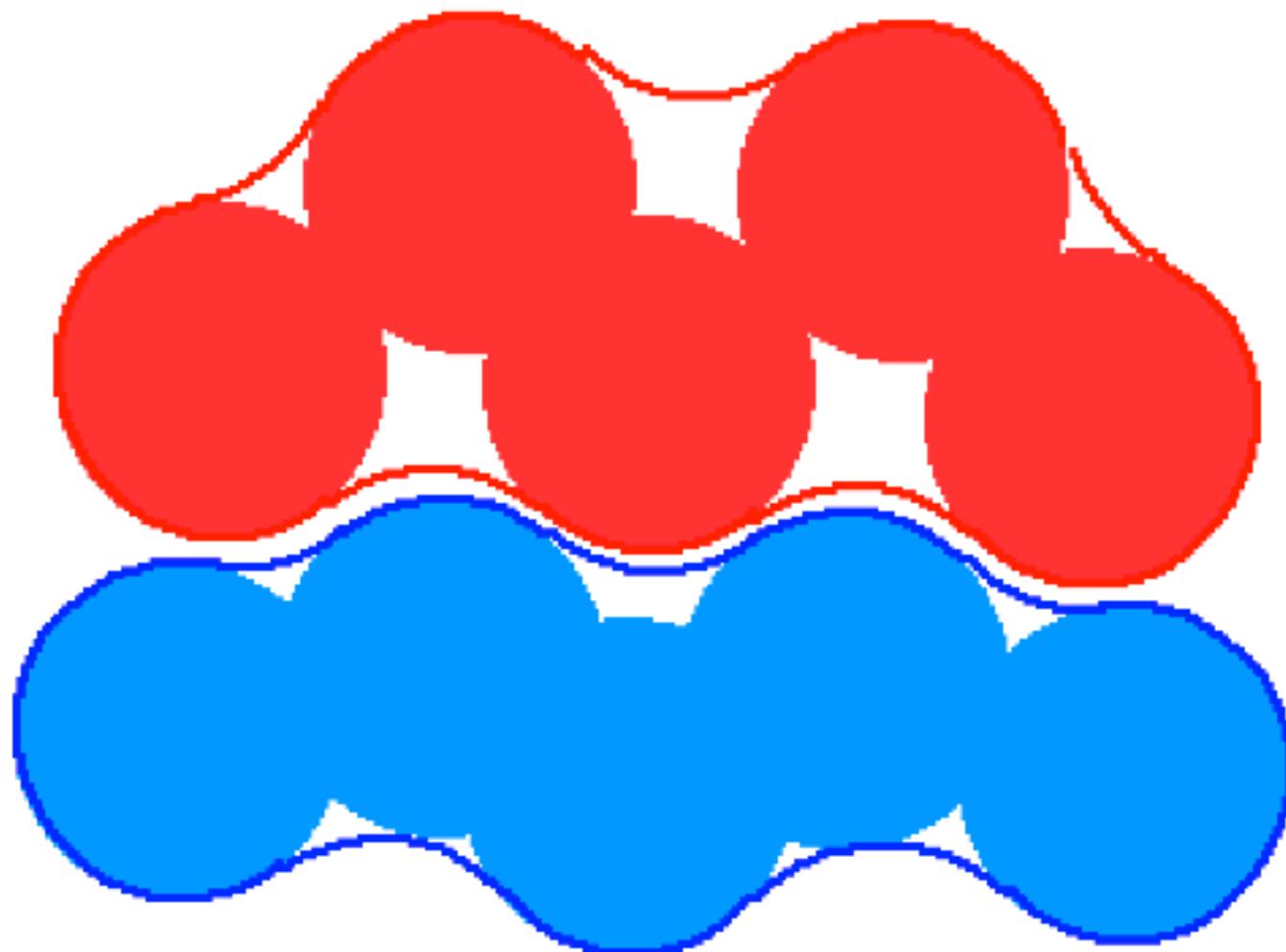
**Yellow:** knob patches  
**Cyan:** hole patches  
**Green:** flat patches  
**Blue:** protein

# Forces governing biomolecular recognition

- Van der Waals
- Electrostatics
- Hydrophobic contacts
- Hydrogen bonds
- Salt bridges

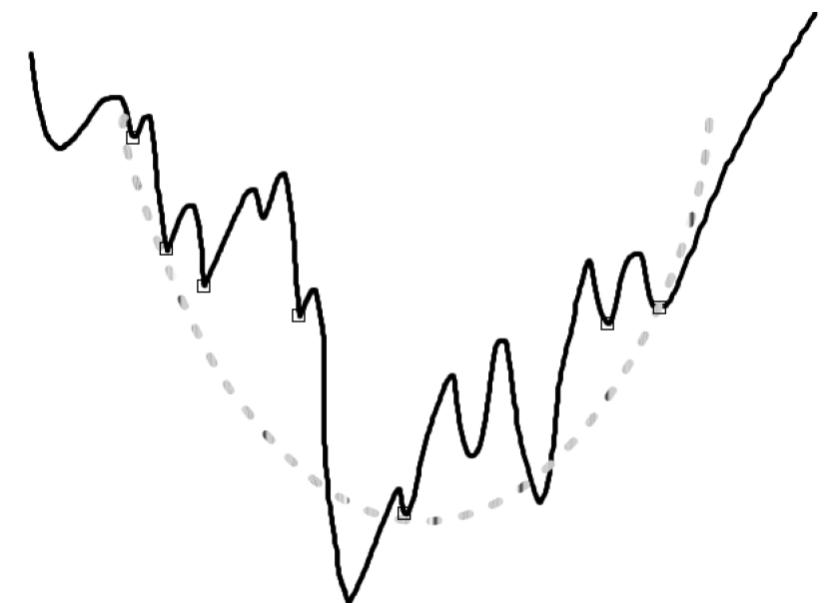
All interactions act at short range → surface complementarity is needed for tight binding

