

M1 ISDD - BI (Feb 2018)

PROTEIN DOCKING

## Lesson 2

# Geometry-based docking methods

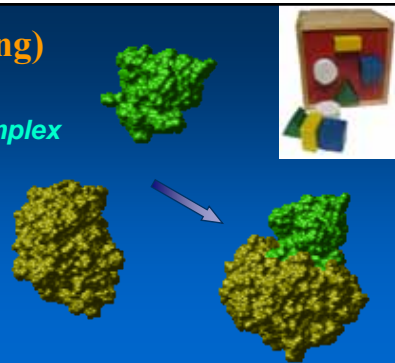
*Juan Fernández-Recio*

CSIC, BSC

[juanf@bsc.es](mailto:juanf@bsc.es)

## Complex structure prediction (docking)

*Generation of the structure of a protein-protein complex  
from the individual protein structures*

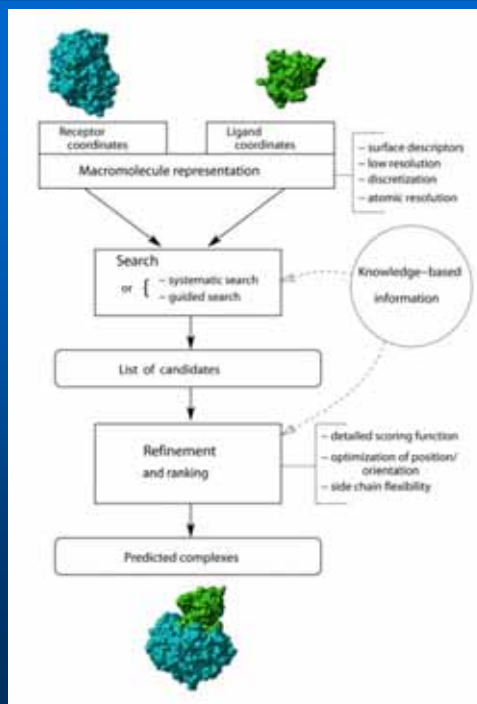


## ab initio docking

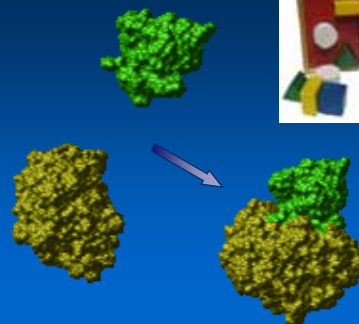


SAMPLING & SCORING

SAMPLING & SCORING

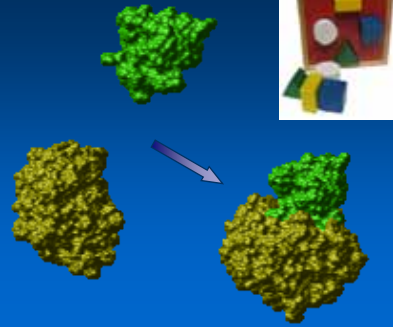


## Docking Search Strategies

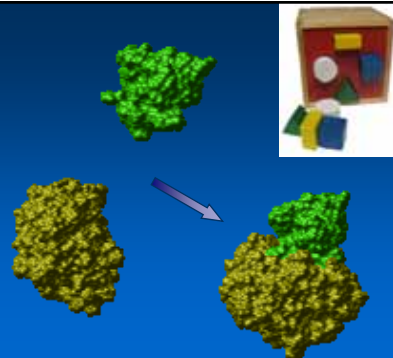


## Docking Search Strategies

- **Pseudo Random**
  - Simulated Annealing / Monte Carlo
  - Genetic Algorithms
- **Directed Search**
  - Geometric Hashing
  - Spherical Harmonic Surface Triangles
- **Brute-Force Search**
  - Explicit Grid Correlations
  - Fast Fourier Transform (FFT) Correlations
  - Spherical Polar Fourier Correlations
- **Refinement Phase**
  - Classical or Soft Potentials (+/- Electrostatics)
  - Desolvation, Solvent Dipoles...
  - Visual Inspection!!

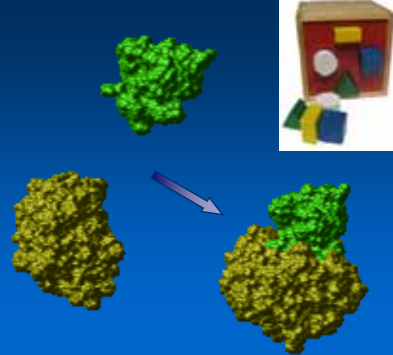


## Criteria for Good Docking Orientations



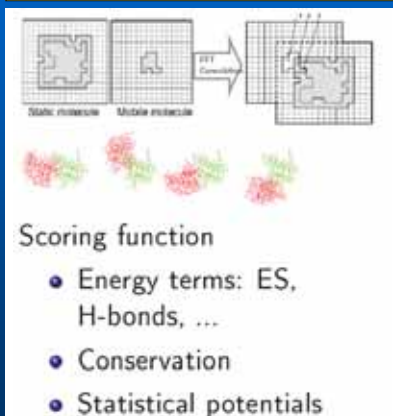
### Criteria for Good Docking Orientations

- Low Free Energy (Difficult!)
- Low Pseudo-Energy (Easy) Based On...
- Large Surface Burial:  $\sim 1600 \pm 400 \text{ \AA}^2$
- Small van der Waals Overlaps
- No Large Cavities in Interface
- Good H-Bonding:  $\sim 1 \text{ HB}/100 \text{ \AA}^2$
- Good Charge Complementarity
- Polar/Polar Contacts Favoured
- Polar/Non-Polar Contacts Disfavoured



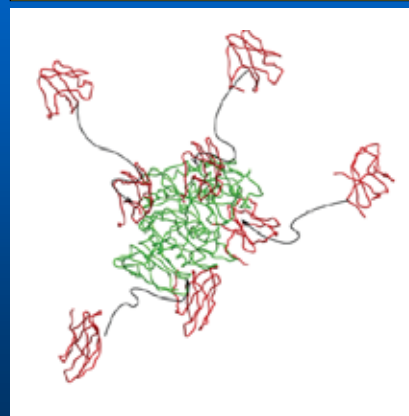
## Protein-protein docking methods

### Exhaustive search (FFT, surface-based)



Geometry-based docking

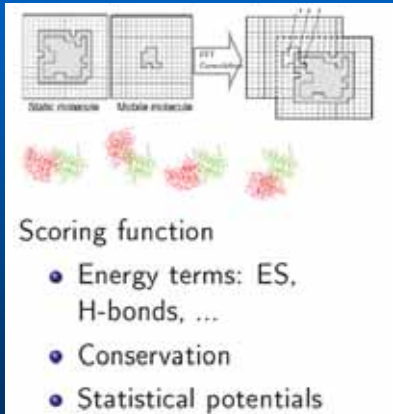
### Stochastic sampling (Monte-Carlo, minimization)



Energy-based docking

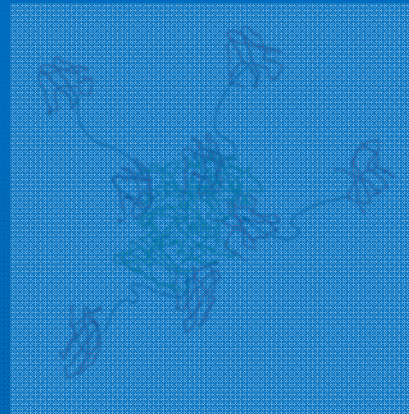
# Protein-protein docking methods

## Exhaustive search (FFT, surface-based)



Geometry-based docking

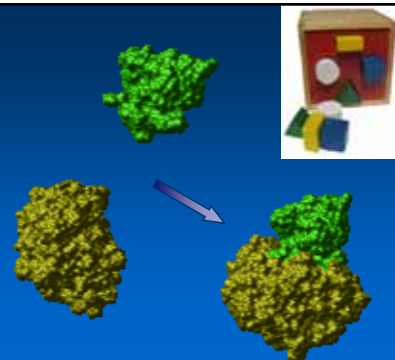
## Stochastic sampling (Monte-Carlo, minimization)



Energy-based docking

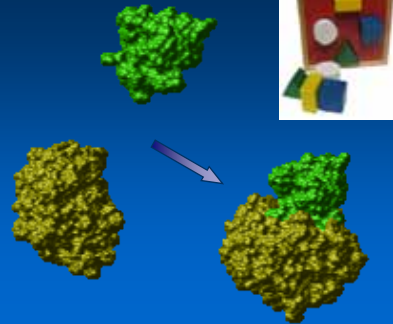
## Docking Search Strategies

- **Pseudo Random**
  - Simulated Annealing / Monte Carlo
  - Genetic Algorithms
- **Directed Search**
  - Geometric Hashing
  - Spherical Harmonic Surface Triangles
- **Brute-Force Search**
  - Explicit Grid Correlations
  - Fast Fourier Transform (FFT) Correlations
  - Spherical Polar Fourier Correlations
- **Refinement Phase**
  - Classical or Soft Potentials (+/- Electrostatics)
  - Desolvation, Solvent Dipoles...
  - Visual Inspection!!



