Within Structural Bioinformatics

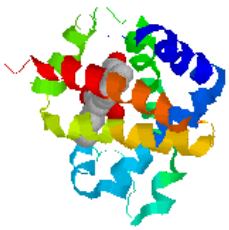
Plant Bioinformatics, Systems and Synthetic Biology Summer School
University of Nottingham, UK
July 2009

Eran Eyal
Cancer Research Center
Sheba Medical Center
Tel Hashomer
Israel

Sequence

```
VAL LEU SER PRO ALA ASP LYS THR ASN VAL LYS ALA ALA
TRP GLY LYS VAL GLY ALA HIS ALA GLY GLU TYR GLY ALA
GLU ALA LEU GLU ARG MET PHE LEU SER PHE PRO THR THR
LYS THR TYR PHE PRO HIS PHE ASP LEU SER HIS GLY SER
ALA GLN VAL LYS GLY HIS GLY LYS LYS VAL ALA ASP ALA
LEU THR ASN ALA VAL ALA HIS VAL ASP ASP MET PRO ASN
ALA LEU SER ALA LEU SER ASP LEU HIS ALA HIS LYS LEU
```

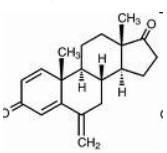
Structure



Function

Dynamics

Drug



Structural Bioinformatics

- **Databases of 3D structures of macromolecules**
- **Structural alignment**
- **Structural classification**
- **Secondary structure prediction**
- **Tertiary structure prediction**
- **Molecular docking**
- **Dynamics**

The structural data – where, what

Description of the databases

How to explore and query the data

Source of the data

Quality of the data

http://www.wwpdb.org//
http://www.rcsb.org/

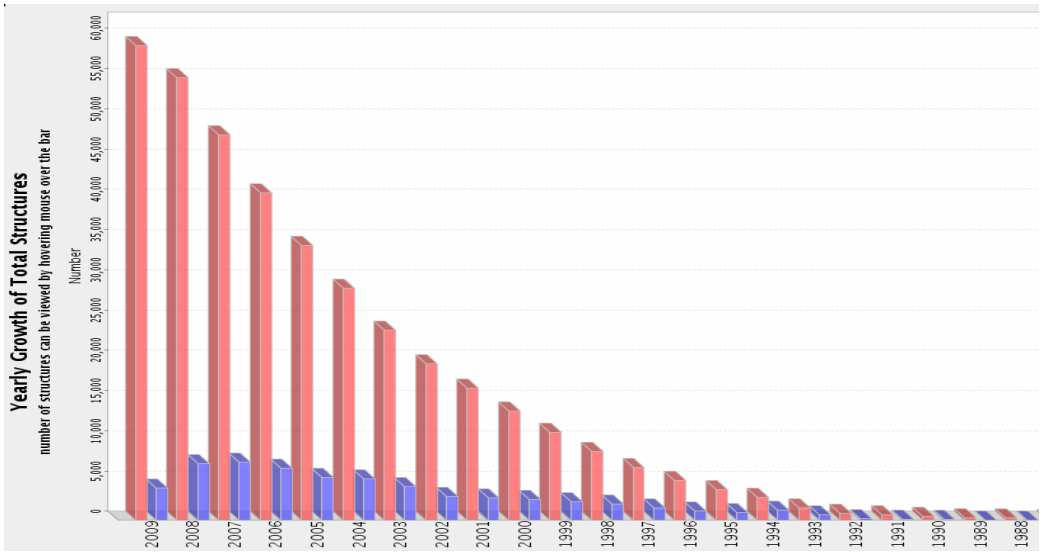
The PDB database is the main repository for the processing and distribution of 3-D biological macromolecular structures

Source of data:

- Crystal structures
- NMR models
- Other

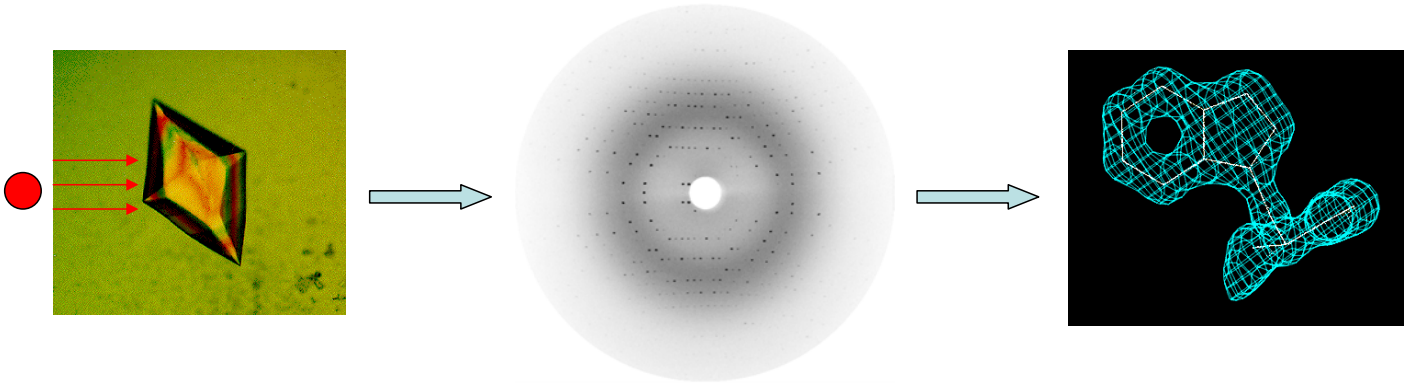
PDB Current Holdings Breakdown

		Proteins	Nucleic Acids	Protein/NA Complexes	Other	Total
Exp. Method	X-ray	47314	1157	2177	17	50665
	NMR	6946	868	147	6	7967
	Electron Microscopy	171	16	61	0	248
	Hybrid	14	1	1	1	17
	Other	113	4	4	9	130
	Total	54558	2046	2390	33	59027



Data Source

X-Ray Crystallography



Clone/Express/Purify
Crystallize

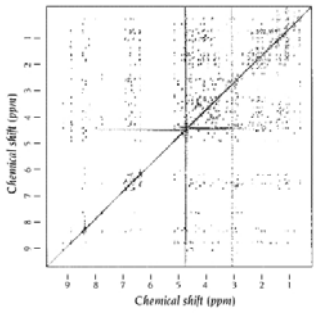
X-Ray diffraction data
+ Solve phase problem

Interpret electron
density map

→ Coordinates of atoms in protein molecule

Data Source

NMR Spectroscopy



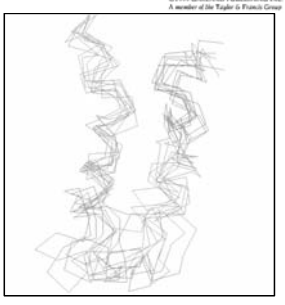
NOESY experiment
information about spatially-closed atoms

↓

J-couplings ↙

list of distance constraints + dihedral angles + …

↓



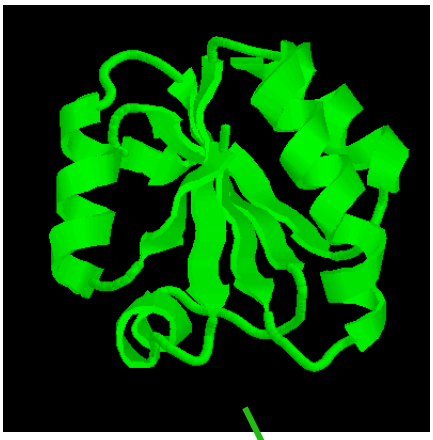
multiple models of protein structure

↓

Coordinates of atoms in protein molecule

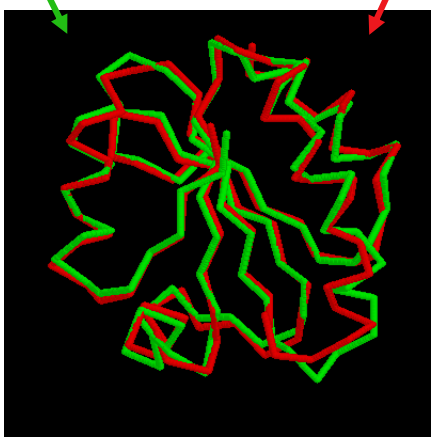
Human thioredoxin structure determined by X-ray and NMR

X-ray
(pdb 1ert)



NMR
(pdb 3trx)

superimposition



	X-ray crystallography	NMR
Atomic resolution	Good	Reasonable
Hydrogen	Rarely determined	Determined
Molecule size	No restriction	Small proteins
Dynamics	Snapshot	Multi models
Membrane proteins	Problematic	Problematic
Procedure	long	long

File Format

```
HEADER    IMMUNOGLOBULIN                          25-OCT-96   1IGT
COMPND    MOLECULE: IGG2A INTACT ANTIBODY - MAB231;
SOURCE    MOUSE (MUS MUSCULUS, STRAIN BALB/C)
KEYWDS    INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
EXPDTA    X-RAY DIFFRACTION
AUTHOR    L.J.HARRIS,S.B.LARSON,K.W.HASEL,A.MCPHERSON
REVDAT   1   07-JUL-97 1IGT    0
JRNL        AUTH   L.J.HARRIS,S.B.LARSON,K.W.HASEL,A.MCPHERSON
JRNL        TITL   REFINED STRUCTURE OF AN INTACT IGG2A MONOCLONAL
JRNL        TITL 2 ANTIBODY
JRNL        REF    BIOCHEMISTRY                  V.  36  1581 1997
JRNL        REFN   ASTM BICHAW  US ISSN 0006-2960               0033
...
REMARK   2 RESOLUTION. 2.8  ANGSTROMS.
...
SEQRES   1 A  214  ASP ILE VAL LEU THR GLN SER PRO SER SER LEU SER ALA
SEQRES   2 A  214  SER LEU GLY ASP THR ILE THR ILE THR CYS HIS ALA SER
SEQRES   3 A  214  GLN ASN ILE ASN VAL TRP LEU SER TRP TYR GLN GLN LYS
...
HET    NAG  D   1      26
HETNAM     NAG N-ACETYL-D-GLUCOSAMINE
FORMUL   5  NAG    8(C8 H15 N1 O6)
...
HELIX    1   1 PRO A   80  ASP A   82  5                                   3
...
SHEET    1   A 4 LEU A   4  SER A   7  0
SHEET    2   A 4 ILE A  19  HIS A  24 -1  N  HIS A  24   0  THR A   5
SHEET    3   A 4 GLY A  70  ILE A  75 -1  N  ILE A  75   0  ILE A  19
SHEET    4   A 4 PHE A  62  SER A  67 -1  N  SER A  67   0  GLY A  70
...
SSBOND   1 CYS A   23    CYS A   88
...
CRYST1   65.820   76.770  100.640  88.05  92.35  97.23 P 1          2
...
ATOM      1  N   ASP A   1       1.600 -85.453  44.624  1.00 43.02           N
ATOM      2  CA  ASP A   1       1.649 -84.304  45.569  1.00 38.99           C
ATOM      3  C   ASP A   1       0.334 -84.255  46.321  1.00 38.23           C
ATOM      4  O   ASP A   1      -0.652 -84.862  45.904  1.00 49.17           O
ATOM      5  CB  ASP A   1       1.826 -82.932  44.807  1.00 45.89           C
ATOM      6  CG  ASP A   1       3.124 -82.933  44.021  1.00 54.91           C
ATOM      7  OD1 ASP A   1       3.551 -83.966  43.460  1.00 58.03           O
ATOM      8  OD2 ASP A   1       3.713 -81.831  43.950  1.00 61.73           O
ATOM      9 1H   ASP A   1       0.744 -85.315  44.045  1.00 15.00           H
ATOM     10 2H   ASP A   1       2.453 -85.542  44.045  1.00 15.00           H
ATOM     11 3H   ASP A   1       1.443 -86.304  45.196  1.00 15.00           H
ATOM     12  N   ILE A   2       0.340 -83.582  47.459  1.00 34.73           N
ATOM     13  CA  ILE A   2      -0.861 -83.420  48.270  1.00 29.53           C
...
TER    2034      CYS A 214
HETATM12748  C1  NAG D   1      15.179   2.409 -21.411  1.00138.48           C
HETATM12749  C2  NAG D   1      14.181   1.365 -20.856  1.00137.31           C
HETATM12750  C3  NAG D   1      14.918   0.088 -20.438  1.00133.46           C
...
CONECT  194  193  814
...
END
```

