

Comparative Modelling

Summary

1. Basic concepts of Homology Modeling
2. Schema of the method
 1. Fold assignment
 2. Template selection
 3. Model building
 4. Evaluation
 5. Improvement

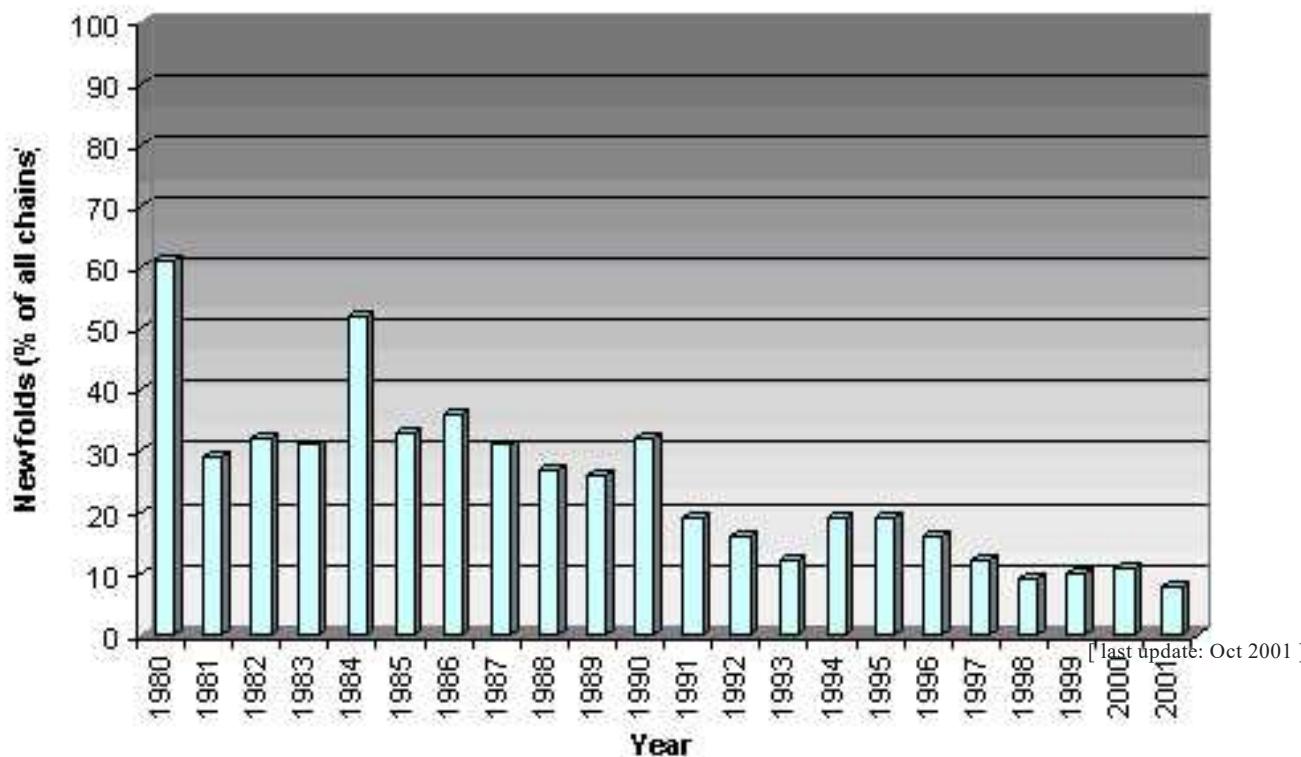
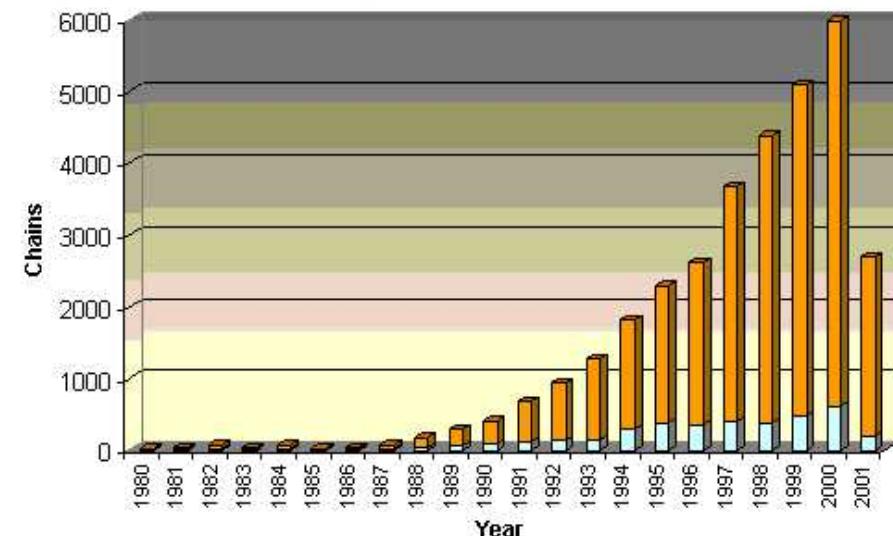
1. Basic concepts of Homology Modeling