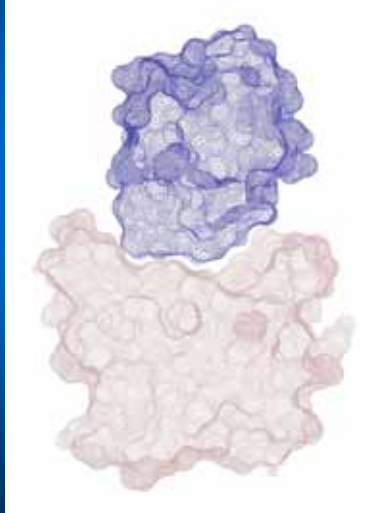
### Criteria for Good Docking Orientations

- Low Free Energy (Difficult!)
- Low Pseudo-Energy (Easy) Based On...
- Large Surface Burial:  $\sim 1600 \pm 400 \text{ \AA}^2$
- Small van der Waals Overlaps
- No Large Cavities in Interface
- Good H-Bonding:  $\sim 1 \text{ HB}/100 \text{ \AA}^2$
- Good Charge Complementarity
- Polar/Polar Contacts Favoured
- Polar/Non-Polar Contacts Disfavoured

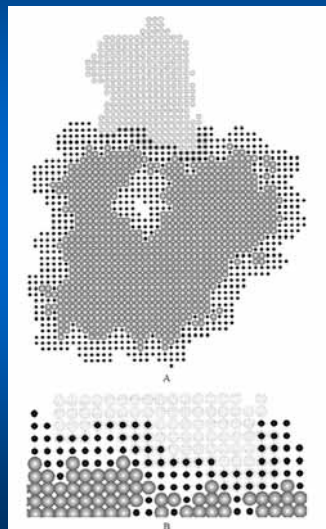
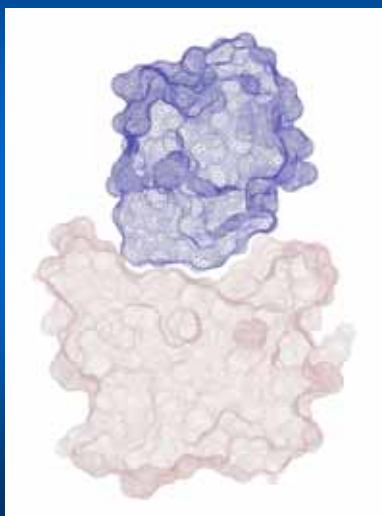


- FFT-based grid search
- Geometric Hashing
- Adding distance-, symmetry- constraints

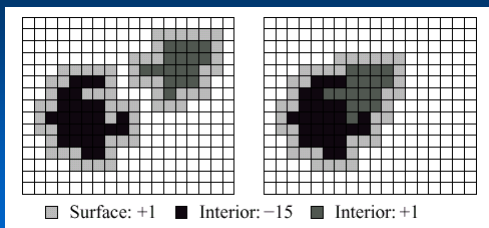
## Rigid-Body Docking: Geometry Approach



## Rigid-Body Docking: Geometry Approach



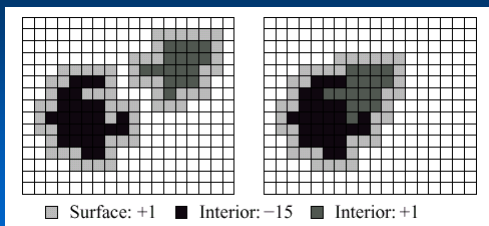
# MolFit



$$\begin{aligned}
 \tilde{a}_{l,m,n} &= \begin{cases} 1 & \text{on the surface of the molecule} \\ \rho & \text{inside the molecule} \\ 0 & \text{outside the molecule} \end{cases} \\
 \tilde{b}_{l,m,n} &= \begin{cases} 1 & \text{on the surface of the molecule} \\ \delta & \text{inside the molecule} \\ 0 & \text{outside the molecule} \end{cases} \\
 l, m, n &= (1, \dots, N) \\
 \tilde{c}_{\alpha,\beta,\gamma} &= \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \tilde{a}_{l,m,n} \cdot \tilde{b}_{l+\alpha, m+\beta, n+\gamma}
 \end{aligned}$$

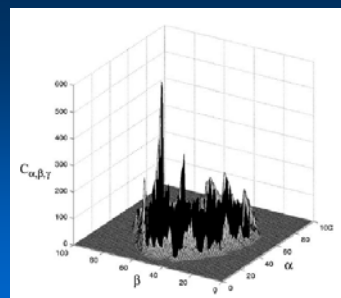
E. Katchalski-Katzir, I. Shariv, M. Eisenstein, A.A. Friesem, C. Aflalo, I.A. Vakser, Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques, Proc. Natl Acad. Sci. USA 89 (1992) 2195–2199.

# MolFit



$$\begin{aligned}
 \tilde{a}_{l,m,n} &= \begin{cases} 1 & \text{on the surface of the molecule} \\ \rho & \text{inside the molecule} \\ 0 & \text{outside the molecule} \end{cases} \\
 \tilde{b}_{l,m,n} &= \begin{cases} 1 & \text{on the surface of the molecule} \\ \delta & \text{inside the molecule} \\ 0 & \text{outside the molecule} \end{cases} \\
 l, m, n &= (1, \dots, N) \\
 \tilde{c}_{\alpha,\beta,\gamma} &= \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \tilde{a}_{l,m,n} \cdot \tilde{b}_{l+\alpha, m+\beta, n+\gamma}
 \end{aligned}$$

E. Katchalski-Katzir, I. Shariv, M. Eisenstein, A.A. Friesem, C. Aflalo, I.A. Vakser, Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques, Proc. Natl Acad. Sci. USA 89 (1992) 2195–2199.



a  
 b  
 ↓  
 Corr(a,b)  
 ↓  
 c



# Fourier Transform

*It re-express a function in terms of sinusoidal basis functions*

Fourier transform equations:

$$F(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i k x} dx$$
$$f(x) = \int_{-\infty}^{\infty} F(k) e^{2\pi i k x} dk$$

$$F(k) = \text{FT}(f(x))$$

$$f(x) = \text{IFT}(F(k))$$

# Fourier Transform

*It re-express a function in terms of sinusoidal basis functions*

Fourier transform equations:

$$F(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i k x} dx$$
$$f(x) = \int_{-\infty}^{\infty} F(k) e^{2\pi i k x} dk$$

$$F(k) = \text{FT}(f(x))$$

$$f(x) = \text{IFT}(F(k))$$

Correlation function:

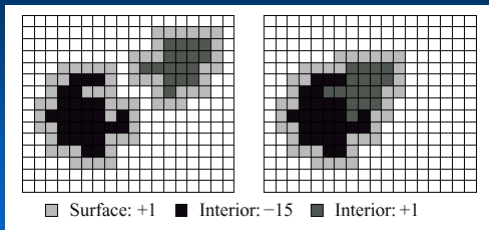
$$c(t) = \text{Corr}(g, h) \equiv \int_{-\infty}^{\infty} g(\tau + t) h(\tau) d\tau$$

“Correlation Theorem”:

$$\text{FT}(c) = \text{FT}(g) \cdot [\text{FT}(h)]^*$$

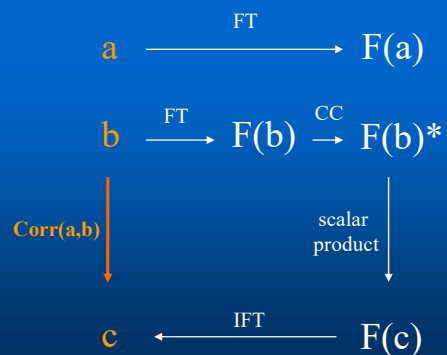
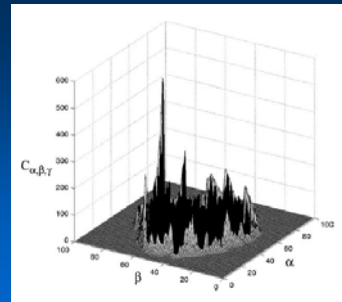
$$c = \text{IFT}(\text{FT}(g) \cdot [\text{FT}(h)]^*)$$

# MolFit



$$\begin{aligned}
 \tilde{a}_{l,m,n} &= \begin{cases} 1 & \text{on the surface of the molecule} \\ \rho & \text{inside the molecule} \\ 0 & \text{outside the molecule} \end{cases} \\
 \tilde{b}_{l,m,n} &= \begin{cases} 1 & \text{on the surface of the molecule} \\ \delta & \text{inside the molecule} \\ 0 & \text{outside the molecule} \end{cases} \\
 l, m, n &= (1, \dots, N) \\
 \tilde{c}_{\alpha,\beta,\gamma} &= \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \tilde{a}_{l,m,n} \cdot \tilde{b}_{l+\alpha, m+\beta, n+\gamma}
 \end{aligned}$$

E. Katchalski-Katzir, I. Shariv, M. Eisenstein, A. A. Friesem, C. Aflalo, I. A. Vakser, Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques, Proc. Natl Acad. Sci. USA 89 (1992) 2195-2199.



## FFT (Fast Fourier Transform)

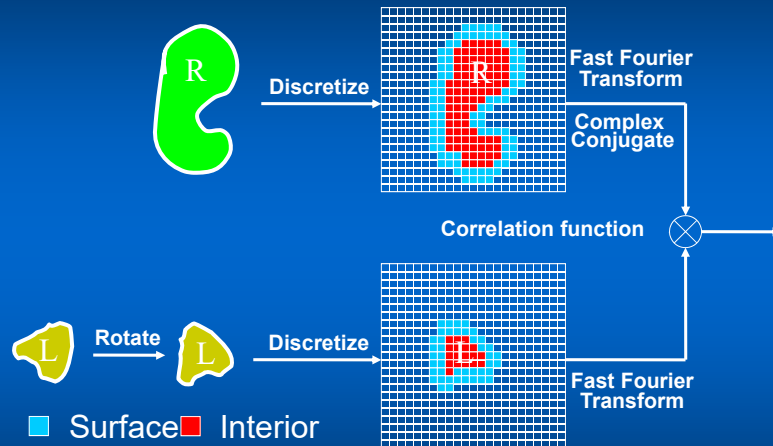
Algorithms for efficient calculation of FT and IFT

- Most common: Cooley-Tukey FFT (divide and conquer)
- Other: Prime-factor, Bruun's, Rader's, Bluestein's

Fourier transform timing:  $N^2$  (if  $N=10^6$  1MHz CPU time ~2 weeks)  
 Fast Fourier transform:  $N \log_2 N$  (if  $N=10^6$  1MHz CPU time ~30 sec)