Header section

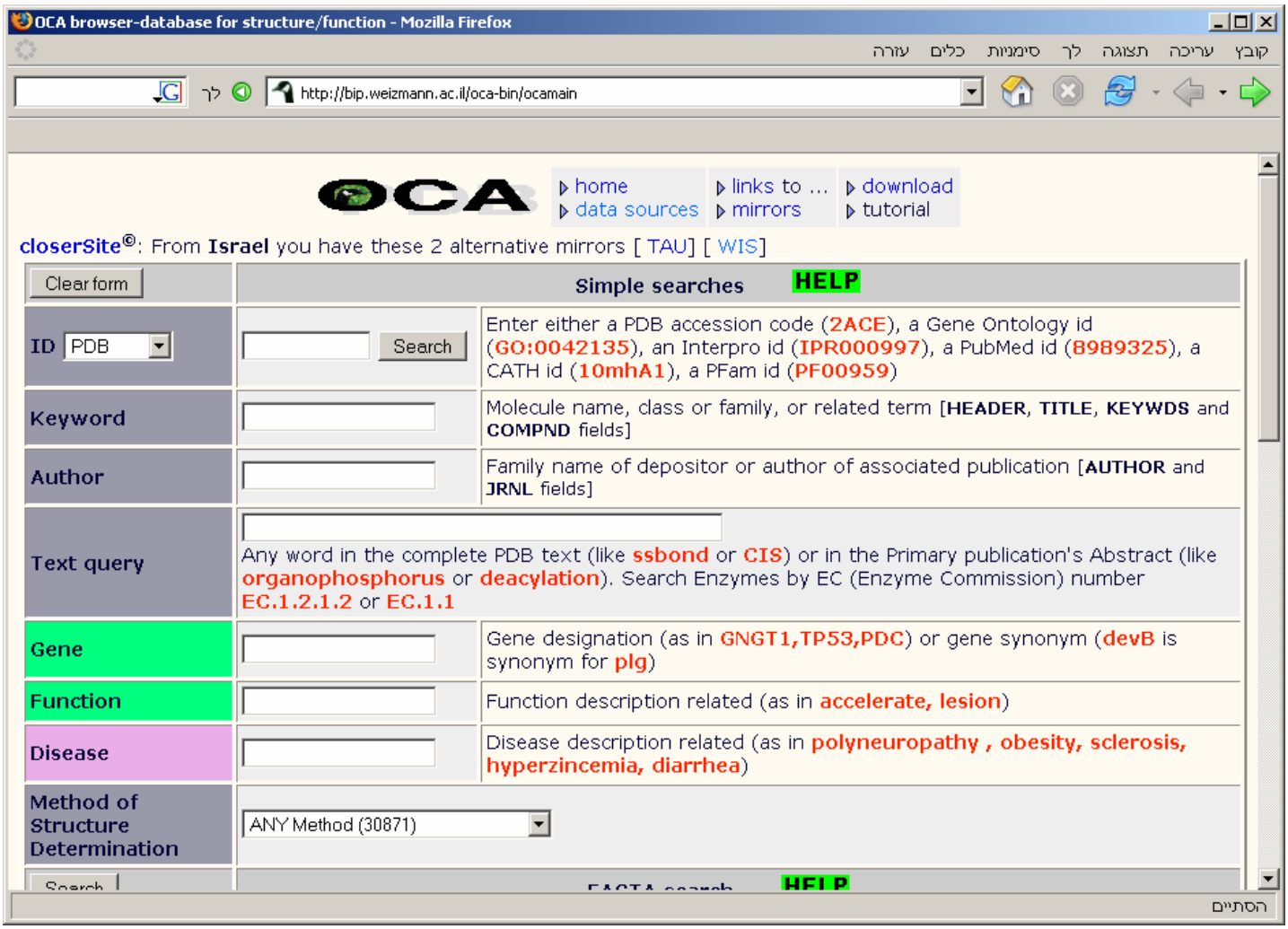
Coordinate section

How to search in the PDB?

The OCA browser developed in the WIS by Jaime Prilusky is one of the best interfaces to the PDB.

Entries can be retrieved by variety of criteria

http://bip.weizmann.ac.il/oca-bin/ocamain

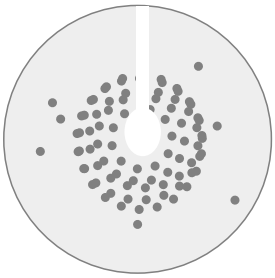


Problems in the PDB database

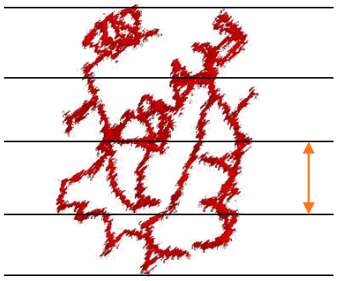
- Missing data
- Quality of data
- Data is often not independent
- Format problems – residue numbers

Diffraction pattern

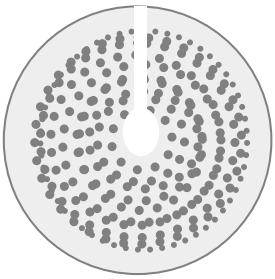
Brag Planes separation



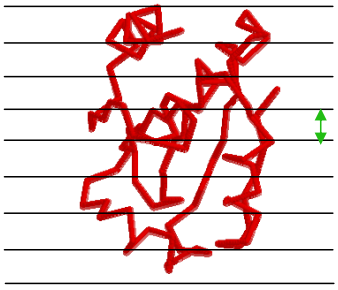
Poor Resolution



3Å resolution



Good Resolution

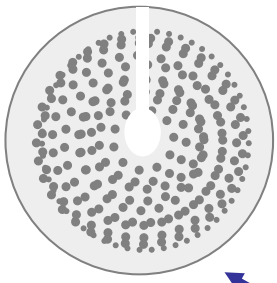


2Å resolution

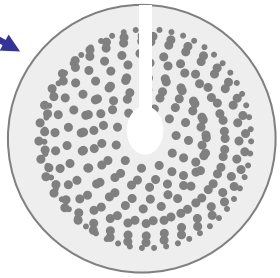
R-factor

Original diffraction pattern

Model



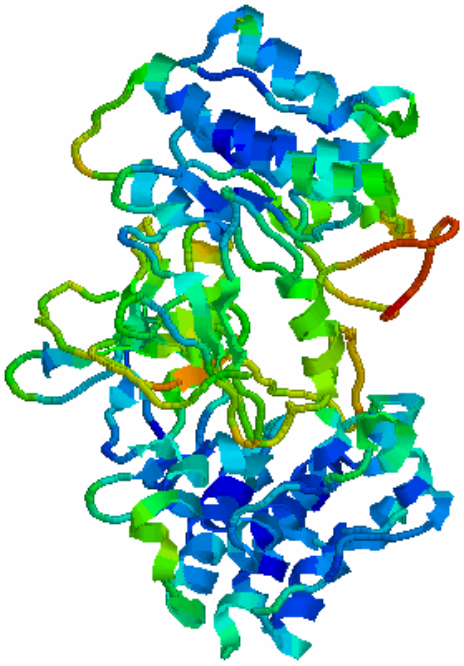
Calculated diffraction pattern based on the model



R factor measures how different is the originated diffraction map from a recalculated one based on a putative model

B-factor

A measure of the uncertainty in the position of individual atoms



TYROSINE-PROTEIN KINASE color by B-factor

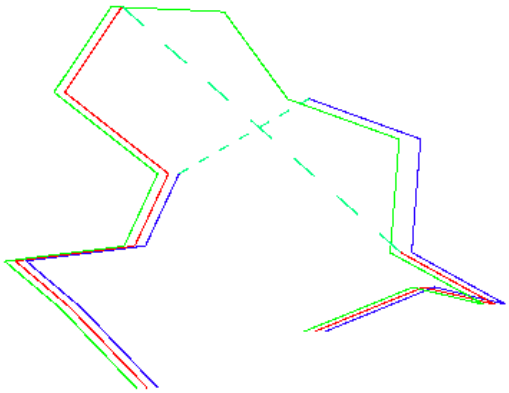
Structural alignment

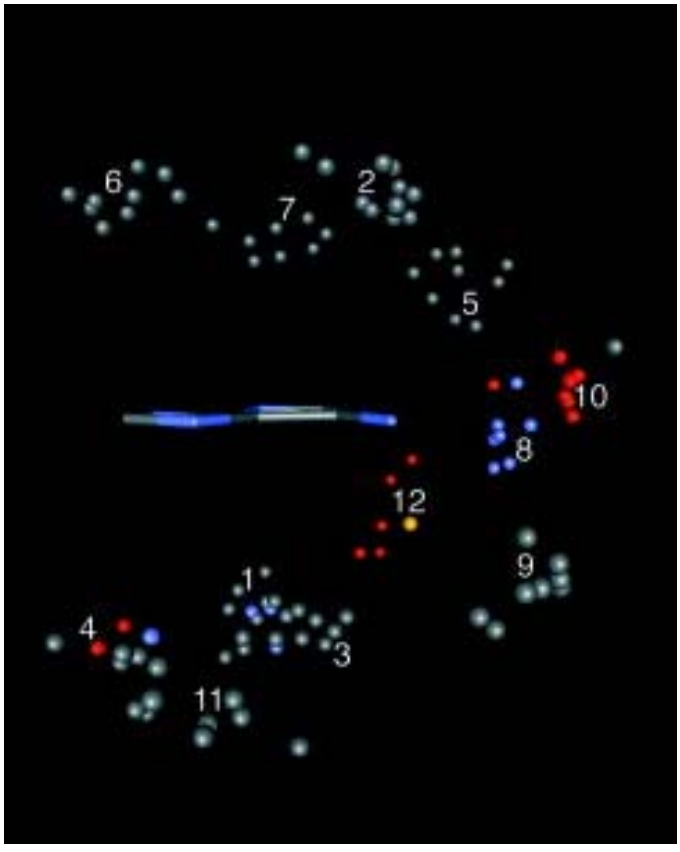
Structures are more conserved throughout evolution than sequences. Two homologous proteins have the same overall structure. It is possible that 2 proteins without detectable sequence similarity will have the same structure.

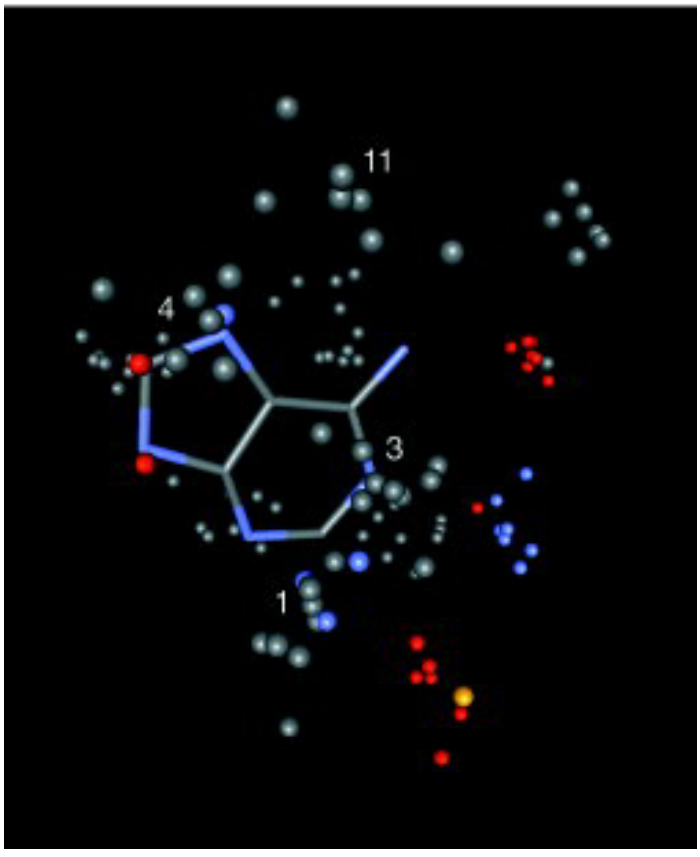
In the twilight zone of sequence similarity, structural alignment might help to correctly determine the relations between 2 proteins

Structural similarity is more sensitive method than sequence alignment to determine protein function

```
1   2    3    4    5    6    7    8    9    10   11   12   13   14
PHE ASP ILE CYS ARG LEU PRO GLY SER ALA GLU ALA VAL CYS
PHE ASN VAL CYS ARG THR PRO --- --- --- GLU ALA ILE CYS
PHE ASN VAL CYS ARG --- --- --- THR PRO GLU ALA ILE CYS
```







Kuttner et al,. 2003

Superposition ⇔ Structural alignment

There are two types of problems related to structural comparison:

- Superposition problem
- Structural alignment problem

In the superposition problem we know in advance the correspondence between the points in the two structures we want to align

What properties of the protein might be used to detect structural similarity to other proteins ?