# Shape complementarity



# Conformational space search

- Efficient search algorithm
- Speed and effectiveness in covering the relevant conformational space
- Computationally difficult - there are many ways to put two molecules together (3 translational + 3 rotational degrees of freedom)
- Goal: locate the most stable state (global minimum) in the energy landscape



# Rigid vs flexible docking

- **Rigid body** is a highly simplistic model that regards the two proteins as two rigid solid bodies
- The **semi-flexible** model is asymmetric; one of the molecules is considered flexible, while the receptor is regarded as rigid
- **Flexible** docking. Both molecules are considered flexible, though flexibility is limited or simplified

# Docking algorithms

- Rigid docking
  - fast → can explore the entire receptor and ligand surfaces
  - Less accurate
  - flexibility = "soft" belt into which atoms can penetrate
- Flexible docking
  - Slower
  - More accurate
  - Can model side-chain/backbone flexibility
  - highly reliable but too slow for extensive ligand docking

# Docking approaches

- 1) scan of the entire solution space in a predefined systematic manner

e.g., complete searches of all orientations between two rigid molecules by systematically rotating and translating one molecule about the other
- 2) a gradual guided progression through solution space. Only part of the solution space is searched, or fitting solutions\* are generated.

e.g., Monte Carlo, simulated annealing, molecular dynamics (MD), and evolutionary algorithms.
- 3) Data-driven docking – it uses the available information about binding site/interface residues

\* Solutions meeting pre-defined criteria

# Docking approaches

**Although each method is optimal for a specific class of protein docking problems, combining computational steps from different methods can improve the reliability and accuracy of results.**

- **Global methods** work well when none of the interacting partners undergo conformational changes larger than 2 Å. In particular, when no prior information on the complex is available, requiring the search of the entire conformational space.
- **Monte Carlo** methods yield good results when side-chain repacking is crucial, e.g., when one of the protein structures is a homology model.
- **Data-driven methods** produce the best results if sufficient information on interface residues is available, even when the binding causes large conformational change, including the backbone.

# Ranking of potential solutions

- A search algorithm may produce an immense number of solutions ( $\sim 10^9$ )
- Goal: discriminate between "correct" native solutions, i.e., with **low rmsd from the crystal structure** and others within reasonable computation time
- good scoring function → fast enough to allow its application to a large number of potential solutions
  - effectively discriminates between native and non-native docked conformations
  - should include and appropriately weight all the energetic ingredients

To solve the docking problem, ideally, the best matching algorithms and scoring schemes should be combined

Two-stage ranking:

1. fast scoring to rapidly scan possible solutions and obtain initial "good" candidates – **mostly geometric criteria**
2. Followed by more advanced methods to further discriminate the limited conformations - **energy criteria**

# Algorithms can be classified by the stage of scoring in the algorithm flow

- **Integrated functions** – scoring is integrated into the search stage and filter emerging solutions – the scoring function is part of the design
- **Edge functions** – scoring is applied at the end of the search stage

# Parameters used for scoring

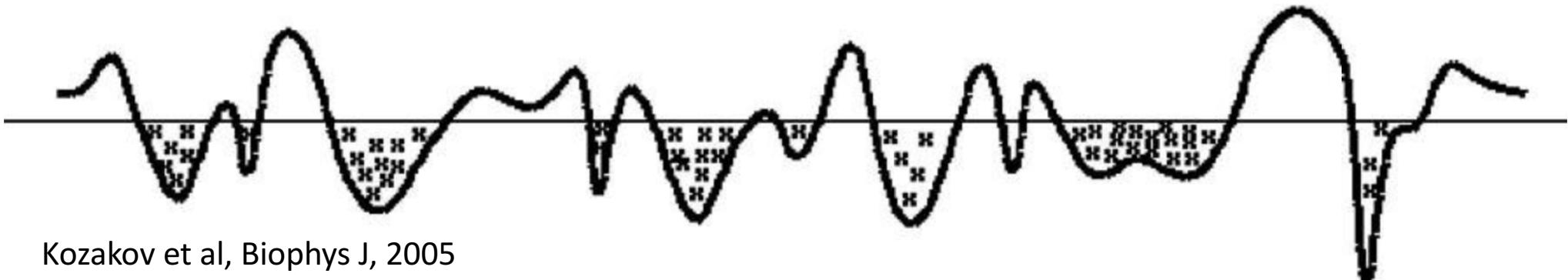
- **Geometric complementarity** - how to score complementarity is strongly coupled with the surface representation
- **Intermolecular overlap** – tolerance to slight interface clashes and penalty for protein interior clashes (surface "belt" of non-penalised penetration area)
- **Intra-molecular overlap** – when backbone flexibility is taken into account
- **Hydrogen bonding**
- **Contact area:** total interactions = hh + pp + hp (h = hydrophobic, p = polar)
- **Pairwise aa and atom-atom contacts** – empirical term derived from observed statistical frequency of aa contacts in X-ray proteins
- **Electrostatic interactions and solvation energy**