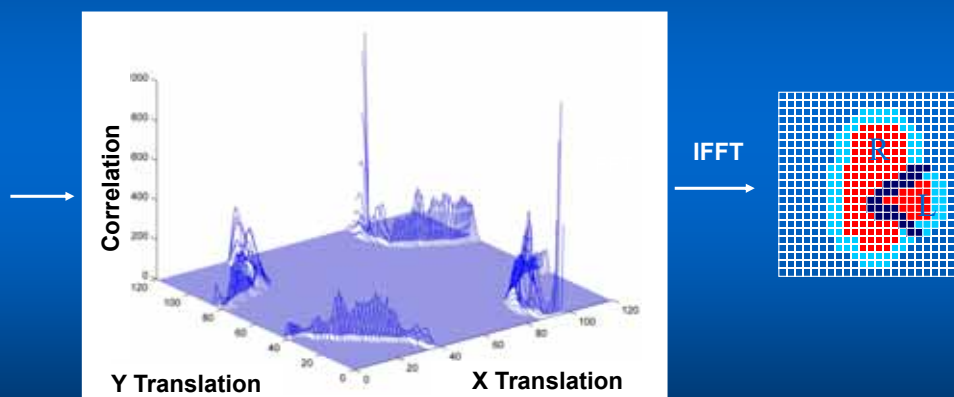
<http://www.fftw.org/>

# Protein Docking Using FFT

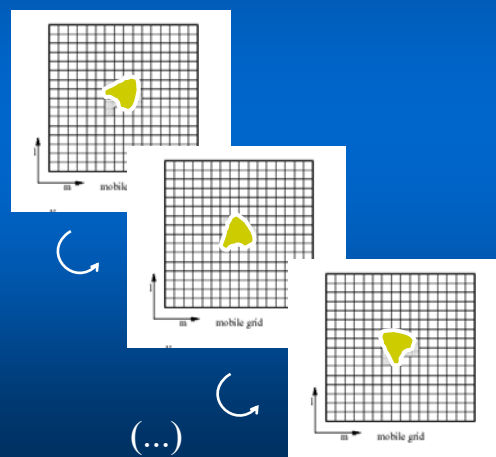
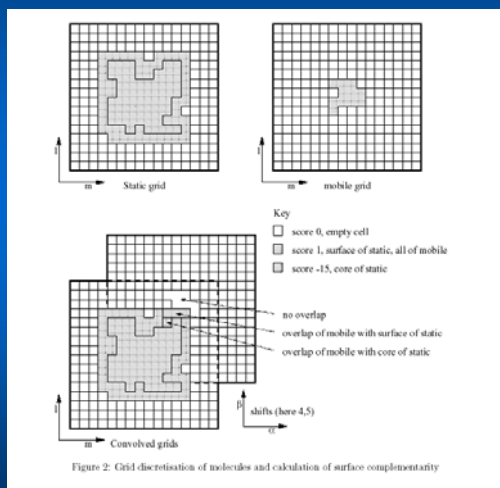


# Protein Docking Using FFT

Comp. cost can decrease by  $>10^4$  (from  $N^6$  to  $N^3 \ln N^3$ )

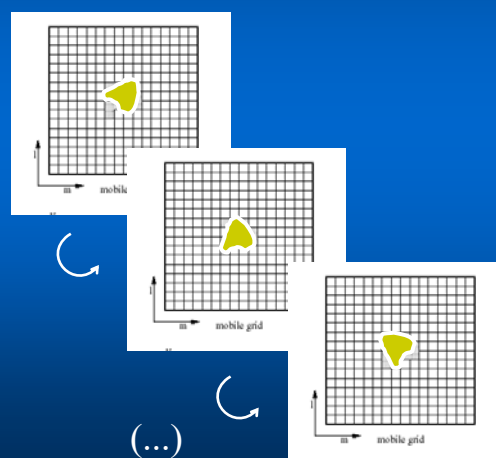
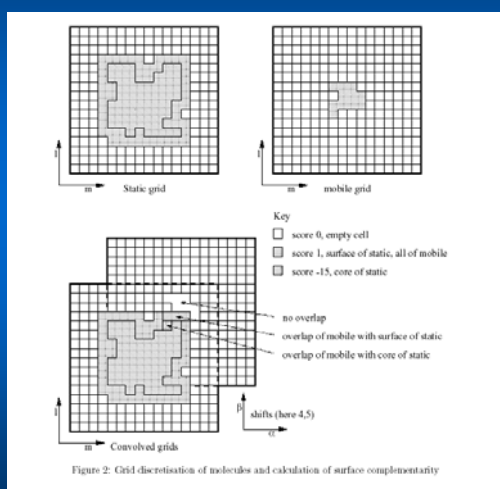


# FTDOCK



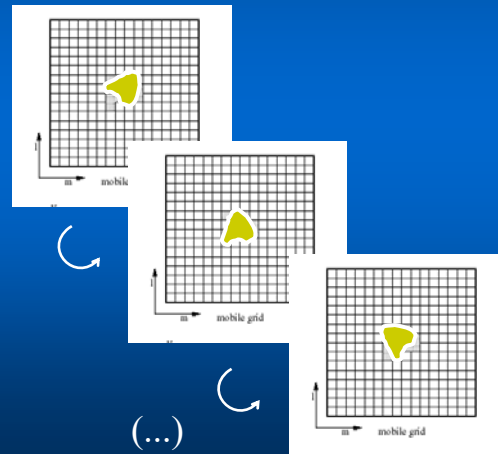
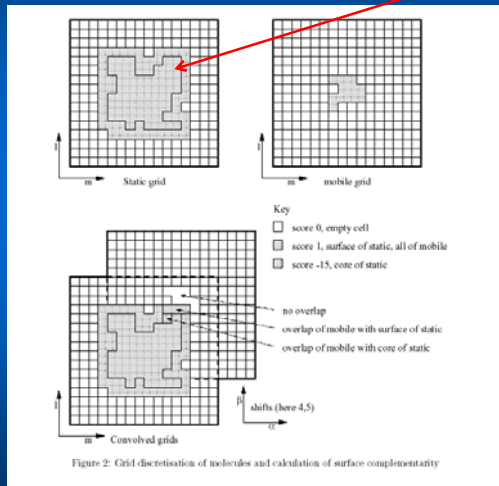
# FTDOCK

```
ftdock -static 1A2P_A.pdb -mobile 1A19_A.pdb
-noelec -calculate_grid 1.2 -angle_step 12
-internal -15 -surface 1.3 -keep 3
-out 1BRS.ftdock > 1BRS.ftdock.log
```



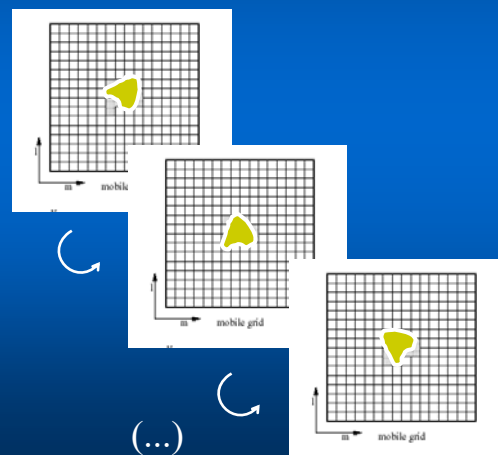
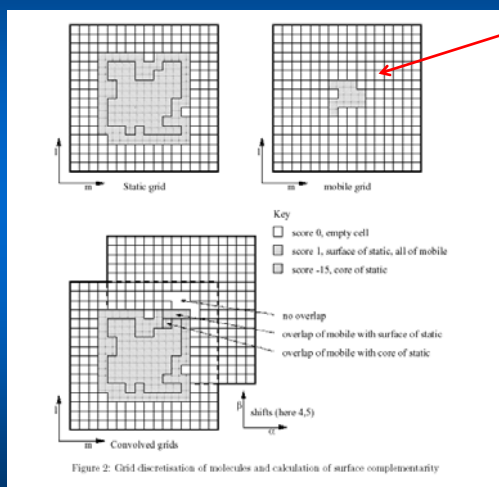
# FTDOCK

```
ftdock -static 1A2P_A.pdb -mobile 1A19_A.pdb
-noelec -calculate_grid 1.2 -angle_step 12
-internal -15 -surface 1.3 -keep 3
-out 1BRS.ftdock > 1BRS.ftdock.log
```



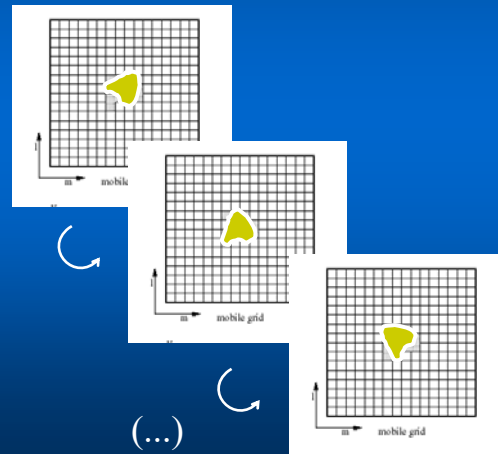
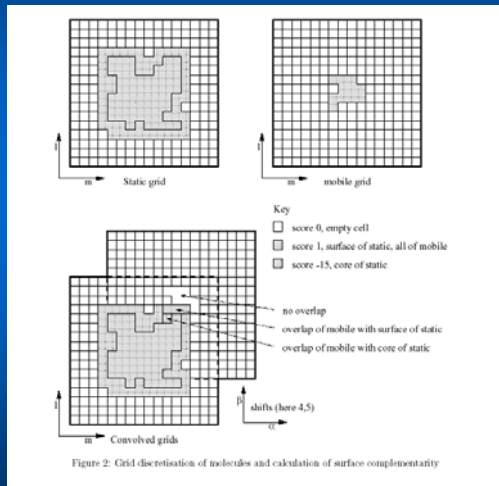
# FTDOCK

```
ftdock -static 1A2P_A.pdb -mobile 1A19_A.pdb
-noelec -calculate_grid 1.2 -angle_step 12
-internal -15 -surface 1.3 -keep 3
-out 1BRS.ftdock > 1BRS.ftdock.log
```