- sequence
- Type and number of secondary structures (sheets, helices)
- Structural arrangement of secondary structures
- Structural attributes of individual amino acids
- Distances between amino acids in the protein

The standard way to quantify similarity between molecules is to measure the positional deviation of the atoms - RMSD

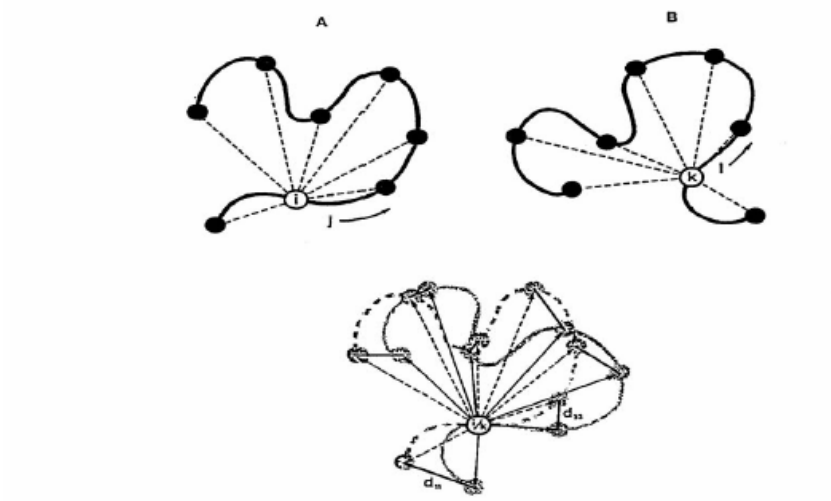
RMSD = root mean square deviation

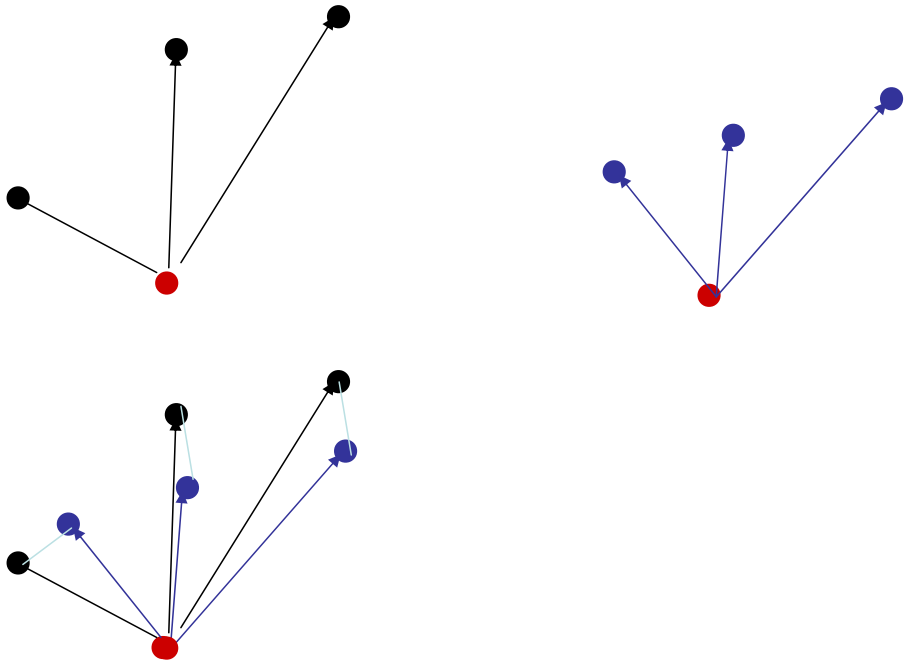
$$RMSD = \sqrt{\frac{\Sigma (X_{i1}-X_{i2})^2+(Y_{i1}-Y_{i2})^2+(Z_{i1}-Z_{i2})^2}{N}}$$

This method amplifies large deviation in local region of the protein

SSAP

http://www.biochem.ucl.ac.uk/cgi-bin/cath/GetSsapRasmol.pl





The method includes 2 steps of dynamic programming. Initial step to obtain the score between each pair of amino acids, and second step in which the best overall alignment in the protein is determined

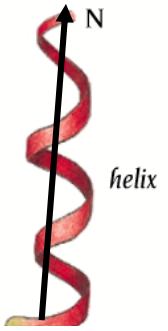
SARF2

http://123d.ncifcrf.gov/sarf2.html

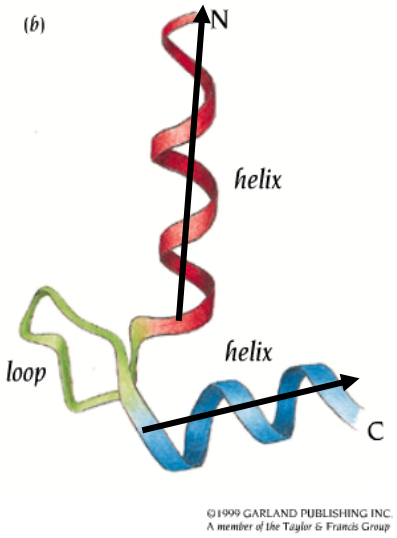
http://carten.gmd.de/ToPign.html

An algorithm to find structural similarity based on comparison of secondary structures.

As such it might be used to compare proteins only, and only proteins with minimal content of defined secondary structures

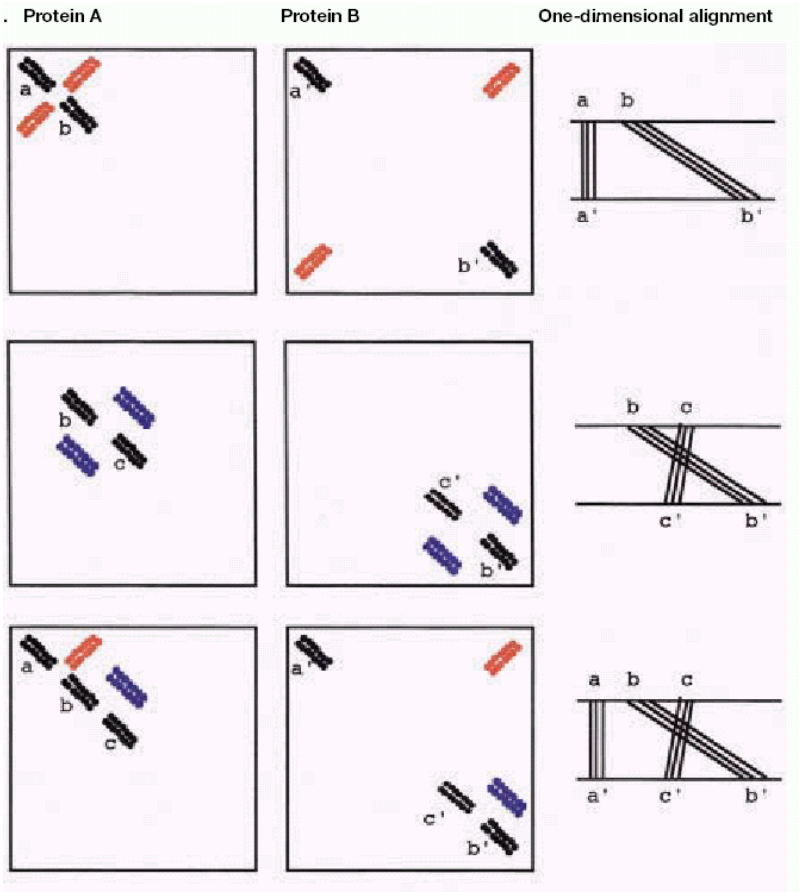


every secondary structure element is represented by a vector



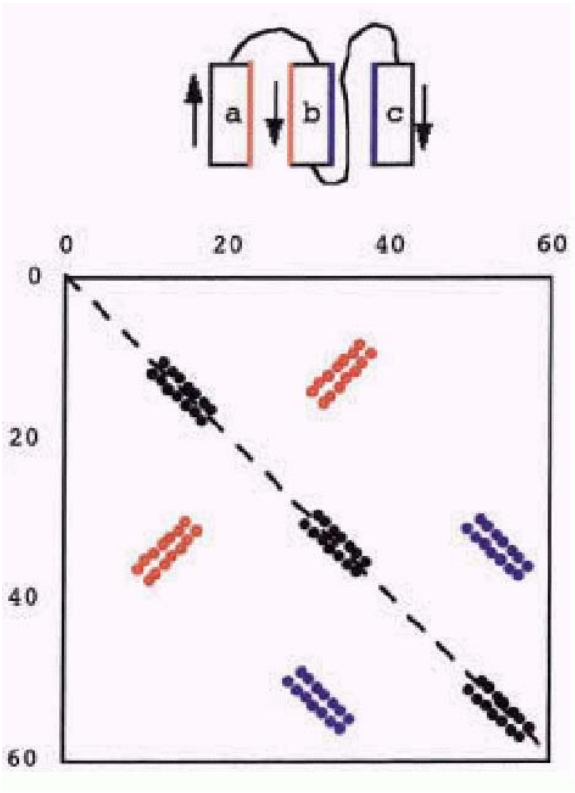
Single SSE does not give any information about the structure of the protein. Two SSEs or more are therefore required.

3-helix-bundle pairwise 3D alignment



DALI: Search for common 3D-pattern of **C_α distance maps**

http://www.ebi.ac.uk/dali/



Structural classification

•Using structural alignment it is feasible to construct a classification system

•Classification helps us understand relations between remote proteins

•Convergence evolution in structures can often hint to the function of the protein

Classification databases

FSSP	**http://www2.ebi.ac.uk/dali/fssp/**
CATH	**http://www.biochem.ucl.ac.uk/bsm/cath_new/index.html**
SCOP	**http://scop.mrc-lmb.cam.ac.uk/scop/**
HOMSTRAD	**http://www-cryst.bioc.cam.ac.uk/data/align/**
MMDB	**http://www.ncbi.nlm.nih.gov/Structure/MMDB/mmdb.shtml**
3Dee	**http://www.compbio.dundee.ac.uk/3Dee/**
CE	**http://cl.sdsc.edu/ce.html**
VAST	**http://www.ncbi.nlm.nih.gov/Structure/VAST/vast.shtml**
SARF	**http://www-lmmb.ncifcrf.gov/~nicka/sarf2.html/**

CATH

http://www.cathdb.info/

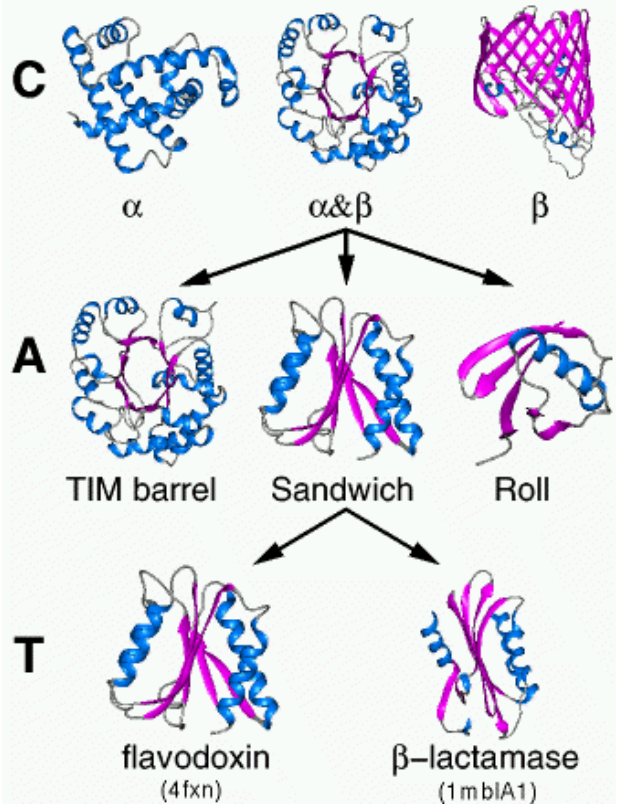
Semi-automatic!

Class – 2D composition – automatic. 4 classes: α, β, αβ, FSS: few 2D structures.

Architecture – manual! Shape created by orientation of 2D units.

Topology – secondary structures connectivity.

Homologous superfamily – high structural and functional similarity. Sequence similarity

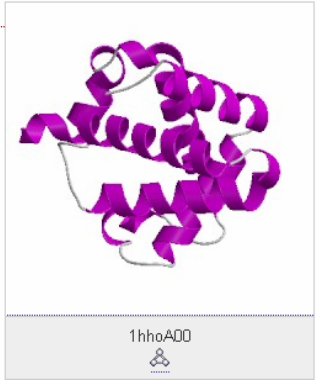


CATH of 1hho

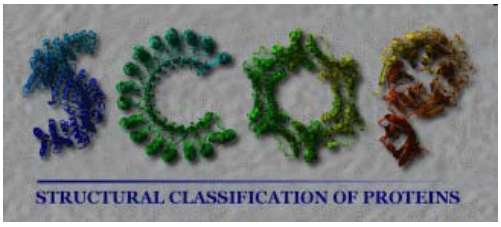
CATH Domain: 1hhoA00 XML

PDB 1hho, Chain A, Domain 0

CATH Code	Level Description	Links
1	Mainly Alpha	
1.10	Orthogonal Bundle	
1.10.490	Globin-like	
1.10.490.10	Globins	[Gene3D]
1.10.490.10.5		
1.10.490.10.5.1		
1.10.490.10.5.1.1		
1.10.490.10.5.1.1.1		
1.10.490.10.5.1.1.1.142		[Gene3D]



SCOP – Structural Classification of Proteins



Manual inspection of automatic output

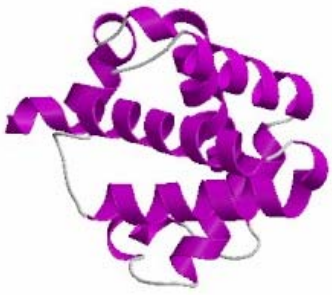
1. 2D content (class)
2. Structural similarity (fold)
3. Remote homology (superfamily/family).
4. Close homology (family)

Structure Prediction

- Databases
- Structural alignment
- Structural classification
- Secondary structure prediction
- Tertiary structure prediction
- Molecular docking
- Visualization
- Dynamics