# Binding site information may be included in scoring

- Knowledge of the **location of the binding site** on one or both proteins drastically reduces the number of possible solutions
- Knowledge of the **specific binding site residues** reduces the search space even further
- Info about active site residues: site directed mutagenesis, chemical cross-linking, phylogenetic data
- Sometimes the binding site **can be predicted**
- For some families the major binding sites are known in advance (e.g. serine proteases and immunoglobulins)

# Clustering of solutions

- **Events that occur in clusters are probably not random**
- The cluster with the largest number of low-energy structures is typically the native fold, the center of the most populated cluster being a structure near the native binding site
- Looking for large clusters is a major tool of finding near-native conformations



# Conclusion I (-)

- The *molecular docking problem* is far from being solved
- It is difficult to find very specific properties of protein-protein interfaces
- Results are generally poor with weakly interacting proteins
- Proteins are flexible and may undergo even large conformational changes upon binding
- Exhaustive space searches provide too many conformations
- Accurate interaction energies are too complicated to compute

# Conclusion II (-)

- Despite some algorithms are able to rank correct solutions within the top ten predictions, for most complexes the highest ranked structures are still false positives (high rmsd from the complex)
- No efficient method for reliable discrimination between correct solutions and FPs is currently available, in particular if the binding site is unknown
- Many FPs displaying good surface complementarity are far from the native complex

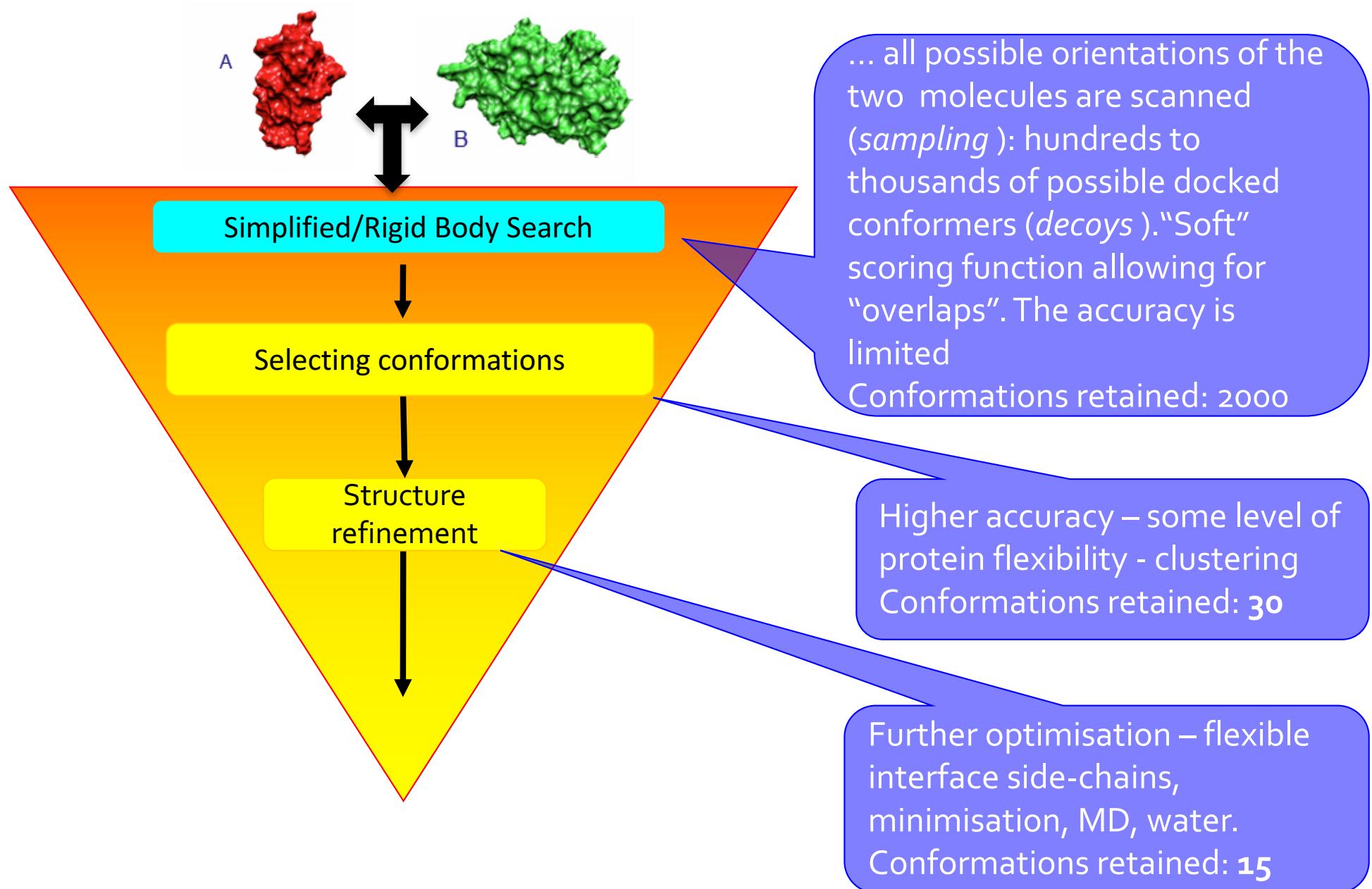
# Conclusion I (+)