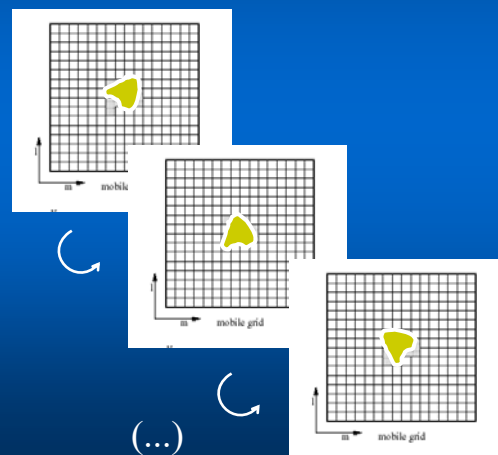
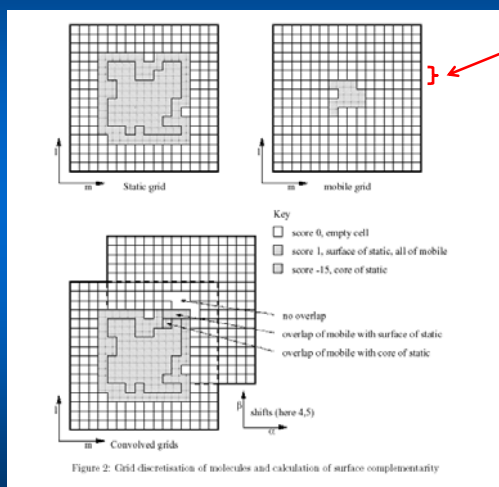
# FTDOCK

```
ftdock -static 1A2P_A.pdb -mobile 1A19_A.pdb
-noelec -calculate_grid 1.2 -angle_step 12
-internal -15 -surface 1.3 -keep 3
-out 1BRS.ftdock > 1BRS.ftdock.log
```



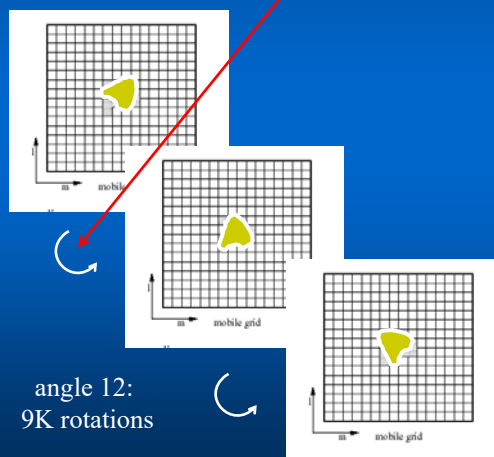
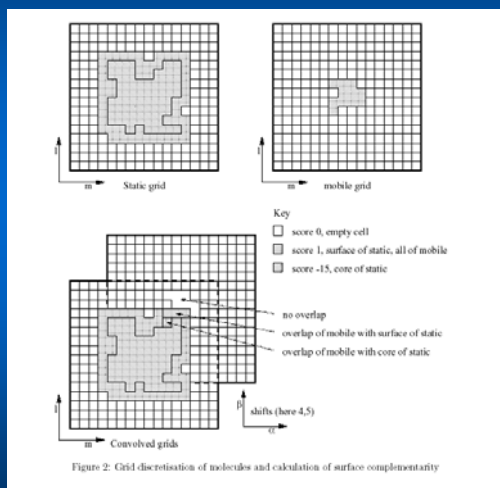
# FTDOCK

```
ftdock -static 1A2P_A.pdb -mobile 1A19_A.pdb
-noelec -calculate_grid 1.2 -angle_step 12
-internal -15 -surface 1.3 -keep 3
-out 1BRS.ftdock > 1BRS.ftdock.log
```



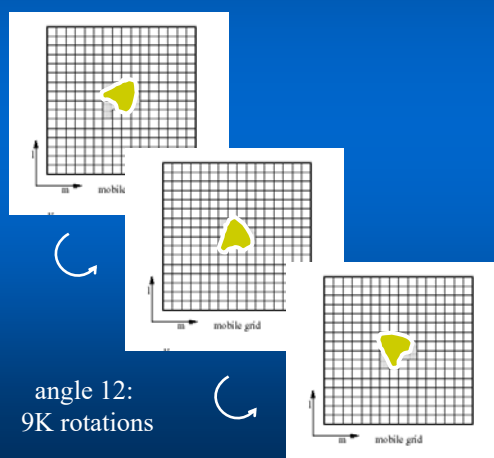
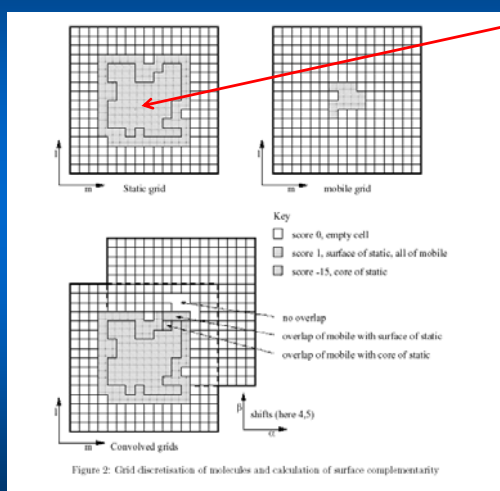
# FTDOCK

```
ftdock -static 1A2P_A.pdb -mobile 1A19_A.pdb
-noelec -calculate_grid 1.2 -angle_step 12
-internal -15 -surface 1.3 -keep 3
-out 1BRS.ftdock > 1BRS.ftdock.log
```



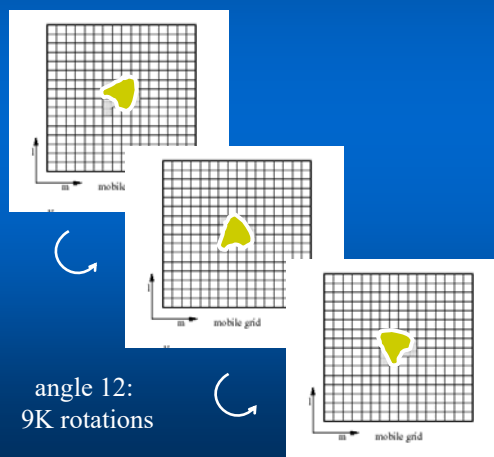
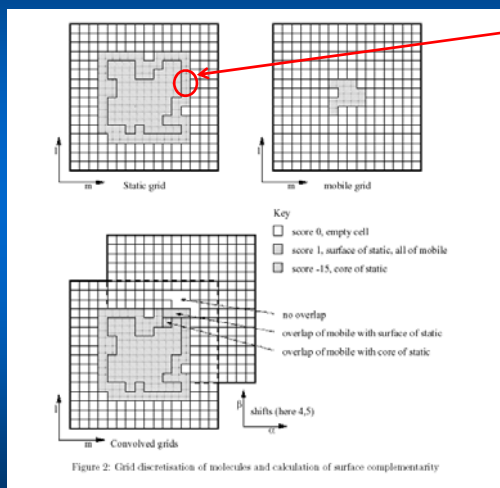
# FTDOCK

```
ftdock -static 1A2P_A.pdb -mobile 1A19_A.pdb
-noelec -calculate_grid 1.2 -angle_step 12
-internal -15 -surface 1.3 -keep 3
-out 1BRS.ftdock > 1BRS.ftdock.log
```



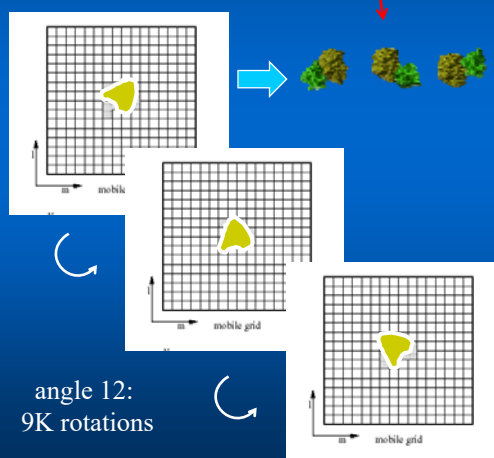
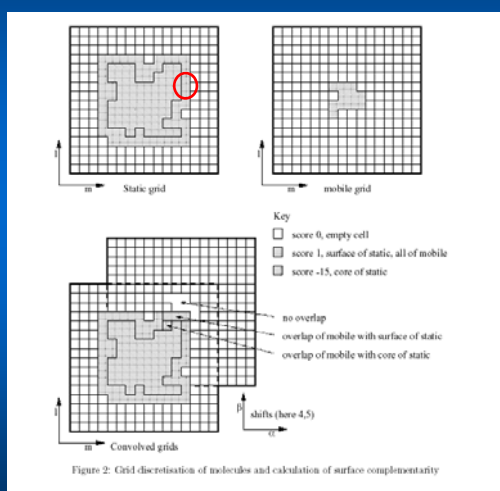
# FTDOCK

```
ftdock -static 1A2P_A.pdb -mobile 1A19_A.pdb
-noelec -calculate_grid 1.2 -angle_step 12
-internal -15 -surface 1.3 -keep 3
-out 1BRS.ftdock > 1BRS.ftdock.log
```



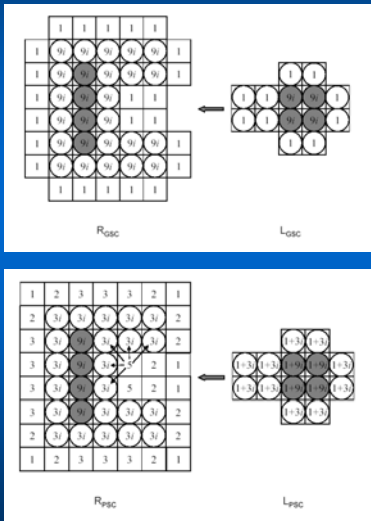
# FTDOCK

```
ftdock -static 1A2P_A.pdb -mobile 1A19_A.pdb
-noelec -calculate_grid 1.2 -angle_step 12
-internal -15 -surface 1.3 -keep 3
-out 1BRS.ftdock > 1BRS.ftdock.log
```