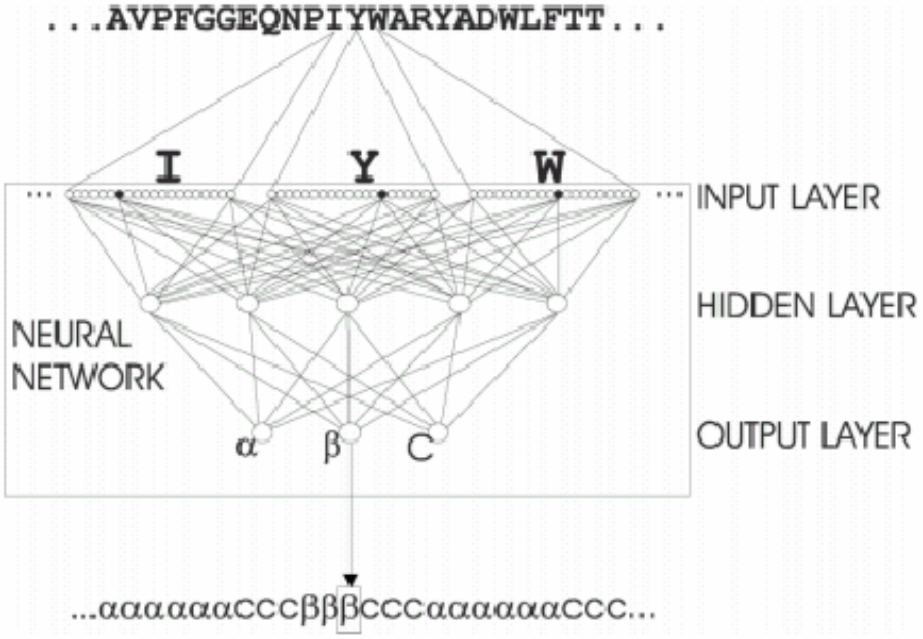
A-C-H-Y-T-T-E-K-R-G-G-S-G-T-K-K-R-E-A

H-H-H-H-H-H-H-H-0-0-0-0-0-S-S-S-S-S-S

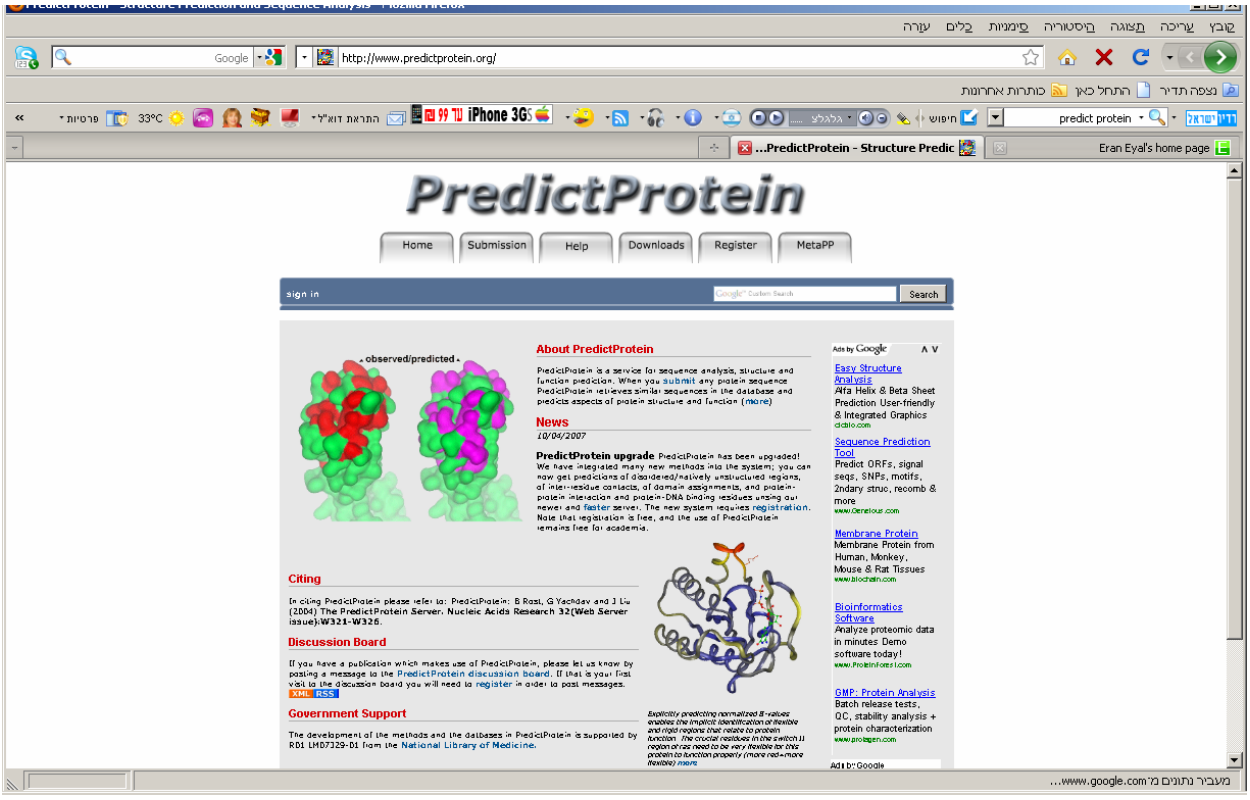


Secondary structure prediction

Tertiary structure prediction



http://www.predictprotein.org/



Why make a structural model for your protein ?

• The structure can provide clues on the function

• With a structure it is easier to guess the location of functional sites and to learn on **the function**

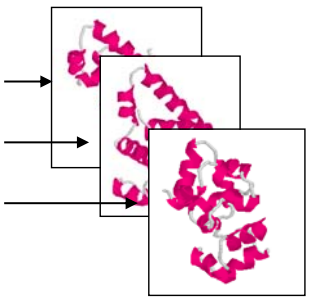
• With a structure we can plan more precise experiments in the lab

• We can do docking experiments (both with other proteins and with small molecules)

Building by homology (Homology modeling)

alignment with proteins of known structure

M	**A**	**A**	**G**	**Y**	**A**	**Y**	**G**	**V**	**L**	**S**
-	A	T	G	F	D	-	-	V	I	D
-	A	S	G	F	E	-	-	V	V	E
-	A	K	A	Y	L	-	-	V	L	S



structural model

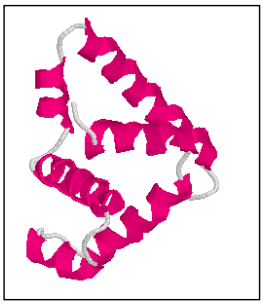
Fold recognition (Threading)

sequence:

M	A	A	G	Y	A	V	L	S

+

known protein folds



→

structural model

Ab initio

sequence

M	A	A	G	Y	A	V	L	S

→

structural model

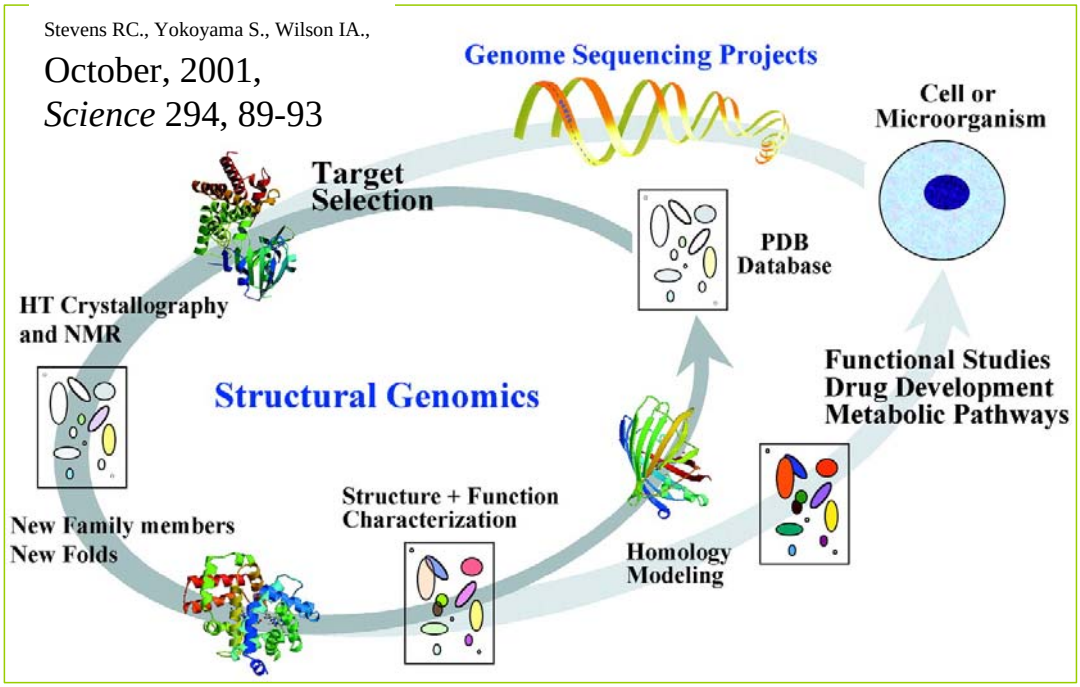
Building by homology

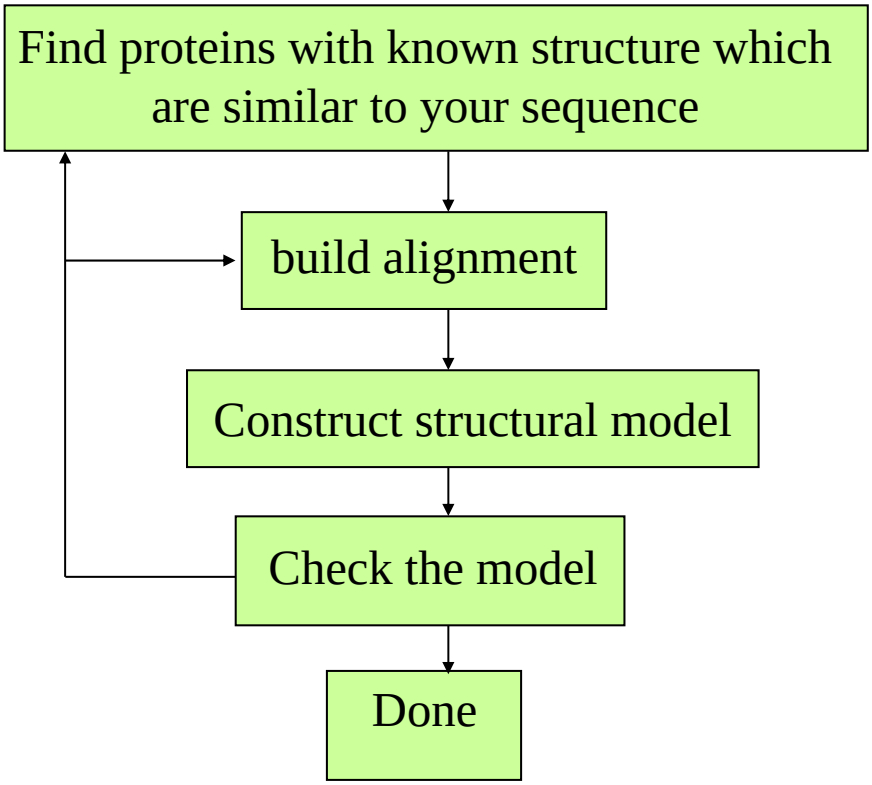
There are hundreds of thousands of protein sequences but only several thousands protein folds

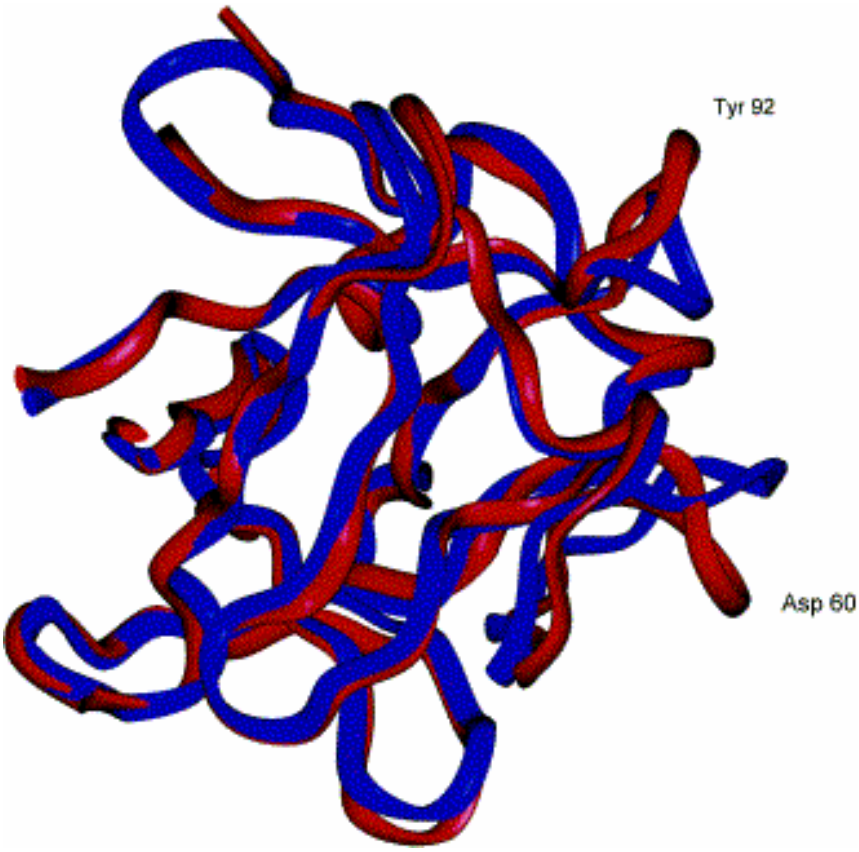
For every second protein that we randomly pick from the structural data base there is “close” homolog (identity > 30%). This homolog almost always has the same fold.