- If the conformational change is limited to surface side-chain atoms, rigid body algorithms have been remarkably successful, even in absence of knowledge of the binding site
- Side-chain flexibility can be handled via a "soft" tolerance belt

# Conclusion II (+)

- "Docking in steps" is a promising strategy: initial rigid-body, entire surface algorithm followed by a dynamic method overcoming energy barriers
- Integration of experimental information produces reliable docking results
- Relatively easy for enzyme-inhibitor complexes
- Sometimes good results with antigen-antibody pairs

# How current protein docking programs work?



# CAPRI: Critical Assessment of PRotein Interactions

Home > Databases > PDBe > Services > Capri-Home

**CAPRI: Critical Assessment of PRediction of Interactions**

CAPRI communitywide experiment on the comparative evaluation of protein-protein docking for structure prediction

Hosted By EMBL/EBI-PDBe Group

PDB idcodes for past targets

<http://www.ebi.ac.uk/msd-srv/capri/>

- CAPRI is a community-wide experiment in modelling the molecular structure of protein complexes
- CAPRI is a **blind prediction experiment** aimed at testing the performance of protein docking methods
- Rounds take place about every six months
- Each round contains between one and six target protein–protein complexes whose structures have been recently determined experimentally
- Targets are unpublished crystal or NMR structures of complexes, whose coordinates are held privately by the assessors, with the co-operation of the structural biologists who determined them
- The atomic coordinates of the two proteins are given to groups for prediction

# CAPRI: Critical Assessment of PRotein Interactions

- The CAPRI experiment is double-blind
  - submitters do not know the solved structure
  - the assessors do not know the correspondence between a submission and the identity of its creator
- International meetings
- *Proteins: Structure, Function, and Bioinformatics*

**Special CAPRI Issue:** Fifth Meeting on the Critical Assessment of Predicted Interactions

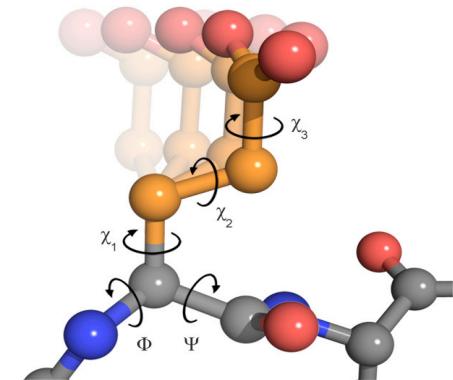


December 2013

# CAPRI drives the community to develop new techniques

## Side-chain flexibility

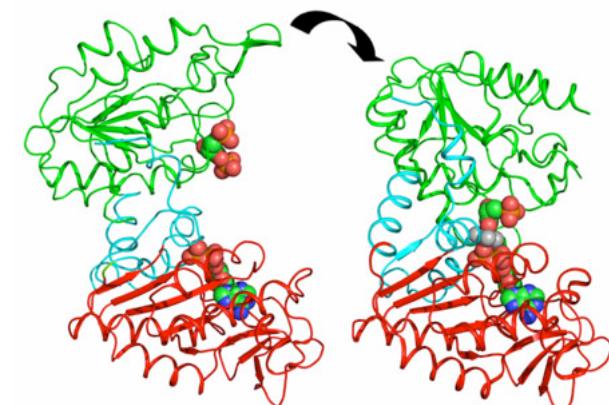
First rounds the big challenge was to model correctly side-chain conformations



Dihedral angles in glutamate

## Domain movements

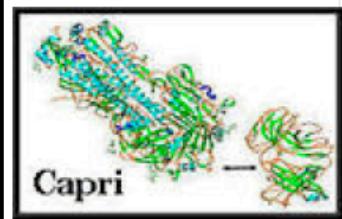
Today the development is focused on backbone flexibility



Domain movements in PGK catalysis.

J. Biol. Chem. (2011) 286, 14040-14048

The clustering application to protein-protein docking was first introduced in CAPRI by Camacho & Gatchell,D. (2003)



Home > Databases > PDBe > Services > Capri-Home

## CAPRI: Critical Assessment of PRediction of Interactions

PDB idcodes for past targets

CAPRI communitywide experiment on the comparative evaluation of protein-protein docking for structure prediction

Hosted By EMBL/EBI-PDBe Group

CAPRI experiment ... who is the winner ??



There is no official ranking and in many cases the differences between the different algorithms are not huge

# CAPRI experiment ... who is the winner ??

- **HADDOCK** (software/web server).  
<http://haddock.chem.uu.nl>
- **CLUSPRO** (software/web server)  
<http://cluspro.bu.edu>
- **ICM-pro** (desktop-modeling environment)  
[http://www.molsoft.com/protein\\_protein\\_docking.html](http://www.molsoft.com/protein_protein_docking.html)
- **ROSETTADOCK** (software/web server)  
<http://graylab.jhu.edu/docking/rosetta/>
  - <http://rosettadock.graylab.jhu.edu/submit>
- **GRAMM-X** (web server)  
<http://vakser.bioinformatics.ku.edu/resources/gramm/grammx>
- **PATCHDOCK/FIREDOCK** (software/web server)  
<http://bioinfo3d.cs.tau.ac.il/PatchDock/>
- **HEX** (software/web server)  
<http://hexserver.loria.fr>