Definition

Extrapolation of the structure for a new (target) sequence from the known 3D-structures of related family members (templates).

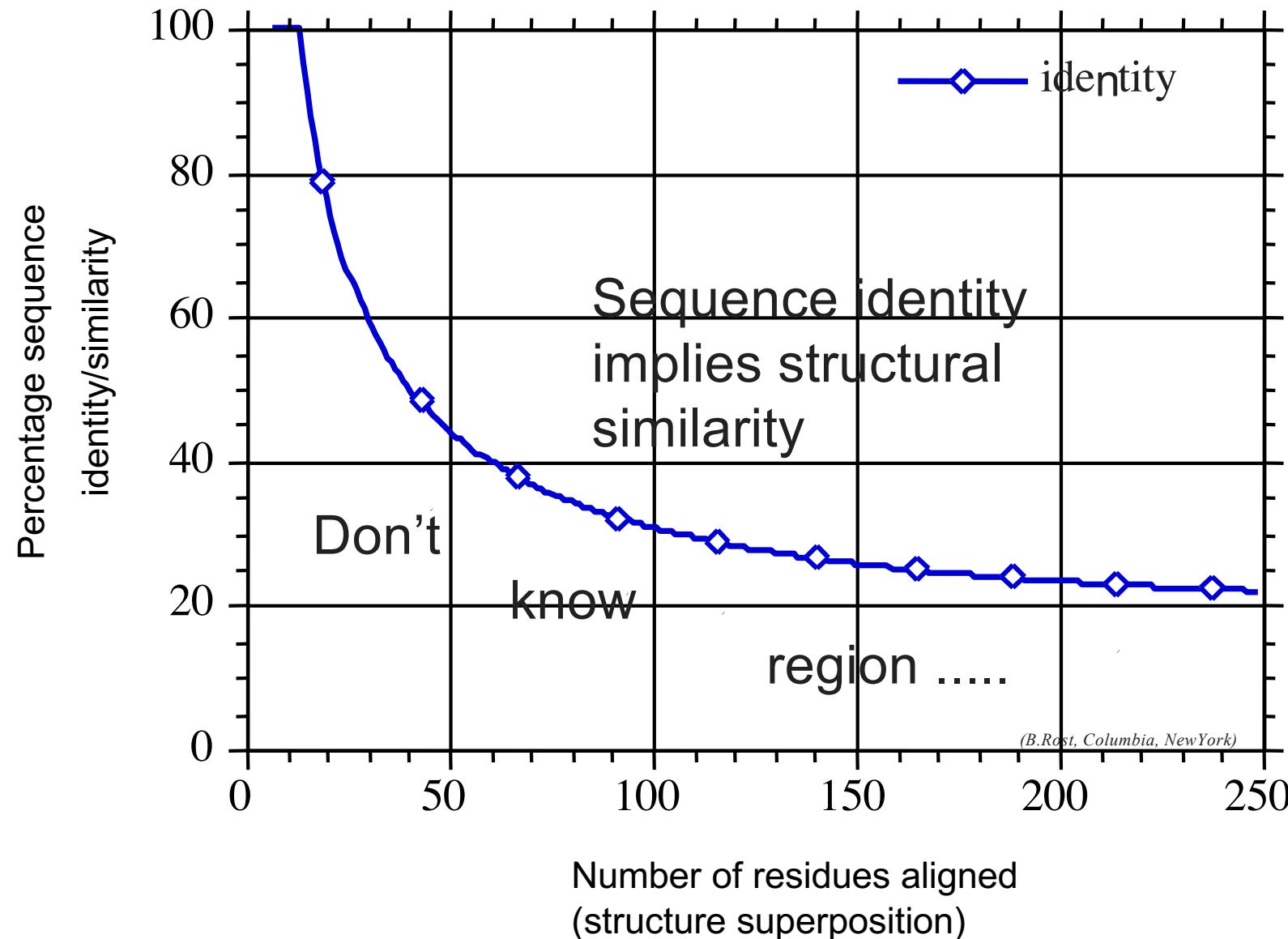
1. Basic concepts of Homology Modeling

The number of different protein folds is limited:



1. Basic concepts of Homology Modeling