In the current projects for experimental determination of protein structures, priority is given to determine structures of protein without homologs in the structural databases (‘structural genomics’)

We believe that in several years we will have almost all the basic folds

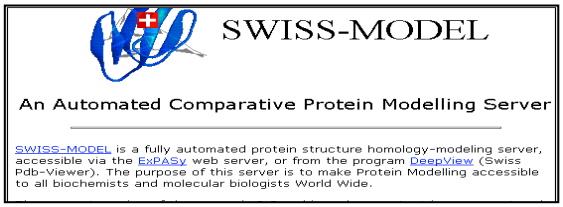



Find proteins with known structure which are similar to your sequence
build alignment
Construct structural model
Check the model
Done


Tyr 92
Asp 60

Swiss-Model

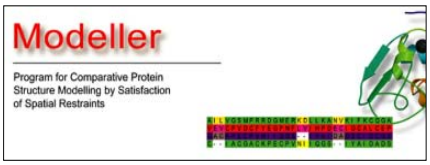
http://www.expasy.ch/swissmod/SWISS-MODEL.html



"Quick and dirty"
The easiest way to do homology modeling

Modeller

http://salilab.org/modeller/



Advanced program for homology modeling.
Implemented in several popular modeling packages such as InsightII

Threading (fold recognition)

The input sequence is threaded on many different folds from library of known folds

Using scoring functions we get a score for the compatibility between the sequence and each structure

Statistically significant score tells that the input protein adopts similar 3D structure to that fold

This method is less accurate but could be applied for more cases

When the “real” fold of the input sequence is not represented in the structural database we can never get correct solution by this method

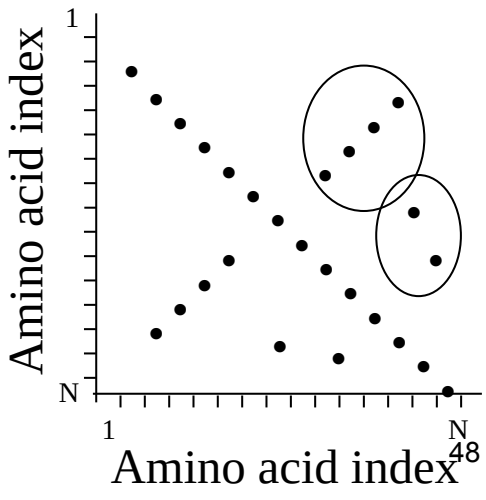
The most important part is the ***accuracy of the scoring function***. The scoring function is the major difference between different programs for fold recognition

Contact potentials

This method is based on predefined tables which include pseudo-energetic scores to each pair-wise interaction of two amino acids.

For each given conformation to be evaluated, a distance matrix can be constructed.

For each pair of amino acids which are close in space the interaction energy is summed. The total is the indication for the fitness of the sequence into that structure

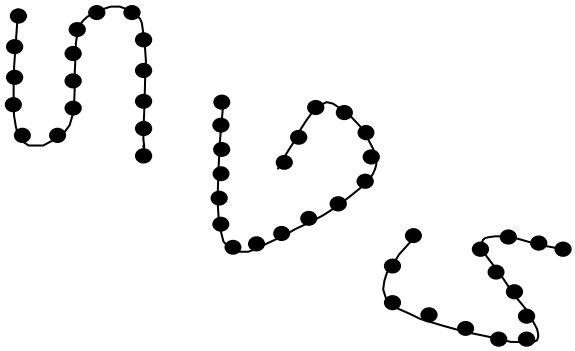


Input:

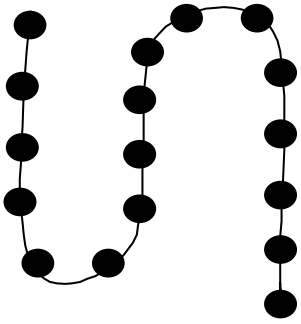
sequence

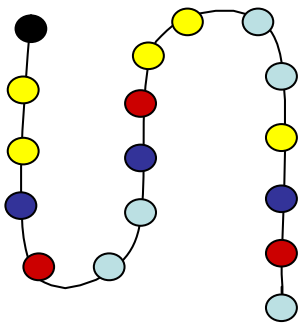
- H bond donor
- H bond acceptor
- Glycin
- Hydrophobic

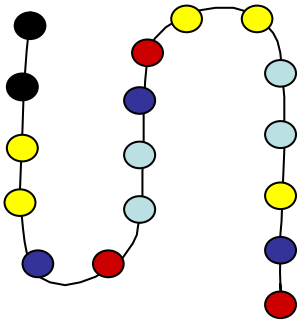
Library of folds of known proteins

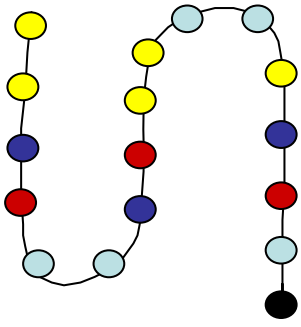


- H bond donor
- H bond acceptor
- Glycin
- Hydrophobic









S=-2
Z= -1

S=5
Z=1.5

S=20
Z=5

Ab initio methods for modeling

This field is of great theoretical interest. Here there is no use of sequence alignments and no direct use of known structures

The basic idea is to build empirical function that simulates real physical forces and potentials of chemical contacts

If we will have perfect scoring function and we will be able to scan all the possible conformations, then we will be able to detect the correct fold

Algorithms for ***Ab initio*** prediction include:
A. Searching procedure that scans many possible structures (conformations)
B. Scoring function to evaluate and rank the structures

Due to the large search space, heuristic methods are usually applied

The parameters in the searching procedure are the dihedral angles which specify the exact fold of the polypeptide chain

Side chain construction

When there is high similarity between the built protein and the templates, construction of the side chains is done using the template structures

Without such similarity the construction can be done using rotamer libraries

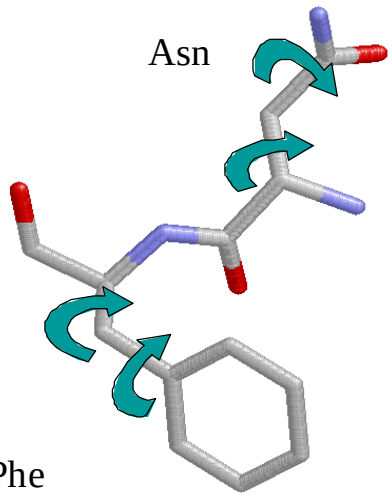
A compromise between the probability of the rotamer and its fitness in specific position determines the score. Comparing the scores of all the rotamers for a given amino acid determines the preferred rotamer.

Conformation - a given set of dihedral angle which defines a structure.

Rotamer - energetically favourable conformation.

Example of a rotamer library:

	χ^1	χ^2	probability
SER	59.6		1.0
SER	-62.5		26.4
SER	179.6		32.6



TYR	63.6	90.5	21.0
TYR	68.5	-89.6	16.4
TYR	170.7	97.8	13.3
TYR	-175.0	-100.7	20.0
TYR	-60.1	96.6	10.0
TYR	-63.0	-101.6	19.3

SCCOMP

SCCOMP is a program for side chain modeling. It uses a scoring function that includes terms for complementarity (geometric and chemical compatability), excluded volume, internal energy based on probability of rotamers, and solvent accessible surface. The program has an accuracy of 92-94% for correct Chi1 prediction (± 40°) of buried residues, 82-84% for all residues, and about 1.7 Å for overall rmsd (not including Cβ). The exact values depend on the searching procedure. A fast iterative search, takes about a minute on the web server for a typical protein. A slower stochastic method takes about 12 minutes and improves the prediction by about 2% and 0.1 Å rmsd.

The program also permits:
1. Modeling only a subset of residues
2. Performing any number of mutations
3. Using an homologous structure as a template.

At this site you can download the source code for sccomp for different Unix/Linux platforms. You can also use our web server for modeling. The results are sent back by Email.

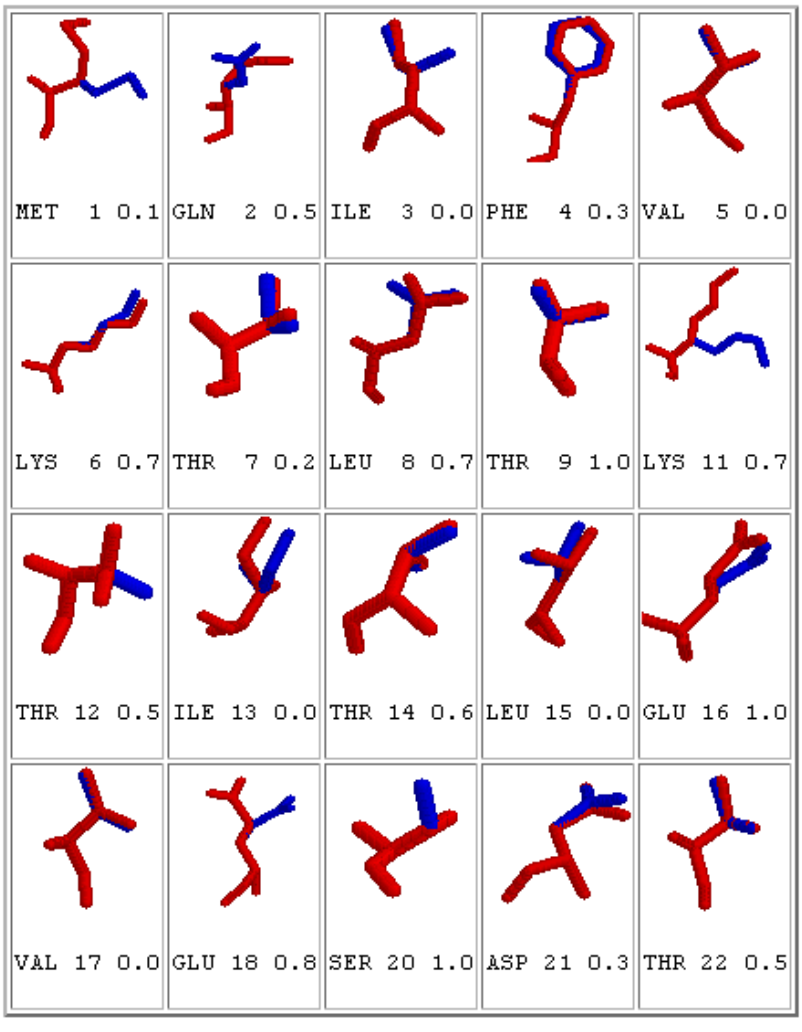
get sccomp source code

To use the web server follow the instructions below. You must fill one of the two first windows, either to submit your own coordinates or to specify a PDB id.

Every residue should include at least the backbone coordinates. The order of the atoms should follow that of the PDB format. The file should end with ".pdb" .

Enter your protein file.

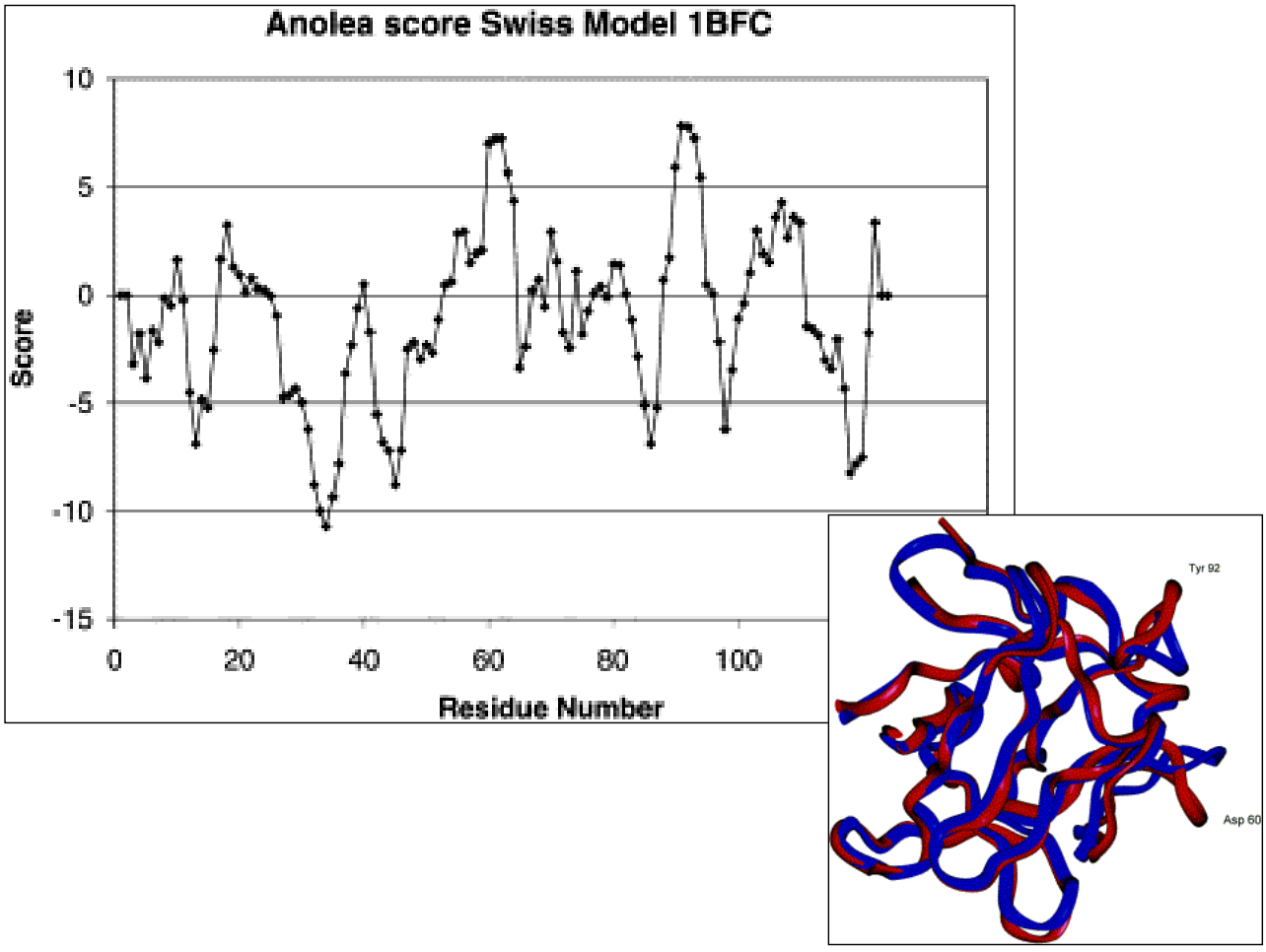
http://ignmtest.ccbb.pitt.edu/cgi-bin/sccomp/sccomp1.cgi



Model evaluation

After the model is built we can check its validity by various ways. We can check that the model has a reasonable shape and that it is usually obey geometric constraints.