HADDOCK webserver

haddock.org

Reader

PubMed dizionari databases BTN News METEO Bandi Sapienza biocomp ISI U-GOV

home

# HADDOCK

## Software web portal

10111  
01001010010010  
10001010010101001011  
0100100100010101001010  
1010010010101101010001010111  
010101001010010010101001010010

Home HADDOCK Whiscy CPoRT DNA SQUEEZE Publications HADDOCK Inc. Contact RSS

WELCOME TO THE UTRECHT BIOMOLECULAR INTERACTION WEB PORTAL >>

The Utrecht Biomolecular Interactions software portal provides access to software tools developed in the Computational Structural Biology group / NMR Research Group of Utrecht University with a main focus on the characterization of biomolecular interactions. Please note that this site is in active development.

### HADDOCK WEB DOCKING



HADDOCK (High Ambiguity Driven protein-protein DOCKing) is an information-driven flexible docking approach for the modeling of biomolecular complexes. HADDOCK distinguishes itself from ab-initio docking methods in the fact that it encodes information from identified or predicted protein interfaces in ambiguous interaction restraints (AIRs) to drive the docking process. HADDOCK can deal with a large class of modelling problems including protein-protein, protein-nucleic acids and protein-ligand complexes. | [GO TO SERVICE >>](#)

**Note:** The default HADDOCK server is now version 2.2. The previous version (2.1) is still available [here](#) (and its grid-enabled version [here](#).)

**Note:** The regular HADDOCK server might be busy at times and/or closed to users

PROFILE >>



Universiteit Utrecht



# The HADDOCK protocol

- Rigid-body docking followed by
- a semi-flexible refinement of the interface in torsion angles' space (of both backbone and side-chains)
- as a final stage, a Cartesian dynamics refinement in explicit solvent

# The HADDOCK services

- the Easy interface
- the Prediction interface
- the Expert interface (requires Expert level access)
- the Refinement interface (requires Expert level access)
- the Guru interface (requires Guru level access)
- the Multi-body interface (requires Guru level access)
- the File upload interface

haddock.scientific.univie.ac.at/HADDOCK2.2/haddock/easy.html

The EASY interface

WELCOME TO THE UTRECHT BIOMOLECULAR INTERACTION WEB PORTAL >>

This is the easy interface to the HADDOCK docking program. Please define the structure for each molecule you want to dock as well as the residues belonging to the interaction interface.

Docking is performed with default settings that work well for average complexes. If you do not have any special wishes for the system you want to have docked, this is the way to go.

Unfold the menus by clicking on the double arrows. Submit your job by providing your username and password and press submit.

*You may supply a name for your docking run (one word)*

Name

**First molecule** 

**Second molecule** 

**Structure definition**

Where is the structure provided?

Which chain of the structure must be used?

PDB structure to submit  no file selected

or: PDB code to download

**Restraint definition**

*Data to drive the docking*

*Please supply residues as comma-separated lists of residue numbers*

Active residues (directly involved in the interaction)

Passive residues (surrounding surface residues)

Define passive residues automatically around the active residues

What kind of molecule are you docking?

The EXPERT interface

home >> HADDOCK2.2 >>

# HADDOCK2.2

## Software web portal

WELCOME TO THE Utrecht BIOMOLECULAR INTERACTION WEB PORTAL >>

This is the expert interface to the HADDOCK docking program. This interface provides more control over HADDOCK parameters and supports additional types of restraints.

Unfold the menus by clicking on the double arrows. Submit your job by providing your username and password and press submit.