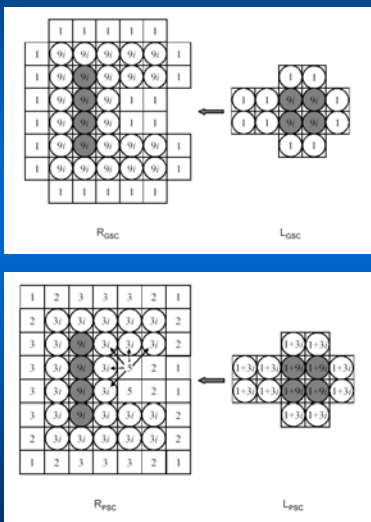




# ZDOCK

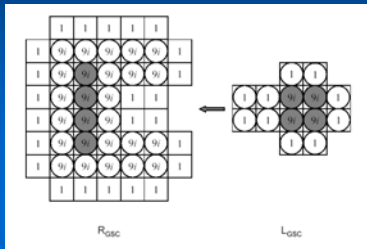


# ZDOCK

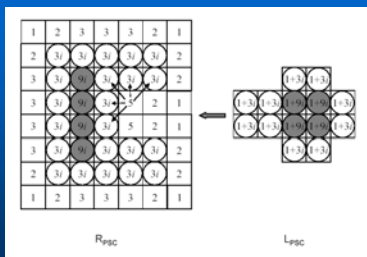


```
mark_sur 1A2P_A.pdb 1BRS_r.pdb
mark_sur 1A19_A.pdb 1BRS_l.pdb
```

# ZDOCK

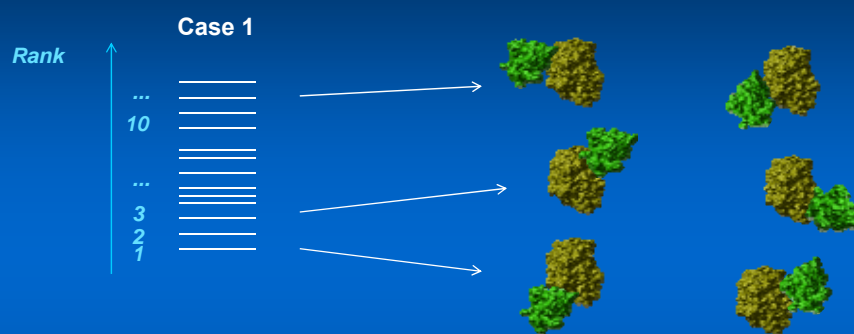


mark\_sur 1A2P\_A.pdb 1BRS\_r.pdb  
mark\_sur 1A19\_A.pdb 1BRS\_l.pdb

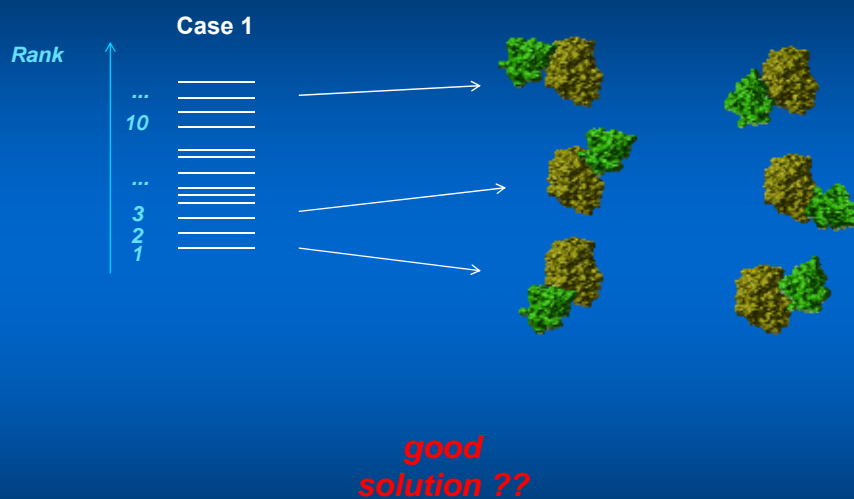


zdock -R 1BRS\_r.pdb -L 1BRS\_l.pdb -o 1BRS\_zdock.out

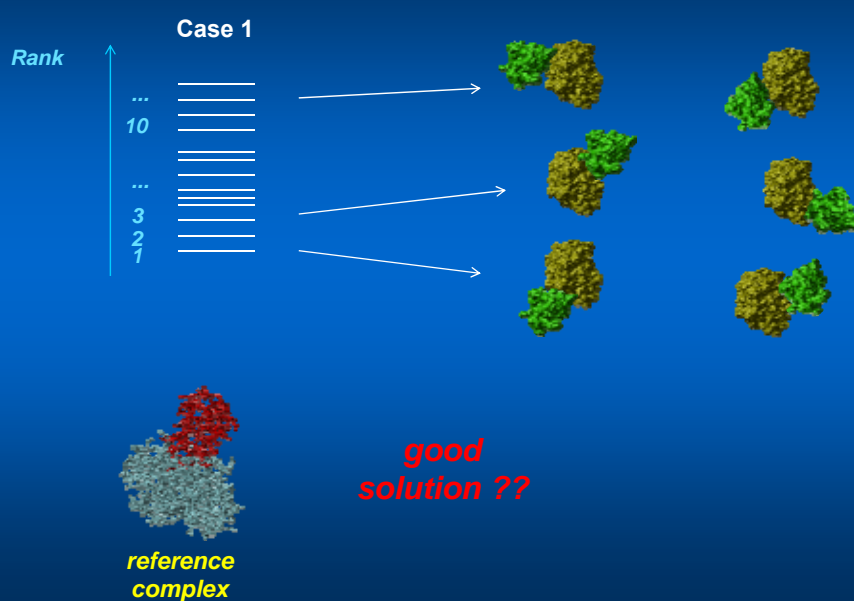
## Assessment of docking performance



## Assessment of docking performance



## Assessment of docking performance



## Assessment of docking performance

- $F_{\text{nat}}$ : fraction of native contacts (within  $5\text{\AA}$ )

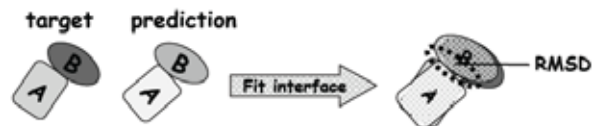
$$F_{\text{nat}} = \frac{\text{Correctly predicted contacts}}{\text{Total number of contacts in the target}}$$

- l-RMSD: RMSD on second protein after superposition on first

near-native:  
L-RMSD  $< 10\text{\AA}$



- i-RMSD: RMSD on interface residues (within  $10\text{\AA}$ )

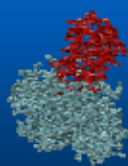


## Assessment of docking performance

Rank	Case 1	L-RMSD
...	...	...
10	...	...
...	...	2.6
...	...	...
3	...	5.4
2	...	25.8
1	...	16.3

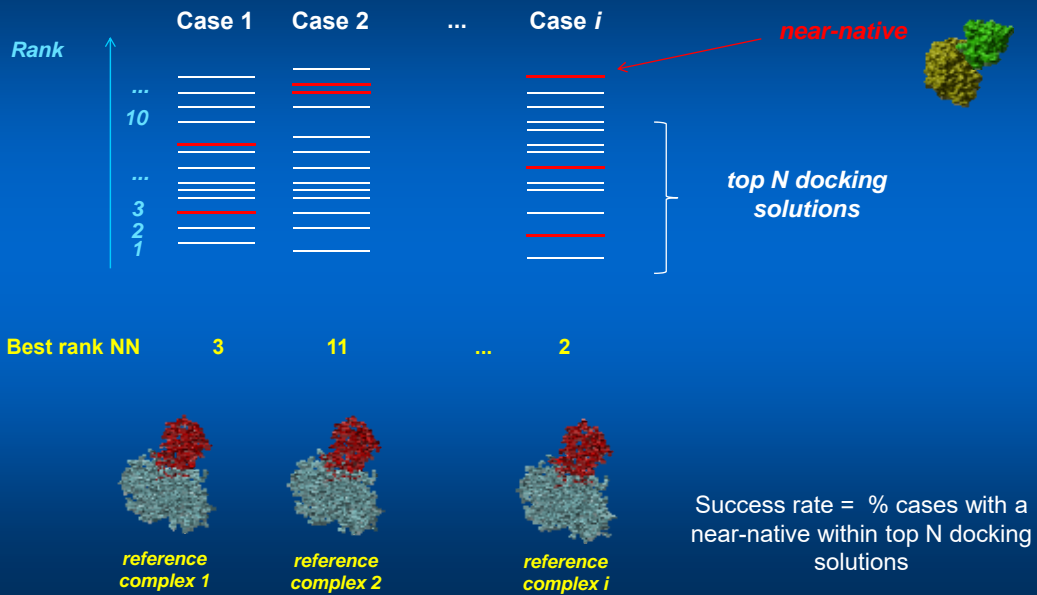
near-native

Best rank NN 3



reference  
complex

## Assessment of docking performance



## Protein-protein Docking Benchmark 4.0 - Weng

[Home](#)
[About](#)
[Contact](#)
[Privacy Policy](#)
[Terms of Service](#)
[Disclaimer](#)
[Feedback](#)

[Download protein as PDB files for the benchmark in a zip file](#)
[Download Table S1 as excel file](#)

[Updates and corrections \(August, 16, 2010\)](#)

Please note that 1A77 protein 1 TRGH is corrected from Ransome to Wasee Sa.

## Supplementary information for

Weng, H., Vreven, T., Jerng, J. and Wang, Z. Protein-Protein Docking Benchmark Version 4.0 Proteins, 2010 (Accepted)

[Download protein as PDB files for the benchmark in a zip file](#)
[Download Table S1 as excel file](#)

[Updates and corrections \(August, 16, 2010\)](#)

Please note that 1A77 protein 1 TRGH is corrected from Ransome to Wasee Sa.