If the model turns out to be bad, it is necessary to repeat several steps of the model building



We can easily assess homology modeling procedures by building models for proteins which have already solved structure and compare between the model and the native structure

It is always possible that information from the native structure will be used in direct or indirect ways for model building

A more objective test is prediction of structures before they are publicly distributed (this is the idea of the CASP competitions)

Docking: finding the binding orientation of two molecules with known structures

According to the molecules involved:
•Protein-Ligand docking
•Protein-Protein docking

Specific docking algorithms usually designed to deal with one of these problems but not with both (different contact area, flexibility, level of representation, etc.)

Why docking?

- Understanding interactions, roles of specific amino acids, design of mutations and changes of activity.

- Prediction affinities

- Drug design

Ligand-Protein docking

Finding the place and the orientation of the interactions

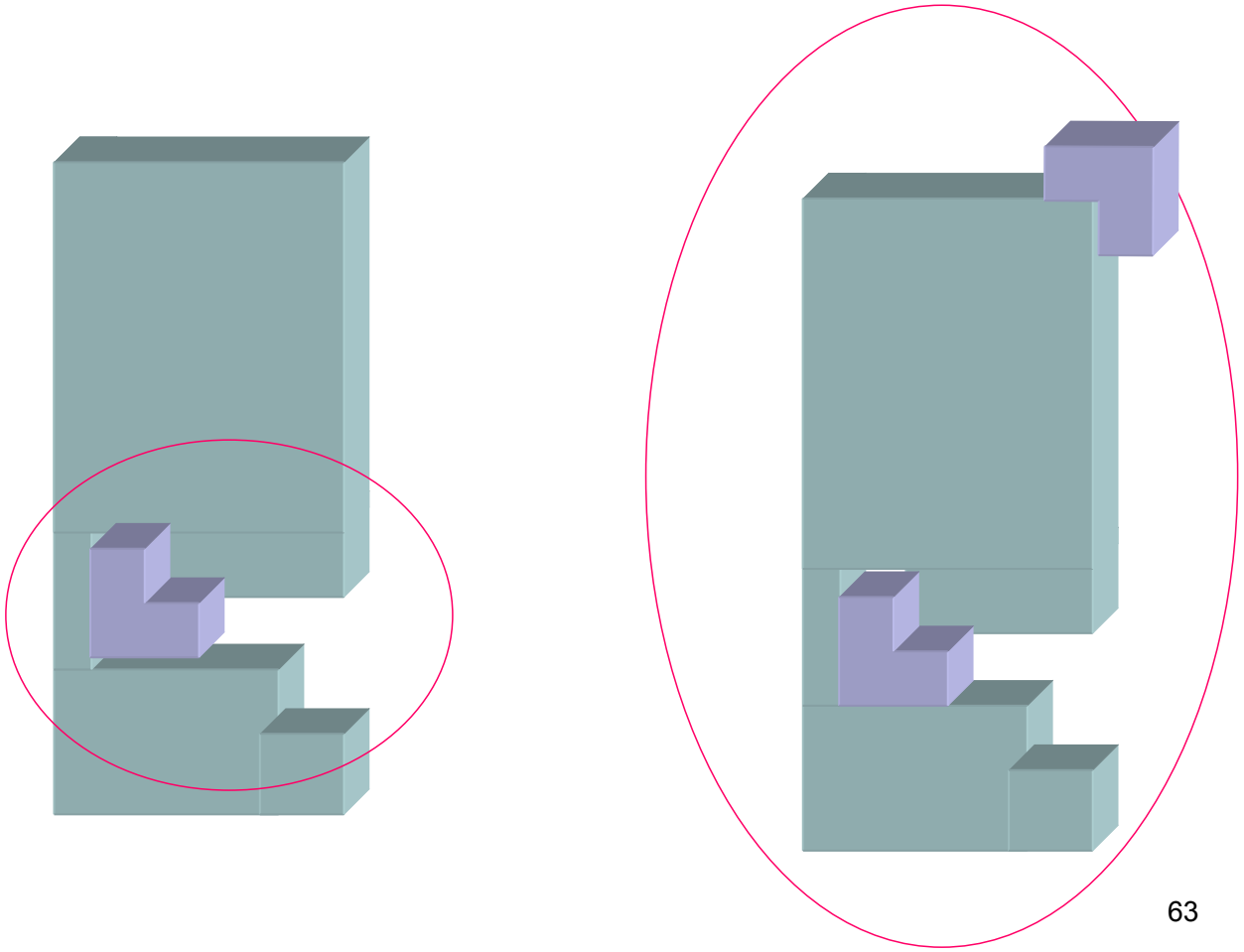
The general problem includes a search for the location of the binding site and a search to figure out the exact orientation of the ligand in the binding site. A program that do both makes a ***Global docking***

Sometimes the location of the binding site is known. In this case we only need to orient the ligand in the binding site. In this case the problem is called ***Local docking***

Global docking is more demanding in terms of computational time and the results are less accurate.

Local docking

Global docking



Rigidity vs flexibility

• Most of the early algorithms assumed that the docked molecules do not change conformations. This assumption allows to treat the molecules as rigid bodies, making the algorithm simpler and faster

• This assumption is problematic and was proven to be wrong in many cases

• New algorithms try to face the flexibility problems.

• Other methods try to handle the flexibility problem indirectly or at least to “minimize the damage” of not incorporating flexibility.

• Docking procedures that perform rigid body search are termed ***rigid docking***

• Docking procedures that consider possible conformational changes are termed ***flexible docking***

Bound and unbound docking

In ***bound docking*** the goal is to reproduce a known complex where the starting coordinates of the individual molecules are taken from the crystal of the complex

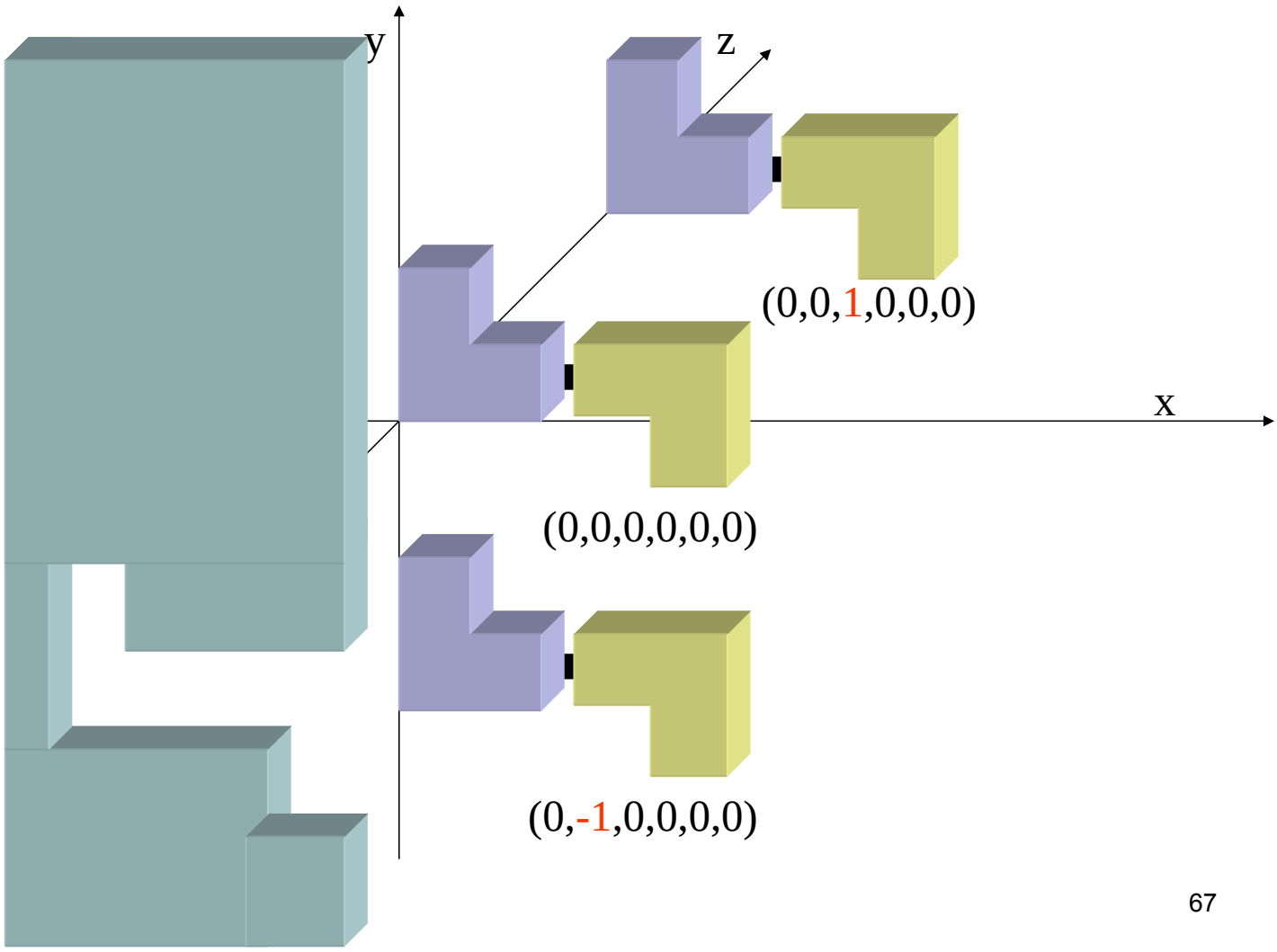
In the ***unbound docking***, which is a significantly more difficult problem, the starting coordinates are taken from the unbound molecules. This is unfortunately a more realistic problem.