*You may supply a name for your docking run (one word)*

Name

**First molecule** 

**Second molecule** 

**Distance restraints** 

**Sampling parameters** 

**Parameters for clustering** 

**Dihedral and hydrogen bond restraints** 

dock.science.uilu.nl/services/HADDOCK2.2/haddockserver\_guru.html

Reader

comp ISI U-GOV

First molecule

Second molecule

Distance restraints

Sampling parameters

Parameters for clustering

Dihedral and hydrogen bond restraints

Noncrystallographic symmetry restraints

Symmetry restraints

Restraints energy constants

Residual dipolar couplings

Pseudo Contact Shifts restraints

Relaxation anisotropy restraints

Energy and interaction parameters

Molecule interaction matrix

Scoring parameters

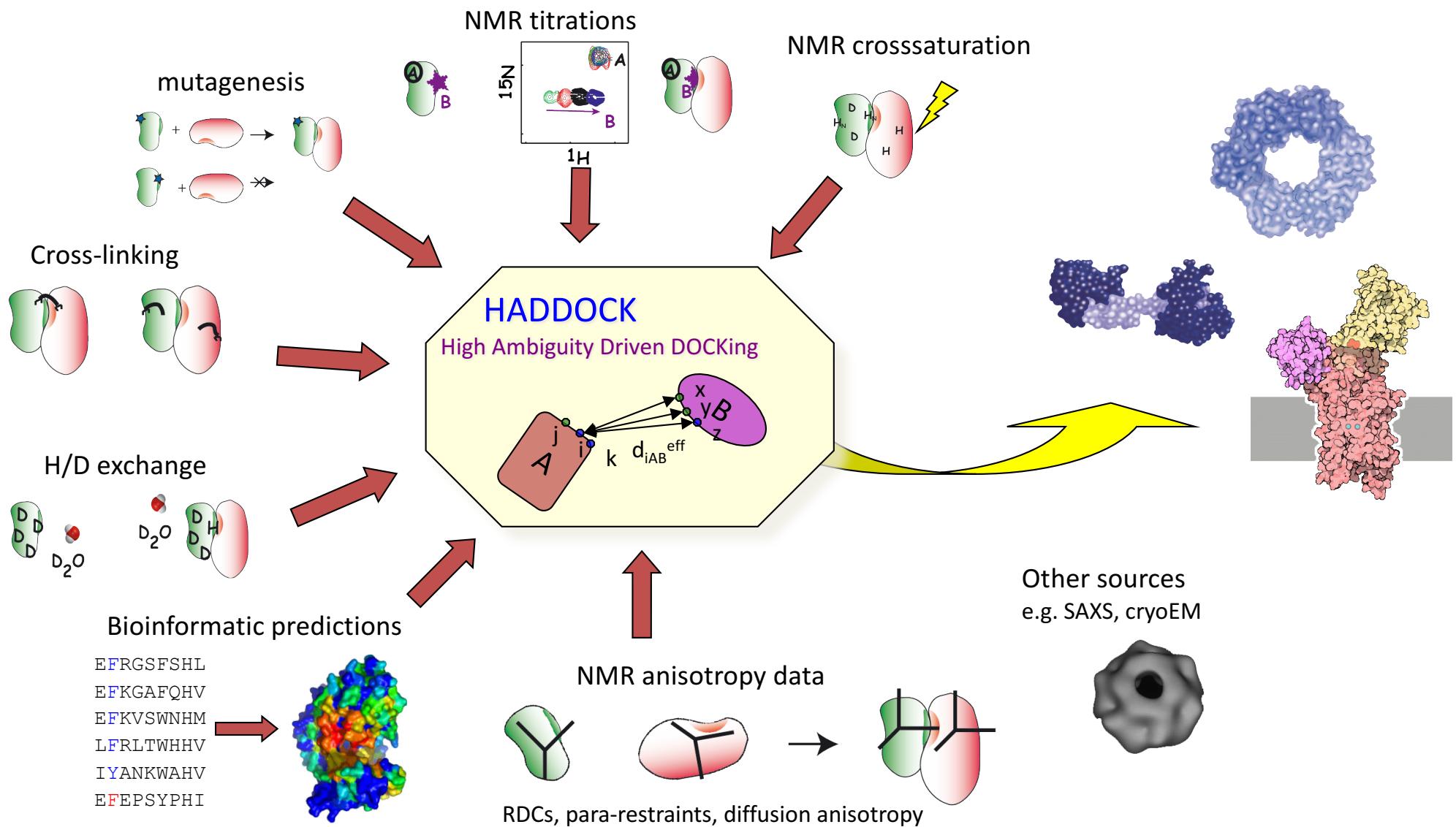
Advanced sampling parameters

Solvated docking parameters

Analysis parameters

Username and password

The screenshot shows a web-based docking parameter configuration interface. At the top, there's a header bar with standard browser controls (back, forward, search, etc.) and tabs for 'PubMed' and 'dizionari'. Below the header, the title 'The GURU interface' is prominently displayed. The main content area is a vertical list of docking parameters, each represented by a dark brown rectangular box containing a white text label and a small upward-pointing double-headed arrow icon on the right side. The parameters listed are: First molecule, Second molecule, Distance restraints, Sampling parameters, Parameters for clustering, Dihedral and hydrogen bond restraints, Noncrystallographic symmetry restraints, Symmetry restraints, Restraints energy constants, Residual dipolar couplings, Pseudo Contact Shifts restraints, Relaxation anisotropy restraints, Energy and interaction parameters, Molecule interaction matrix, Scoring parameters, Advanced sampling parameters, Solvated docking parameters, and Analysis parameters. At the very bottom of the page, there's a small instruction 'Username and password'.



# Parameters supported by the various HADDOCK web interfaces

de Vries et al., Nature Protocols, 2010

**TABLE 1** | Data-driven docking parameters supported by the various HADDOCK web interfaces.

	Easy interface	Expert interface	Guru/upload interface
Active and passive residue lists	Supported	Supported	Supported
Docking from ensemble structures	Supported	Supported	Supported
Protein–nucleic acid docking	Supported	Supported	Supported
Cofactors and modified amino acids <sup>a</sup>	Supported	Supported	Supported
Semi-flexible segment definition	Automatic	Automatic or manual	Automatic or manual
Fully flexible segment definition	No	Supported	Supported
Histidine protonation states	Automatic <sup>b</sup>	Automatic <sup>b</sup> or manual	Automatic <sup>b</sup> or manual
Custom CNS distance restraints	No	Supported	Supported
Custom hydrogen bond restraints	No	Supported	Supported
Custom dihedral angle restraints	No	Supported	Supported
Preservation of nucleic acid base pairing and backbone conformation	Automatic	Automatic or manual	Automatic or manual
<i>Ab initio</i> docking	No	Supported	Supported
Nonpolar hydrogens	No	Supported	Supported
Random removal of restraints	Automatic	Manual	Manual
Number of structures to dock and to refine	Automatic	Manual	Manual
DMSO refinement	No	Supported	Supported
Solvated docking	No	Supported, automatic parameters	Supported, manual parameters
Epsilon (electrostatic scaling constant)	Automatic	Manual	Manual
Clustering	Automatic	Manual	Manual
Symmetry	No	No	Supported
Restraints energy constants	Automatic	Automatic	Manual
Relaxation anisotropy restraints	No	No	Supported
Residual dipolar couplings	No	No	Supported
Energy and interaction parameters	Automatic	Automatic	Manual
Scoring parameters	Automatic	Automatic	Manual
Randomization of starting orientations	Automatic	Automatic	Manual
Refinement-only protocol	No	No	Supported
Temperature and timesteps of the various refinement stages	Automatic	Automatic	Manual