Follow the "info" link in the leftmost column to view bound/unbound chain ECAT alignments and potentially useful comments about individual benchmark cases

**Table S1. Protein-Protein Docking Benchmark 4.0**

Info	Complex	Protein 1	Protein 2	Protein 3	Protein 4	Protein 5	Protein 6	Protein 7	Protein 8	Protein 9	Protein 10	Protein 11	Protein 12	Protein 13	Protein 14	Protein 15	Protein 16	Protein 17	Protein 18	Protein 19	Protein 20	Protein 21	Protein 22	Protein 23	Protein 24	Protein 25	Protein 26	Protein 27	Protein 28	Protein 29	Protein 30	Protein 31	Protein 32	Protein 33	Protein 34	Protein 35	Protein 36	Protein 37	Protein 38	Protein 39	Protein 40	Protein 41	Protein 42	Protein 43	Protein 44	Protein 45	Protein 46	Protein 47	Protein 48	Protein 49	Protein 50	Protein 51	Protein 52	Protein 53	Protein 54	Protein 55	Protein 56	Protein 57	Protein 58	Protein 59	Protein 60	Protein 61	Protein 62	Protein 63	Protein 64	Protein 65	Protein 66	Protein 67	Protein 68	Protein 69	Protein 70	Protein 71	Protein 72	Protein 73	Protein 74	Protein 75	Protein 76	Protein 77	Protein 78	Protein 79	Protein 80	Protein 81	Protein 82	Protein 83	Protein 84	Protein 85	Protein 86	Protein 87	Protein 88	Protein 89	Protein 90	Protein 91	Protein 92	Protein 93	Protein 94	Protein 95	Protein 96	Protein 97	Protein 98	Protein 99	Protein 100																																																																																																					
1A77	1A77_A	A	1A77_B	Protein 1	1A77_C	Protein 2	1A77_D	Protein 3	1A77_E	Protein 4	1A77_F	Protein 5	1A77_G	Protein 6	1A77_H	Protein 7	1A77_I	Protein 8	1A77_J	Protein 9	1A77_K	Protein 10	1A77_L	Protein 11	1A77_M	Protein 12	1A77_N	Protein 13	1A77_O	Protein 14	1A77_P	Protein 15	1A77_Q	Protein 16	1A77_R	Protein 17	1A77_S	Protein 18	1A77_T	Protein 19	1A77_U	Protein 20	1A77_V	Protein 21	1A77_W	Protein 22	1A77_X	Protein 23	1A77_Y	Protein 24	1A77_Z	Protein 25	1A77_AA	Protein 26	1A77_AB	Protein 27	1A77_AC	Protein 28	1A77_AD	Protein 29	1A77_AE	Protein 30	1A77_AF	Protein 31	1A77_AG	Protein 32	1A77_AH	Protein 33	1A77_AI	Protein 34	1A77_AJ	Protein 35	1A77_AK	Protein 36	1A77_AL	Protein 37	1A77_AM	Protein 38	1A77_AN	Protein 39	1A77_AO	Protein 40	1A77_AP	Protein 41	1A77_AQ	Protein 42	1A77_AR	Protein 43	1A77_AS	Protein 44	1A77_AT	Protein 45	1A77_AU	Protein 46	1A77_AV	Protein 47	1A77_AW	Protein 48	1A77_AX	Protein 49	1A77_AY	Protein 50	1A77_AZ	Protein 51	1A77_BA	Protein 52	1A77_BB	Protein 53	1A77_BC	Protein 54	1A77 BD	Protein 55	1A77_BE	Protein 56	1A77_BF	Protein 57	1A77_BG	Protein 58	1A77_BH	Protein 59	1A77_BI	Protein 60	1A77_BJ	Protein 61	1A77_BK	Protein 62	1A77_BL	Protein 63	1A77_BM	Protein 64	1A77_BN	Protein 65	1A77_BO	Protein 66	1A77_BP	Protein 67	1A77_BQ	Protein 68	1A77_BR	Protein 69	1A77_BS	Protein 70	1A77_BT	Protein 71	1A77_BU	Protein 72	1A77_BV	Protein 73	1A77_BW	Protein 74	1A77_BX	Protein 75	1A77_BY	Protein 76	1A77_BZ	Protein 77	1A77_CA	Protein 78	1A77_CB	Protein 79	1A77_CC	Protein 80	1A77_CD	Protein 81	1A77_CE	Protein 82	1A77_CF	Protein 83	1A77.CG	Protein 84	1A77_CH	Protein 85	1A77_CI	Protein 86	1A77_CJ	Protein 87	1A77_CK	Protein 88	1A77_CL	Protein 89	1A77_CM	Protein 90	1A77_CN	Protein 91	1A77_CO	Protein 92	1A77_CP	Protein 93	1A77_CQ	Protein 94	1A77_CR	Protein 95	1A77_CS	Protein 96	1A77_CT	Protein 97	1A77_CU	Protein 98	1A77_CV	Protein 99	1A77_CW	Protein 100

176 cases

<http://zlab.umassmed.edu/benchmark/>

# ZDOCK performance

## A Novel Shape Complementarity Function

Number of Predictions	GSC (%)	GSC+Desolvation+Electrostatics (%)	PSC (%)	PSC+Desolvation+Electrostatics (%)
1	5	18	12	18
10	12	32	28	45
100	30	50	45	68
1000	60	78	80	85

$\text{Ref}[R_{\text{ref},i,j,\text{dist}}] = \text{Ref}[L_{\text{ref},i,j,\text{dist}}] = \begin{cases} 3.5 & \text{solvent excluding surface layer of the protein} \\ 3.5^2 & \text{protein core} \\ 0 & \text{open space} \end{cases}$

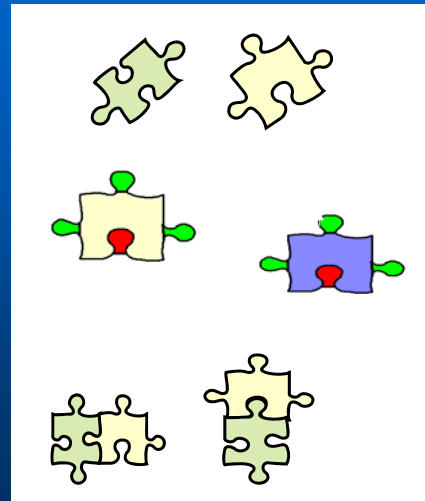
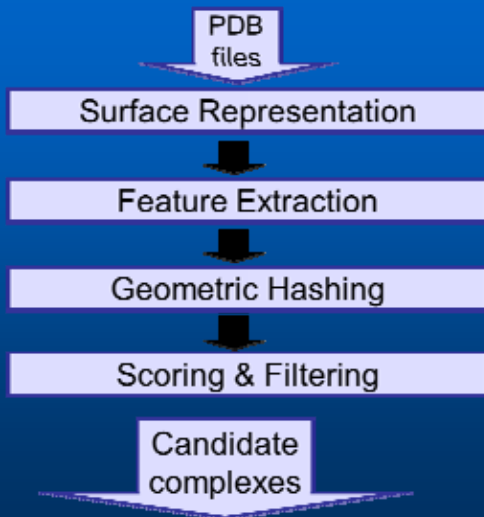
$\ln[R_{\text{ref},i,j,\text{dist}}] = \begin{cases} \beta \times (\text{electric potential of all receptor atoms}) & \text{open space} \\ 0 & \text{otherwise} \end{cases}$

$\ln[L_{\text{ref},i,j,\text{dist}}] = \begin{cases} -1 \times (\text{atom charge}) & \text{if this grid point is the nearest grid point of a ligand atom} \\ 0 & \text{otherwise} \end{cases}$

$S_{\text{ref},i,j,\text{dist}} = \text{Ref}[R_{\text{ref},i,j,\text{dist}}] \times L_{\text{ref},i,j,\text{dist}} + \frac{1}{2} \times \ln[R_{\text{ref},i,j,\text{dist}}]$

- FFT-based grid search
- **Geometric Hashing**
- Adding distance-, symmetry- constraints

## Geometrical hashing - PatchDock

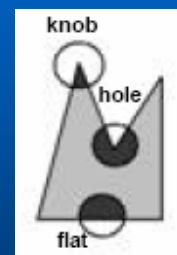
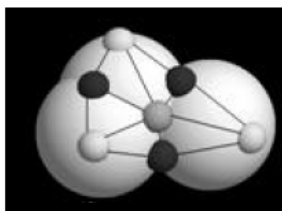
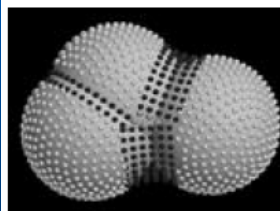


Schneidman-Duhovny et al. Proteins 2003  
Duhovny (Schneidman), D., Nussinov, R. Wolfson, H.J. WABI 2002

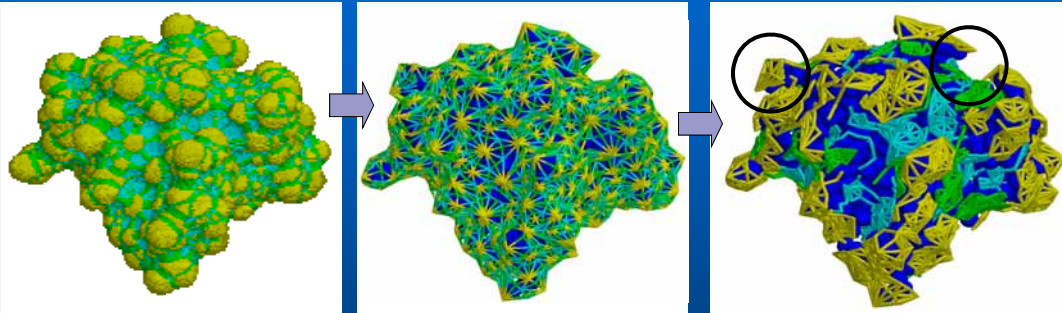
## Surface Representation

### Surface Representation

- Dense MS surface (Connolly)
- Sparse surface (Shuo Lin et al.)



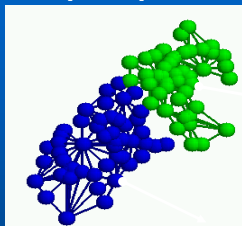
## Feature Extraction



- We focus on sparse surface features, preserving the quality of shape representation.
- The sparse features reduce the complexity of the matching step.