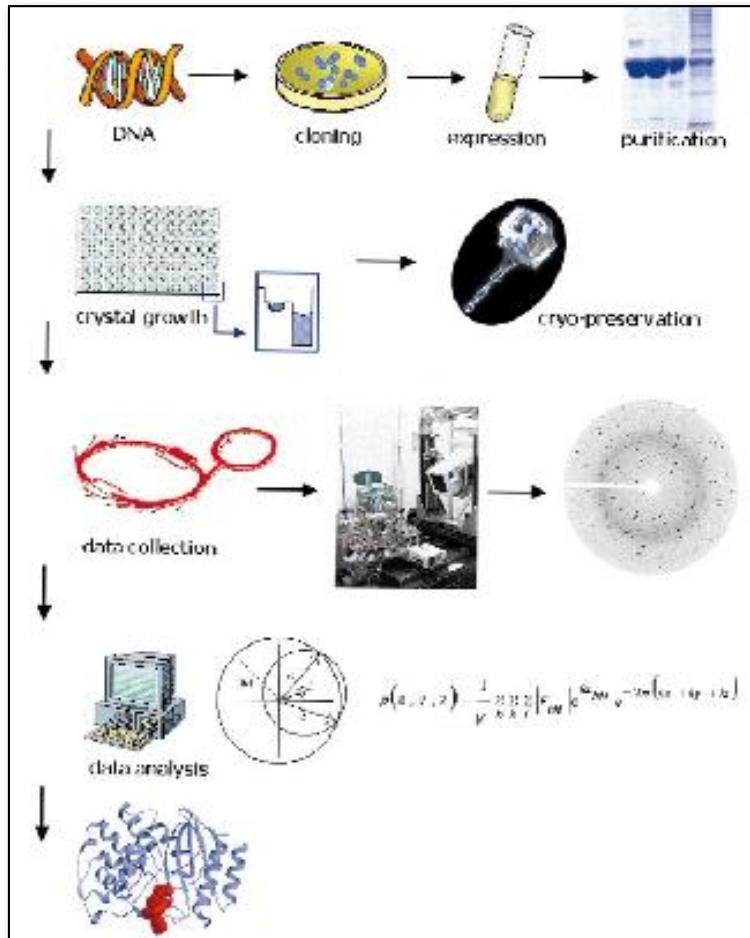
Sequence similarity implies structural similarity?



1. Basic concepts of Homology Modeling

- Fold is more conserved than sequence.
- Secondary structure are the most conserved parts
- Loops have the higher variability in structure.