<sup>a</sup>Small molecule parameters are automatically retrieved from the PRODRG server<sup>35</sup>. <sup>b</sup>The most likely protonation states are automatically retrieved from the WHATIF server<sup>25</sup>.

# Example of HADDOCK output

In this case additional experimental information has to be taken into account to completely rule out alternatives to the highest-scoring solution

**CLUSTER 1**

HADDOCK score	-154.5 +/- 4.4
Cluster size	126

**CLUSTER 3**

HADDOCK score	-141.1 +/- 8.7
Cluster size	20
RMSD from the overall lowest-energy structure	8.9 +/- 0.1
Van der Waals energy	-41.0 +/- 5.4
Electrostatic energy	-554.1 +/- 26.6
Desolvation energy	3.6 +/- 5.4
Restraints violation energy	71.2 +/- 10.83
Buried Surface Area	1693.0 +/- 78.4

**HADDOCK**  
Software web portal

Home HADDOCK Whispy DNA Publications Forum Contact

WELCOME TO THE Utrecht BIOMOLECULAR INTERACTION WEB PORTAL >>

HADDOCK server status for docking run e2ahpr

Status: FINISHED

The run has been completed. The complete run can be downloaded as a gzipped tar file [here](#) (unavailable). The structures have been clustered, and the top 10 clusters are shown here. The top cluster is the most stable.

**CLUSTER 1**

HADDOCK score	-154.5 +/- 4.4
Cluster size	126
Energy structure	0.6 +/- 0.4
Van der Waals energy	-39.0 +/- 1.3
Electrostatic energy	-591.8 +/- 15.3
Desolvation energy	-0.4 +/- 4.8
Restraints violation energy	32.6 +/- 12.40
Buried Surface Area	1648.4 +/- 40.3

**CLUSTER 3**

HADDOCK score	-141.1 +/- 8.7
Cluster size	20
Energy structure	0.6 +/- 0.4
Van der Waals energy	-39.0 +/- 1.3
Electrostatic energy	-591.8 +/- 15.3
Desolvation energy	-0.4 +/- 4.8
Restraints violation energy	32.6 +/- 12.40
Buried Surface Area	1648.4 +/- 40.3

**3D Structure Viewers**

Jmol structure viewer. Your browser must be Java enabled:

[Nr 1 best structure](#) [View structure](#) [Download structure](#)

[Nr 2 best structure](#) [View structure](#) [Download structure](#)

[Nr 3 best structure](#) [View structure](#) [Download structure](#)

[Nr 4 best structure](#) [View structure](#) [Download structure](#)

**3D Structure Viewers**

Jmol structure viewer. Your browser must be Java enabled:

[Nr 1 best structure](#) [View structure](#) [Download structure](#)

[Nr 2 best structure](#) [View structure](#) [Download structure](#)

[Nr 3 best structure](#) [View structure](#) [Download structure](#)

[Nr 4 best structure](#) [View structure](#) [Download structure](#)

ClusPro 2.0: protein-protein docking

cluspro.bu.edu/login.php?redir=/home.php

Reader

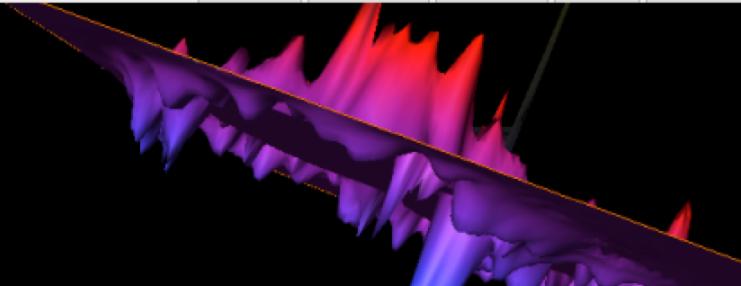
PubMed dizionari databases BTN News METEO Bandi Sapienza biocomp ISI U-GOV Fisica ISI Web

ClusPro 2.0: protein-protein docking ClusPro 2.0: protein-protein docking

Login Sign Up Papers Help Contact

# ClusPro

protein-protein docking



## Welcome to Cluspro 2.0

Recent news: [ClusPro server tops the competition in the latest rounds of CAPRI experiment](#)

**Use Without an Account**

[Use the server without the benefits of your own account](#)

--or--

**Login**

Username:

Password:

--or--

[\*\*Sign up for an account\*\*](#)

Forgot Password?  
[Reset Password](#)

Forgot Username?  
[Retrieve Username](#)