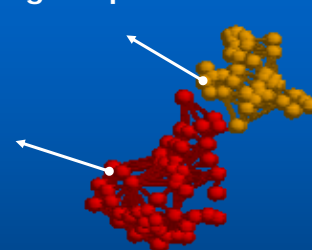
## Patch-Pair Matching

Receptor patches



Transformation

Ligand patches



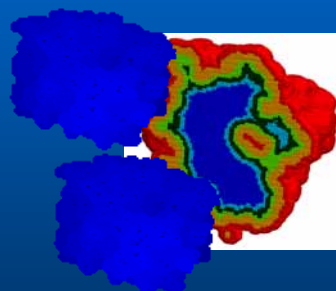
**Base:** 1 critical point with its normal from one patch and  
1 critical point with its normal from a neighboring patch.

- Match every base from the receptor patches with all the bases from complementary ligand patches.
- Compute the transformation for each pair of matched bases.



## Geometric Scoring

- > The surface of the **receptor** is divided into five shells according to the distance function: **S1-S5**
- > The number of **ligand** surface points in every shell is counted.
- > The geometric score is a weighted sum of the number of ligand surface points inside every shell.



- > **Multi-resolution surface** data structure was developed to speed up this stage.

<http://bioinfold.cs.tau.ac.il/PatchDock>

**PATCHDOCK**

Molecular Docking Algorithm Based on Shape Complementarity

Principles

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

Type PDB codes of receptor and ligand molecules or upload files in PDB format:

Receptor Molecule:  (PDB:chainId e.g. Zka1:AB) or upload file:  Browse...

Ligand Molecule:  (PDB:chainId e.g. Zka1:I) or upload file:  Browse...

E-mail address:  (the results are sent to this address)

Clustering RMSD:  4.0

Complex Type:  Default

Submit form Clear

Be sure to give receptor and ligand in the corresponding order!

3 mandatory fields

Clustering RMSD determines the number of output solutions

Complex Type adjusts parameter set for the specific complex type, such as enzyme-inhibitor, antibody-antigen, protein-small ligand

Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. NAR 2005

# PatchDock - server

**Input data**  
**Top 20 solutions**

Receptor	Ligand	Complex Type	Clustering RMSD	User e-mail	Receptor Site	Ligand Site
A.pdb.tr	B.pdb.tr	Default	4.0	duhovka@tau.ac.il	-	-

Solution No	Score	Area	ACE	Transformation	PDB file of the complex
1	10166	1528.90	-12.07	-1.15 -0.37 -0.32 -11.03 -15.21 22.35	<a href="#">result_1.pdb</a>
2	9740	1285.70	129.52	-2.07 -0.40 -1.44 -3.43 9.43 18.35	<a href="#">result_2.pdb</a>
3	9380	1284.70	-8.60	0.77 0.16 -0.54 -1.20 8.06 -10.85	<a href="#">result_3.pdb</a>
4	9122	1298.80	203.84	2.45 -0.57 -1.59 13.45 2.35 14.25	<a href="#">result_4.pdb</a>
5	9048	1185.20	-243.82	3.11 0.11 1.04 -16.19 -29.02 2.25	<a href="#">result_5.pdb</a>
6	9034	1094.70	301.73	2.09 -0.49 -1.37 6.24 3.03 10.96	<a href="#">result_6.pdb</a>
7	8972	1102.40	-39.65	-2.07 -1.05 -2.05 6.41 -17.96 20.59	<a href="#">result_7.pdb</a>
8	8886	1163.80	165.53	0.50 0.03 2.13 13.63 23.66 10.70	<a href="#">result_8.pdb</a>
9	8886	1195.40	-31.63	0.92 0.49 3.07 18.21 -10.46 -11.65	<a href="#">result_9.pdb</a>
10	8850	1240.20	-149.90	-1.43 0.72 0.57 -14.80 -28.15 -1.02	<a href="#">result_10.pdb</a>
11	8768	1062.70	-10.42	-1.12 -0.24 -0.21 12.24 -0.22 21.55	<a href="#">result_11.pdb</a>
12	8740	1110.70	2.06	0.34 0.78 2.56 2.98 3.29 -6.48	<a href="#">result_12.pdb</a>
13	8700	973.80	-79.03	-1.32 -0.69 -0.86 -7.61 -8.97 26.23	<a href="#">result_13.pdb</a>
14	8696	1012.60	-161.10	0.23 0.68 2.24 14.51 -28.88 -3.17	<a href="#">result_14.pdb</a>
15	8644	1206.50	-46.74	-1.89 -0.36 1.91 5.66 -34.25 17.10	<a href="#">result_15.pdb</a>
16	8628	1192.00	330.84	-0.75 0.74 0.78 -0.58 0.95 23.63	<a href="#">result_16.pdb</a>
17	8602	1285.20	17.88	-2.71 0.40 -1.13 0.42 4.95 -2.29	<a href="#">result_17.pdb</a>
18	8574	1342.30	118.56	2.27 -0.07 0.36 -10.58 13.53 29.87	<a href="#">result_18.pdb</a>
19	8526	1039.50	-111.44	-0.65 0.10 -1.60 -11.41 7.24 13.05	<a href="#">result_19.pdb</a>
20	8524	1098.60	-27.75	0.71 -0.70 2.28 24.54 7.03 47.47	<a href="#">result_20.pdb</a>

[NEW: Jmol view](#)

**Download output**

**Send for refinement**

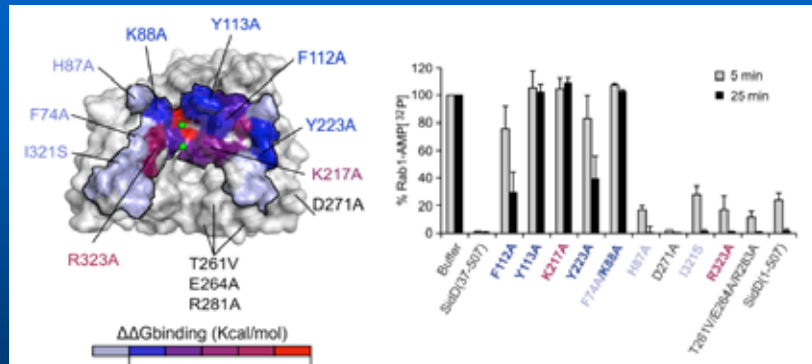
**DOWNLOAD best solutions as a ZIP file:**  (solutions number, from 2 to 100)  (this takes few seconds, please wait patiently)

**DOWNLOAD** [solutions table](#) [transformations file](#)

**REFINE best solutions with FireDock:**  (solutions number, from 1 to 1000)

- FFT-based grid search
- Geometric Hashing
- Adding distance-, symmetry- constraints

## Docking with known interface residues



## Optional parameters: binding site

Advanced Options:

[\[Show\]](#) [\[Hide\]](#)

Receptor Binding Site:   upload receptor binding site file

Ligand Binding Site:   upload ligand binding site file

Distance Constraints:   upload distance constraints file

## Optional parameters: binding site

Advanced Options:

[\[Show\]](#) [\[Hide\]](#)

Receptor Binding Site:

Browse...

upload receptor binding site file

Ligand Binding Site:

Browse...

upload ligand binding site file

Distance Constraints:

Browse...

upload distance constraints file

[Submit Form](#) [Clear](#)

**Binding Site** — list the residues of the receptor potential binding site. Only patches that include these residues will be used for matching.

Format: *[residue index] [chain ID]*

Example:

74 A

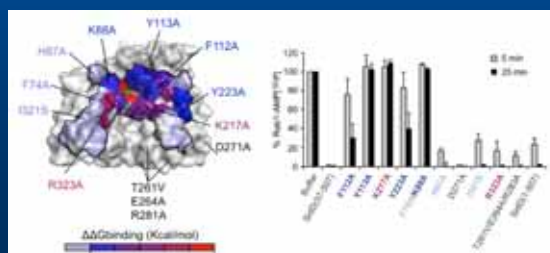
88 A

112 A

113 A

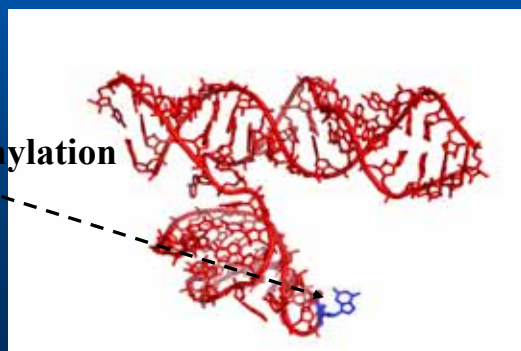
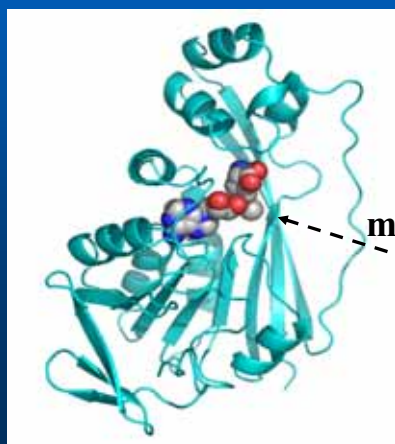
217 A

223 A



## Docking with Distance Constraints

A complex of Rlma2 methyltransferase of *S. pneumoniae* and a 74 nucleotide RNA transcript



methylation

## Optional parameters

Advanced Options:

[\[Show\]](#) [\[Hide\]](#)

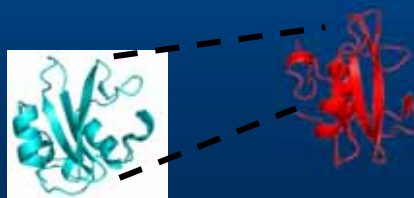
Receptor Binding Site:  Browse... upload receptor binding site file  
Ligand Binding Site:  Browse... upload ligand binding site file  
Distance Constraints:  Browse... upload distance constraints file