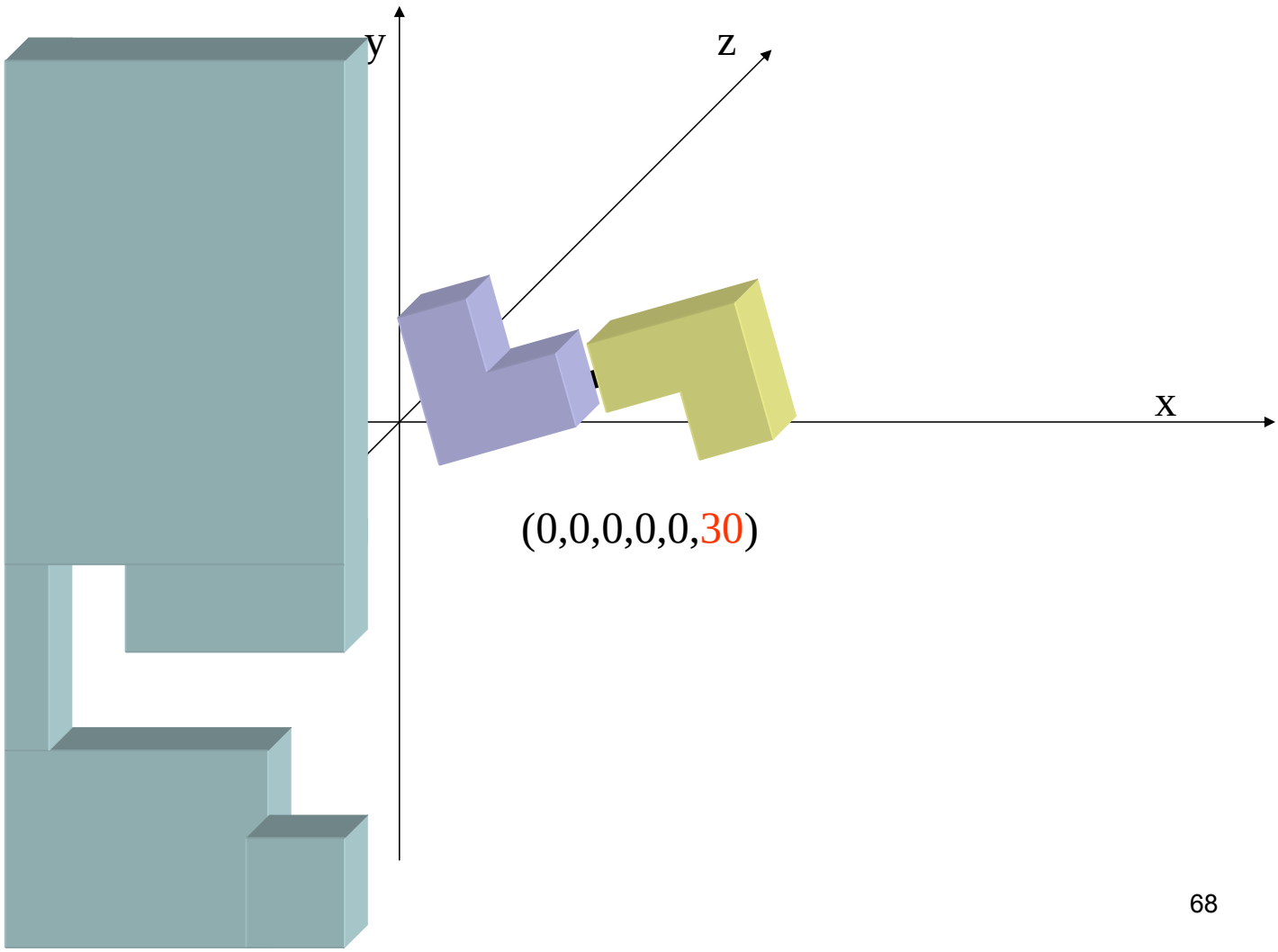
Components of the problem

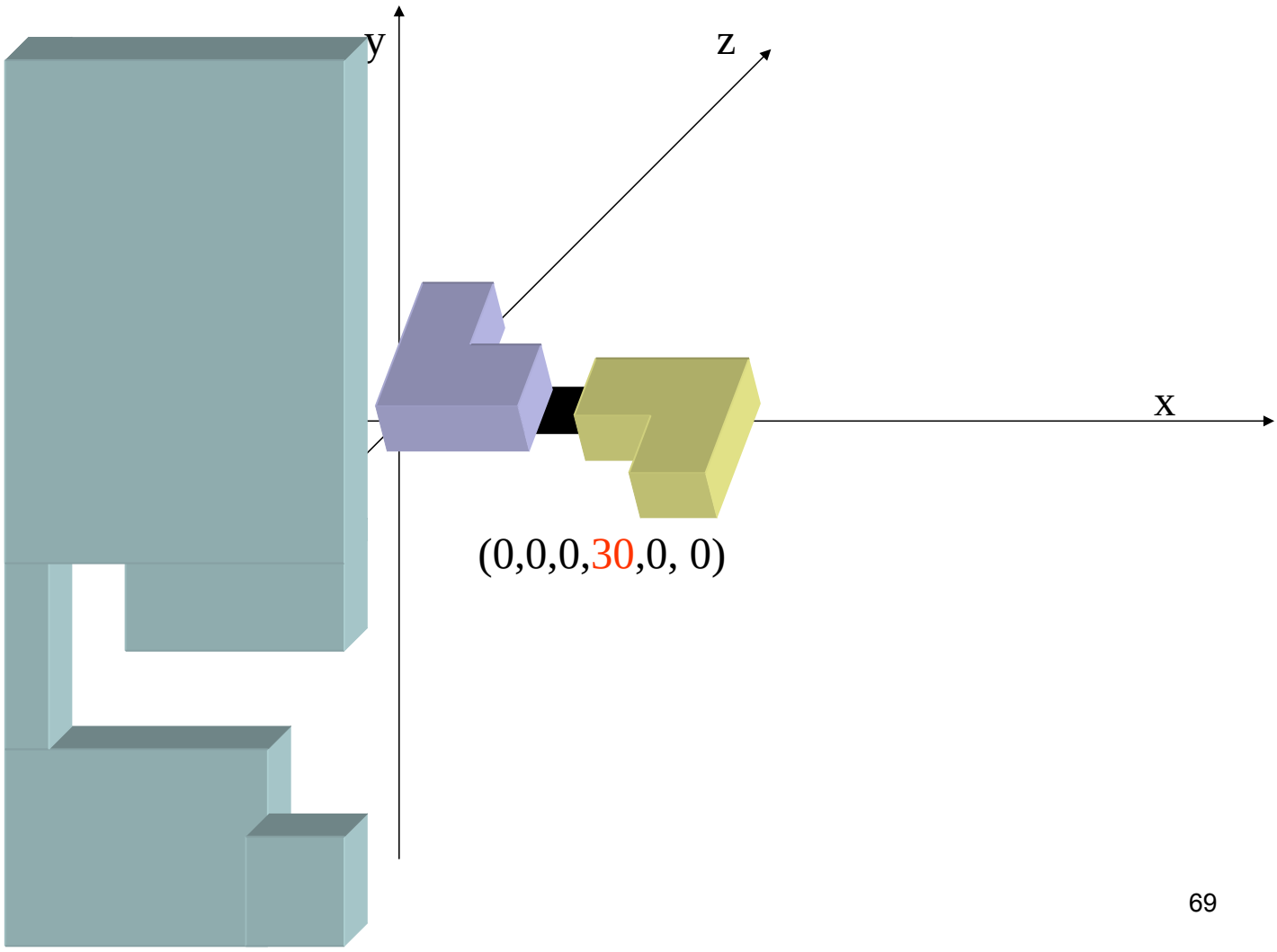
Algorithms to dock molecules need:

A. System representation
B. Searching procedure
C. Scoring function
D. Clustering procedure

The parameters of the problem for docking of 2 rigid bodies are 3 angles (rotations) and 3 distances (translations)


y
z
x
(0,0,1,0,0,0)
(0,0,0,0,0,0)
(0,-1,0,0,0,0)


y
z
x
(0,0,0,0,0,30)



Usually the ligand is not rigid and few other parameters are required

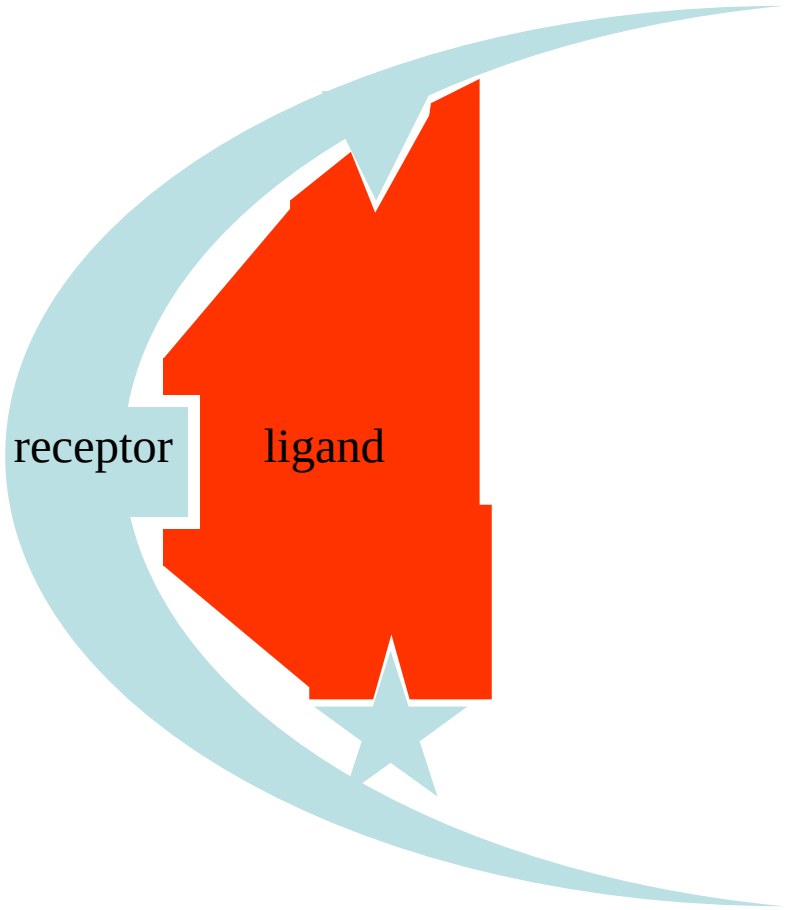
$$N_p = 3 + 3 + N_{fb}$$

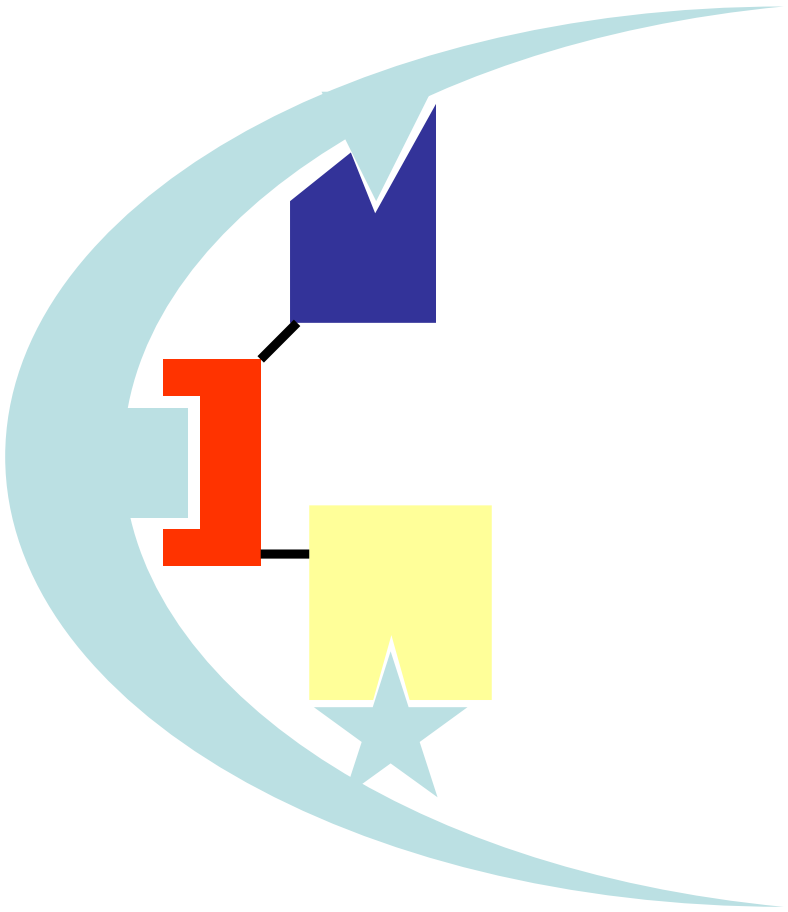
Number of parameters needed to fully describe ligand position

Position

Orientation

Number of flexible bonds





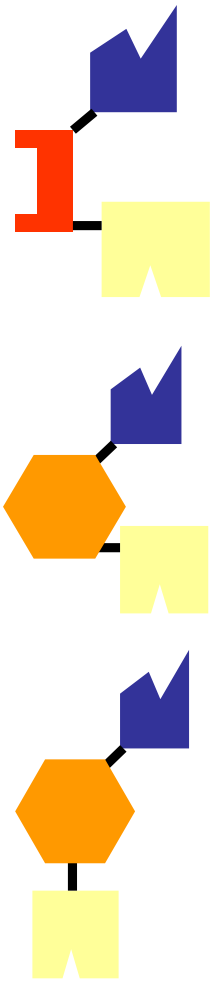
ספריית קבוצות כימיות:

1 5

2 6

3 7

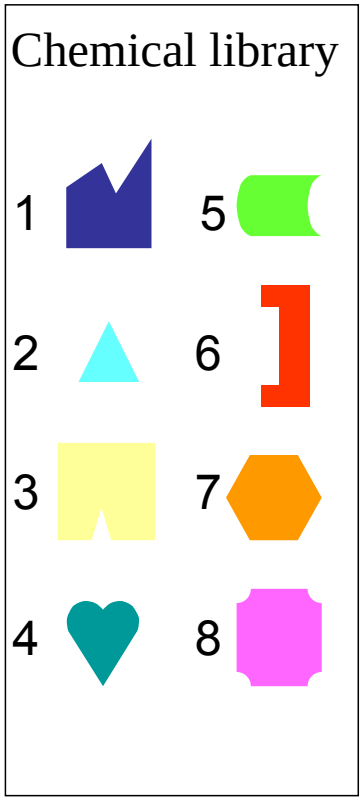
4 8



1 045 6 090 3

1 045 7 090 3

1 045 7 180 3



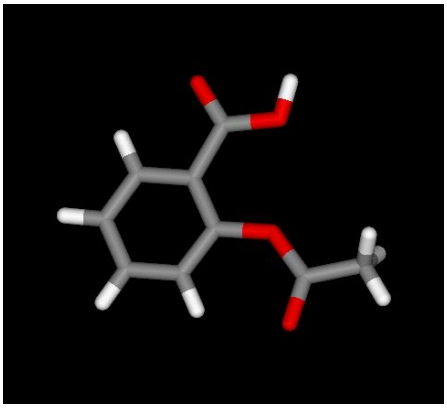
- Databases
- Structural alignment
- Structural classification
- Secondary structure prediction
- Tertiary structure prediction
- Molecular docking
- Visualization
- Dynamics

Visualization – Molecular graphics

What do we need?