ClusPro 2.0: protein-protein docking

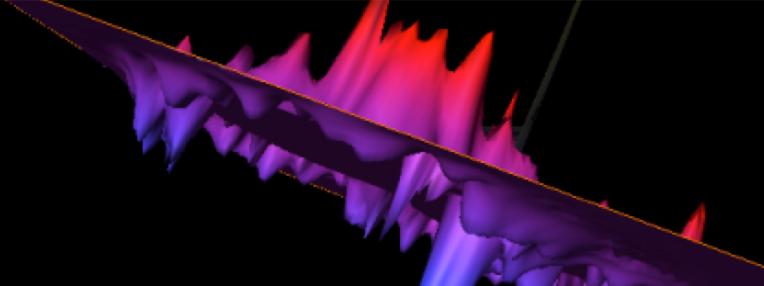
cludpro.bu.edu/home.php Reader

PubMed dizionari databases BTN News METEO Bandi Sapienza biocomp ISI U-GOV

ClusPro 2.0: protein-protein docking ClusPro 2.0: protein-protein docking +

Dock Queue Results Papers Help Contact

**ClusPro**  
protein-protein docking



sign out

## Dock

Note: all jobs by non logged in users will be publicly accessible. Please create an account if data is embargoed and needs to remain confidential

Job Name:

Server:

Accepted PDB Input:  
20 standard amino acids and RNA (as receptor only), ref: [RNA](#) Select Heparin Mode to use Heparin as Ligand.

<b>Receptor</b>	<b>Ligand</b>
PDB ID: <input type="text"/>	PDB ID: <input type="text"/>
<a href="#">Upload PDB</a>	<a href="#">Upload PDB</a>
Chains: <input type="text"/>	Chains: <input type="text"/>

Whitespace separate desired chains. Leave chains blank to use all chains.

» Advanced Options

I agree to use ClusPro only for noncommercial purposes.

# ClusPro

<http://nrc.bu.edu/cluster>

- **The first fully automated, web-based program for the computational docking of protein structures**
- The docking algorithms evaluate billions of putative complexes, retaining a preset number with favorable surface complementarities
- A filtering method is then applied to this set of structures, selecting those with good electrostatic and desolvation free energies for further clustering
- The program output is a short list of putative complexes ranked according to their clustering properties

# ClusPro

<http://nrc.bu.edu/cluster>

- **User Input:** PDB files of the 2 protein structures
- **Output:** 10 (default) top predictions of docked conformations close to the native structure
- Rigid body docking is performed using 2 established FFT-based docking programs (DOT and ZDOCK)
- The scoring is solely based on the surface complementarity between the two structures
- Over  $2.7 \cdot 10^{10}$  structures are evaluated, retaining 20 000 structures with the best surface complementarity scores, which are then further subjected to an empirical free energy filtering algorithm

# ClusPro

<http://nrc.bu.edu/cluster>

- The free energy of the complex structure is typically dominated by van der Waals interactions, which is a very noisy function and difficult to calculate, especially with incorrect side-chain rotamers
- Therefore, other components of the binding free energy are used to account for the noise of the van der Waals energies
  - Desolvation free energy using the atomic contact potential (statistical measure of the desolvation free energy)
  - Electrostatic free energy using a Coulombic model with a distance-dependent dielectric of  $4r$

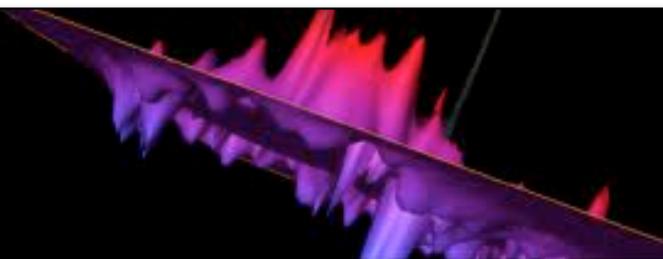
# ClusPro

<http://nrc.bu.edu/cluster>

- The top 2000 energetically favorable structures are clustered on the basis of a pairwise binding site root mean squared deviation (RMSD) criterion
- Clusters are then formed by selecting the ligand that has the most neighbors below a previously selected clustering radius.
- The ligand with the most neighbors is the cluster center, and is the representative structure for the cluster

# ClusPro

protein-protein docking



## Dock

Announcement: New [Downloads section](#) offers KVM image of ClusPro and PIPER binary for your use on your machine.

Job Name:

Accepted PDB Input:  
20 standard amino acids and RNA (as receptor only), ref: [RNA](#)

### Receptor

PDB ID:  [Sfoglia...](#)



[Use PDB ID](#) [Upload PDB](#)

Chains:

### Ligand

PDB ID:  [Sfoglia...](#)



[Use PDB ID](#) [Upload PDB](#)

Chains:

Whitespace separate desired chains. Leave chains blank to use all chains.

### Advanced Options

### Attraction and Repulsion

Enter attraction and repulsion of residues as whitespace separated "chain-residue" entries.  
eg. a-23 a-25 a-26 a-27

Attraction:

Attraction:

Repulsion:

Repulsion:



# Course materials

<https://github.com/Protein-Interactions/BolognaMaster2016/>

## Suggested readings:

- La et al. Predicting permanent and transient protein-protein interfaces. *Proteins* 2013
- Perkin et al. "Transient Protein-Protein Interactions: Structural, Functional, and Network Properties" *Structure* 2010
- Bendell et al. Transient protein-protein interface prediction: datasets, features, algorithms, and the RAD-T predictor. *BMC Bioinformatics* 2013