**Distance Constraints** —between pairs of atoms, one in the receptor and one in the ligand.

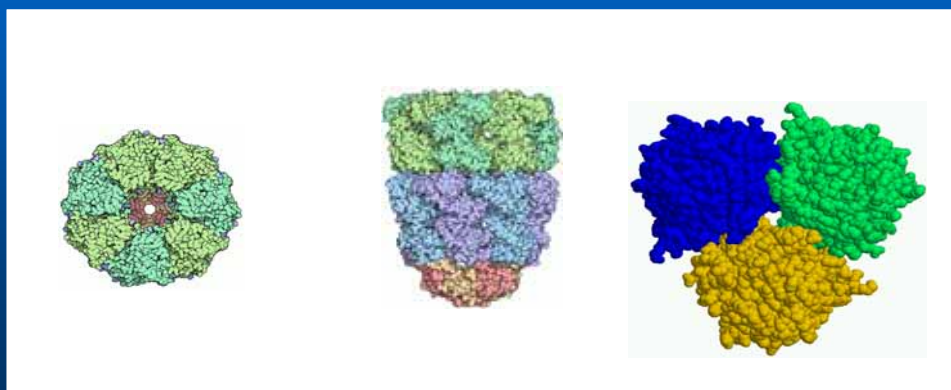
Format: *[receptor\_atom\_index] [ligand\_atom\_index] [max\_dist]*

For example:

25 377 5.0  
340 5603 10.0



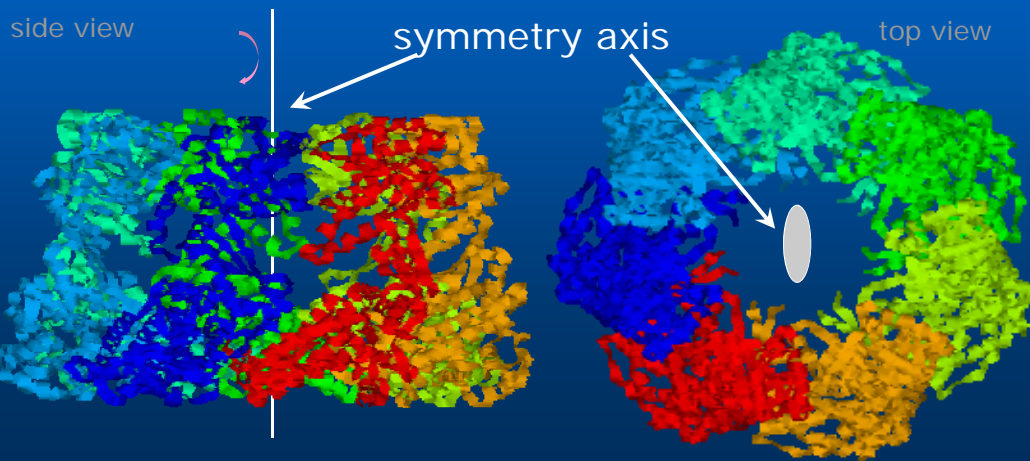
## Symmetric Docking



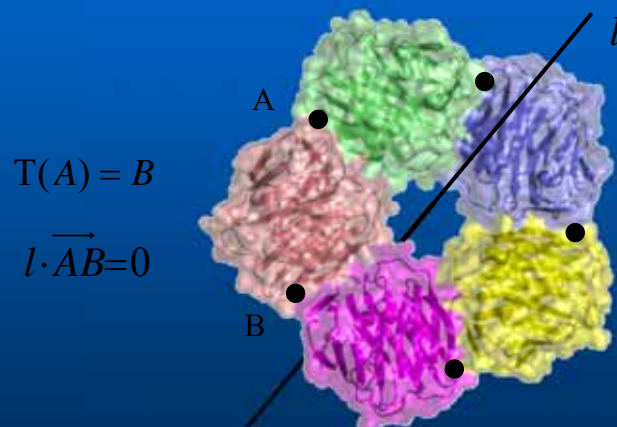
Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ Proteins 05

## Cyclic Symmetry

- Cyclic symmetry is defined by rotation of a single unit around an **axis**.
- The angle is determined by a number of units **n**.

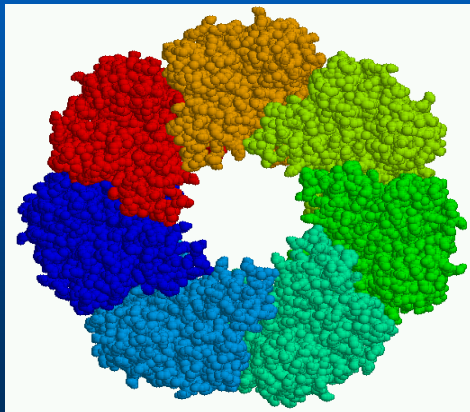


## Geometrical view

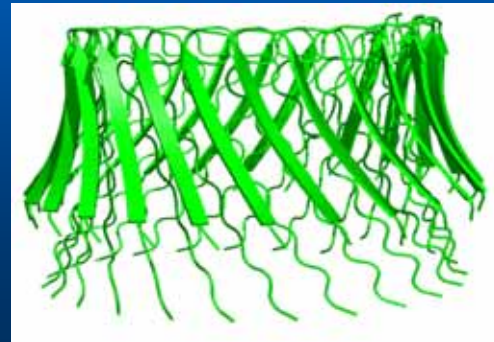




Chaperon: 2.5 Å RMSD prediction for the homo-heptamer.



Alzheimer's disease amyloid protein. Prediction of the membrane channel



N=24

WWW server: <http://bioinfo3d.cs.tau.ac.il/SymmDock>

# SymmDock

3 mandatory fields

**Unit Molecule:** the asymmetric unit of the multimer complex (PDB format)

Type the PDB code  (PDB-code:chain Id, e.g. 1f23:A)

or upload a file

Browse...

**Symmetry Order:**  (any number from 2 to 20: 3 for trimer, 4 for tetramer etc.)

**e-mail address**  (the results are sent to this address)

**Binding Site:**  Browse... binding site file (optional)

Option to specify interface residues  
(file format as in PatchDock)

Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. NAR 2005

# SymmDock Output Page

Top 20 solutions

Solution No	Score	Area	ACE	Transformation	PDB file of the complex
1	11858	1635.30	370.62	1.41 1.39 0.37 54.14 103.65 17.77	<a href="#">result.1.pdb</a>
2	10512	1545.10	292.99	2.13 0.50 -1.39 42.10 38.18 -0.03	<a href="#">result.2.pdb</a>
3	10468	1347.40	359.91	1.69 1.36 0.06 52.63 102.43 19.92	<a href="#">result.3.pdb</a>
4	10284	1731.50	118.78	0.33 1.04 1.71 67.49 68.41 -23.84	<a href="#">result.4.pdb</a>
5	9994	1808.40	228.47	1.59 1.28 0.26 45.22 104.90 15.19	<a href="#">result.5.pdb</a>
6	9748	1420.40	641.49	2.17 -0.52 0.70 -54.29 101.23 -6.25	<a href="#">result.6.pdb</a>
7	9478	1411.20	561.82	2.18 -0.53 0.68 -53.30 101.31 -2.72	<a href="#">result.7.pdb</a>
8	9312	1307.00	346.75	0.34 0.04 2.07 50.38 58.31 16.20	<a href="#">result.8.pdb</a>
9	9212	1528.30	761.87	2.03 0.43 0.17 9.12 119.03 -11.22	<a href="#">result.9.pdb</a>
10	9186	1312.60	36.15	0.58 0.96 1.60 57.31 63.89 -28.67	<a href="#">result.10.pdb</a>
11	9148	1444.50	99.71	0.12 1.13 1.88 67.67 63.51 17.33	<a href="#">result.11.pdb</a>
12	9130	1290.40	221.23	1.40 1.13 0.66 24.37 95.76 8.39	<a href="#">result.12.pdb</a>
13	9084	1105.10	319.00	2.41 0.83 -1.82 45.17 23.56 20.96	<a href="#">result.13.pdb</a>
14	8922	1070.60	306.21	2.11 0.43 -1.24 40.73 44.84 -2.10	<a href="#">result.14.pdb</a>
15	8888	1137.10	166.91	2.07 -0.14 0.72 -21.26 50.08 0.49	<a href="#">result.15.pdb</a>
16	8818	1190.60	101.84	-1.01 -1.29 -2.49 -27.92 40.88 20.18	<a href="#">result.16.pdb</a>
17	8798	1347.60	476.51	1.11 0.42 1.60 -1.25 53.16 -18.73	<a href="#">result.17.pdb</a>
18	8784	1452.70	375.98	-0.53 0.79 2.18 86.39 90.79 21.56	<a href="#">result.18.pdb</a>
19	8452	1410.70	499.66	2.08 -0.35 1.17 -52.48 58.60 19.25	<a href="#">result.19.pdb</a>
20	8452	1134.10	342.38	0.31 0.04 2.07 81.88 66.44 -18.55	<a href="#">result.20.pdb</a>

DOWNLOAD best solutions as a ZIP file.

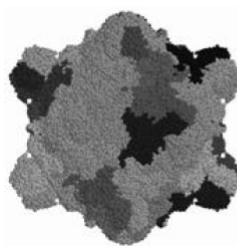
how many solutions to zip?  (any number from 2 to 100)

(this takes few seconds, please wait patiently.)

Download  
output

## Rigid-body docking: symmetry constraints

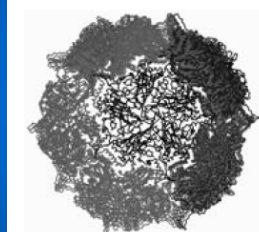
### Virus structure prediction



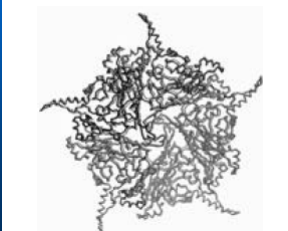
Icosahedron:  
12 nodes,  
20 faces,  
30 edges  
60 monomeric  
units

5-fold, 3-fold,  
2-fold cyclic  
symmetry

### Cyclic symmetry n=3



### Cyclic symmetry n=5





## Geometry-based docking: conclusions

- *Many available methods*
- *Very fast!*
- *In most of the cases, reasonable success rates*
- *Unbound much worse than bound*
- *Some times, missing native (even with bound subunits)*
- *Good as first docking step before refinement*