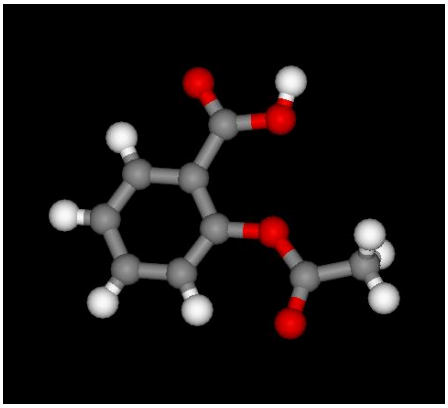
- Rotation & translation
- Color specific parts of the molecule
- Labeling of residues and atoms
- Geometrical measurements (distances & angles)
- Schematic representation:
  Atoms/Bonds/Secondary structures, …
- Molecular surfaces
- Compare structures
- Saving pictures

Representation of molecules (1)



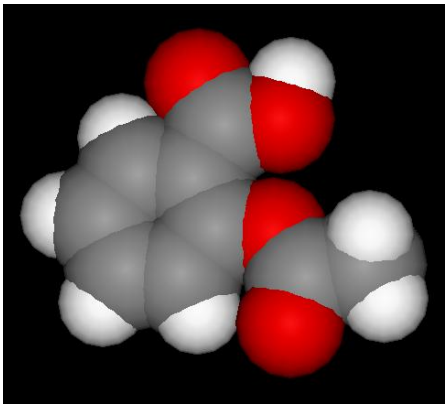
Stick-model

Ball size: 0
Stick size: 0.2



Ball & Stick

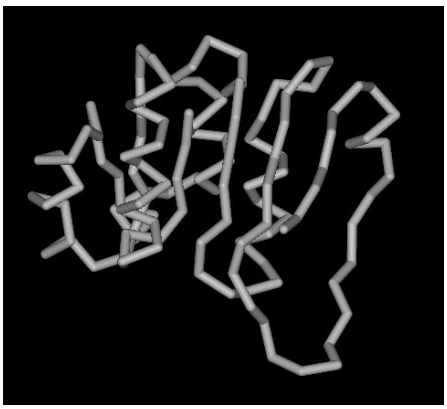
Ball size: 0.4
Stick size: 0.2



Space-filled model

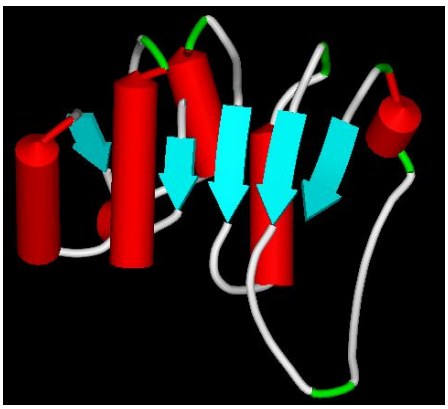
Ball size: 0.8
Stick size: 0

Representation of molecules (2)



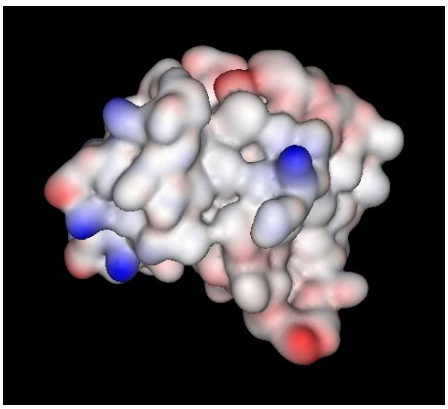
Backbone

only connections
between C-alpha atoms



Schematic

helix – cylinder
strand – arrow



Surface

color indicate
electrostatic
potentials

Molecular Graphics Software Links

The PDB does not distribute software for molecular graphics but we maintain a list of widely used molecular graphics programs. These programs take PDB coordinate files as input.

- **BioEditor**
  A tool for creating and viewing dynamic, formatted structure annotations; for Windows
- **BRAGI**
  A protein visualization and modeling program
- **Chemscape Chime**
  From MDL Information Systems. This program allows visualisation of structures with Windows, Macintosh (with Netscape 4.x only), and (as version 0.9z) SGI/Irix. For further information about Chime see the **UMass Chime Resources Page**
- **Chimera**
  Interactive molecular modeling system, free to academic/non-profit; displays multiple sequence alignments and associated structures, atom-type and H-bond identification, molecular dynamics trajectories (AMBER format), and offers ligand-screening interface (DOCK), filter by number/position of H-bonds, and extensibility to create custom modules - for Windows, Linux, Mac OS X, IRIX, and Tru64 Unix
- **Cn3D**
  Simultaneously displays structure, sequence, and alignment, with annotation and alignment editing features, for use with 3-D structures from NCBI's Entrez; available for Windows, Macintosh, and Unix
- **CrystalMaker**
  A program for building, displaying and manipulating all kinds of crystal and molecular structures.
- **iMol**
  Open GL graphics program displays small, large, and multiple molecules; measures distances and angles,

- **ProteinScope**
  Free viewer to display and manipulate PDB files and create animations and slides of proteins for Windows
- **PyMOL**
  A free and open-source molecular graphics system for visualization, animation, editing, and publication-quality imagery. PyMOL is scriptable and can be extended using the Python language. Supports Windows, Mac OSX, Unix, and Linux
- **QuteMol**
  An open source (GPL), interactive, high quality molecular visualization system. QuteMol exploits the current GPU capabilites through OpenGL shaders to offers an array of innovative visual effects.
- **RasMol**
  A free viewing system for PDB coordinate files that runs on Mac (PPC), Windows, Unix, and Linux systems. **Open source versions** are also available.
- **Raster3D**
  A set of tools for generating high quality raster images of proteins or other molecules. Freeware for Mac OSX, Windows, Unix, and Linux
- **RasTop (v. 2.0)**
  A free user-friendly graphical interface to RasMol molecular visualization software (v. 2.7.2.1), available for Windows and Linux
- **Ribbons**
  A program for molecular illustration and error analysis,

http://www.rcsb.org/pdb/static.do?p=software/software_links/molecular_graphics.html

displays dynamics trajectories. For Mac OS X incl. 10.2

- **Java3D Molecular Visualisation System**
  Free Java/Java3D progam and source code for Windows or Unix
- **Jmol**
  Jmol is a free, open source molecule viewer for students, educators, and researchers in chemistry and biochemistry. It is cross-platform, running on Windows, Mac OS X, and Linux/Unix systems.
- **Mage and Kinemages**
  Interactive molecular display for research and educational uses. Free, open source for Windows and Mac (OSX or PPC), Unix, and Linux. A Java version does 3-D Web display without plug-ins.
- **MOLMOL**
  A program for displaying, analyzing, and manipulating the 3-D structure of biological macromolecules, with special emphasis on the study of protein or DNA structures determined by NMR; for Mac OSX, Windows, Unix, Linux
- **MolScript**
  A program for displaying structures in both detailed and schematic formats and writing images in various formats for Unix
- **MolView and MolView Lite**
  Free molecular visualization programs for the Mac (PPC)
- **MVM**
  Molecular Visualization Program and GUI of ZMM. MVM is a free molecular viewer that can be used to display protein, nucleic acids, oligosacharides, small and macromolecules. It has an intuitive interface. In addition to being a molecular viewer, it is the user interface of a very powerful molecular mechanics engine (ZMM).
- **PDB2MGIF**
  Free, user-friendly server that converts PDB files to animated gif files that can be used in Web pages and

A Tcl/Tk script responsible to redirect PDB files or RasMol scripts to multiple RasMol sessions; can be used as a Web browser helper application or as a standalone program for Mac (OSX or PPC), Windows, or Unix

- **Sirius**
  An extensible molecular graphics and analysis environment developed at San Diego Supercomputer Center
- **SPADE**
  The Structural Proteomics Application Development Environment (SPADE) provides community tools for development and deployment of essential structure and sequence equipment. Includes a chemical probing suite to support experimental verification of predicted structural models. Written in Python with scripting tools available. Runs on Windows, Linux and Mac.
- **STRAP**
  Align proteins by sequence and 3D structure. STRAP was developed by Christoph Gille at the Institut für Biochemie, Charité
- **Swiss PDB viewer** available from **Australia** | **Bolivia** | **Canada** | **China** | **Korea** | **Switzerland** | **Taiwan** | **USA** |
  A 3D graphics and molecular modeling program for the simultaneous analysis of multiple models and for model-building into electron density maps. The software is available for Mac (OSX or PPC), Windows, Linux, or SGI
- **Uppsala Electron Density Server**
  Generated density maps
- **VMD**
  VMD (Visual Molecular Dynamics) runs on many platforms including MacOS X, and several versions of Unix and Windows. VMD provides visualization, analysis, and Tcl/Python scripting features, and has recently added sequence browsing and volumetric rendering features. VMD is distributed free of charge.

Dynamics of proteins

• Dynamics of proteins is clearly related to their function.

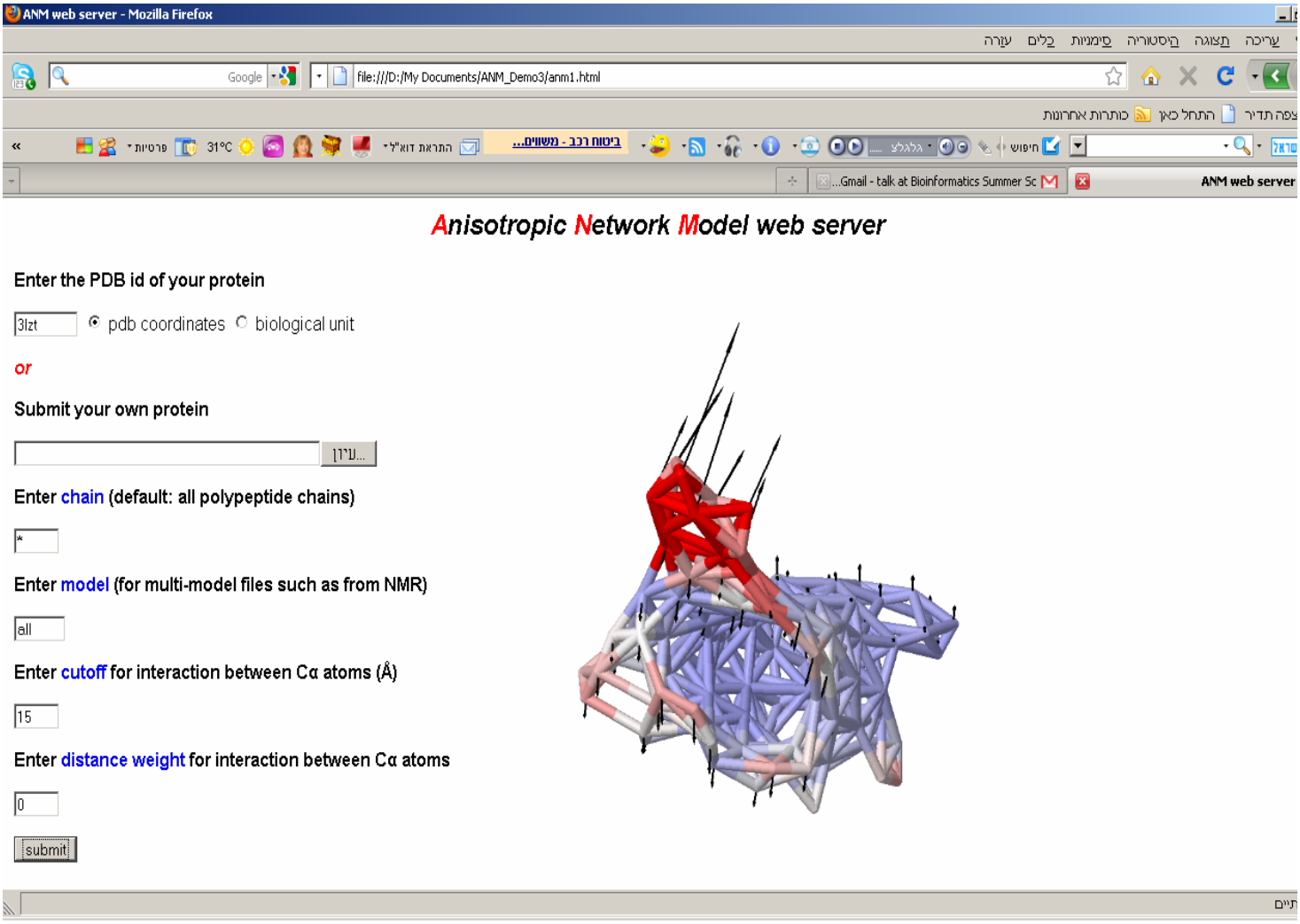
• Understanding the relation between the two is a main challenge in the field of biophysics

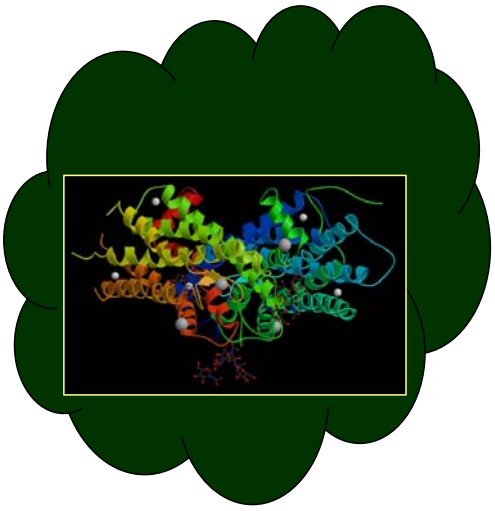
• ***Molecular Dynamics*** provides a way to conduct non-equilibrium simulations but only for short time scales (10^{-7} s)

• ***Normal Mode Analysis*** provides a way to analyze equilibrium motion for longer time scales

Type of motion	Functionality examples	Times and Amplitude scales
Local Motions: •Atomic fluctuation •Side chain motion	•Ligand docking flexibility •Temporal diffusion pathways	fs - ps (10^{-15} - 10^{-12} s) less than 1 Å
Medium Scale Motions: •Loop motion •Terminal-arm motion •Rigid-body motion (helices)	•Active site conformation adaptation •Binding specificity	ns - μs (10^{-9} - 10^{-6} s) 1 - 5 Å
Large Scale Motions: •Domain motion •Subunit motion	•Hinge bending motion •Allosteric transitions	μs - ms (10^{-6} - 10^{-3} s) 5 - 10 Å
Global Motions: •Heix-coil transition •Folding/unfolding •Subunit association	•Hormone activation •Protein functionality	ms - h (10^{-3} - 10^{4} s) more than 10 Å

Modified after: Becker & Watanabe (2001). Dynamic Methods.
In Computational & Biochemistry & Biophysics (Edited by Becker *et al.*)





Thanks

- The organizers
- Dr. Jaume Bacardit

You!