1. Basic concepts of Homology Modeling Structural Genomics



express & purify

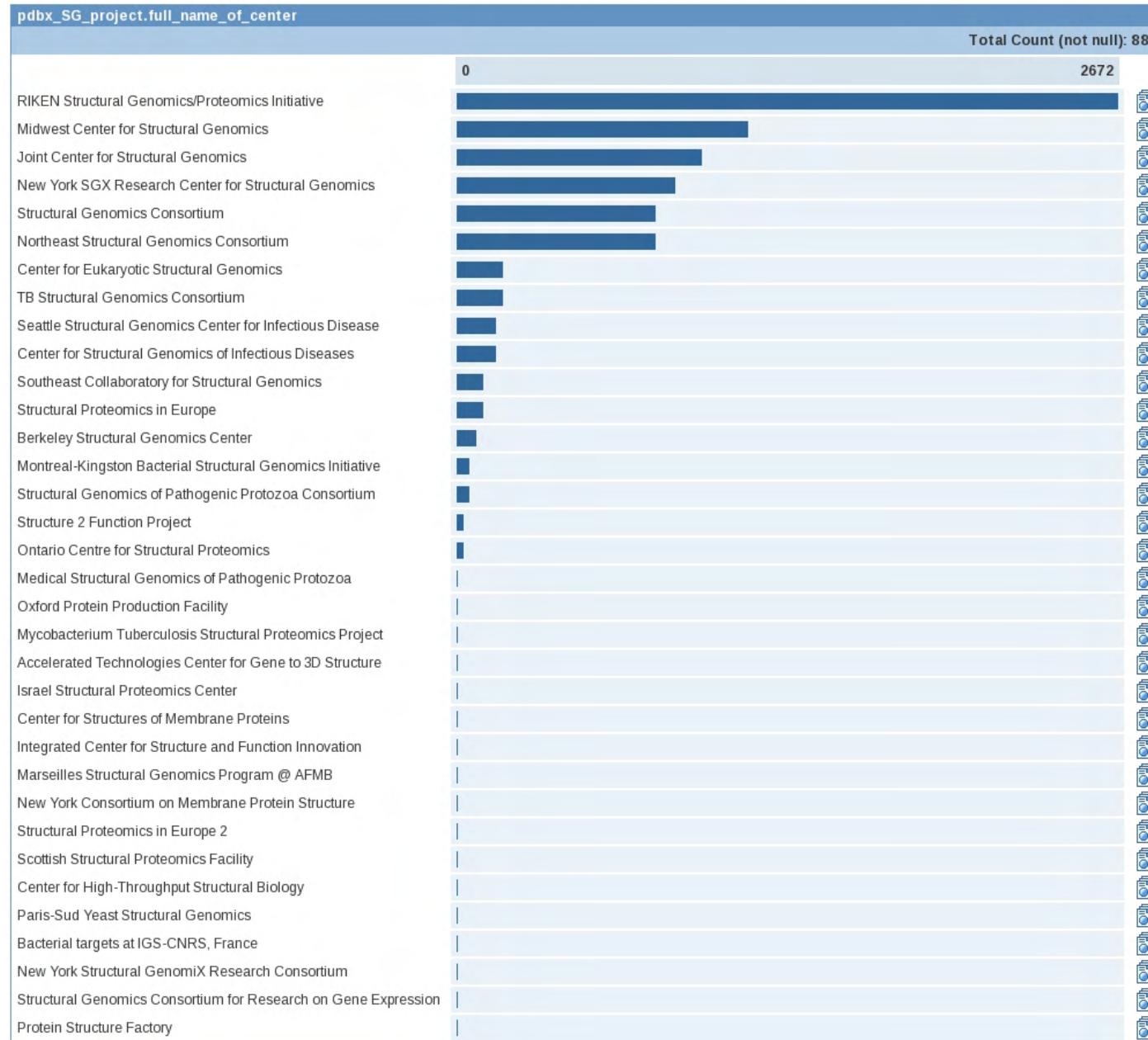
cristallize

X-ray

analyses

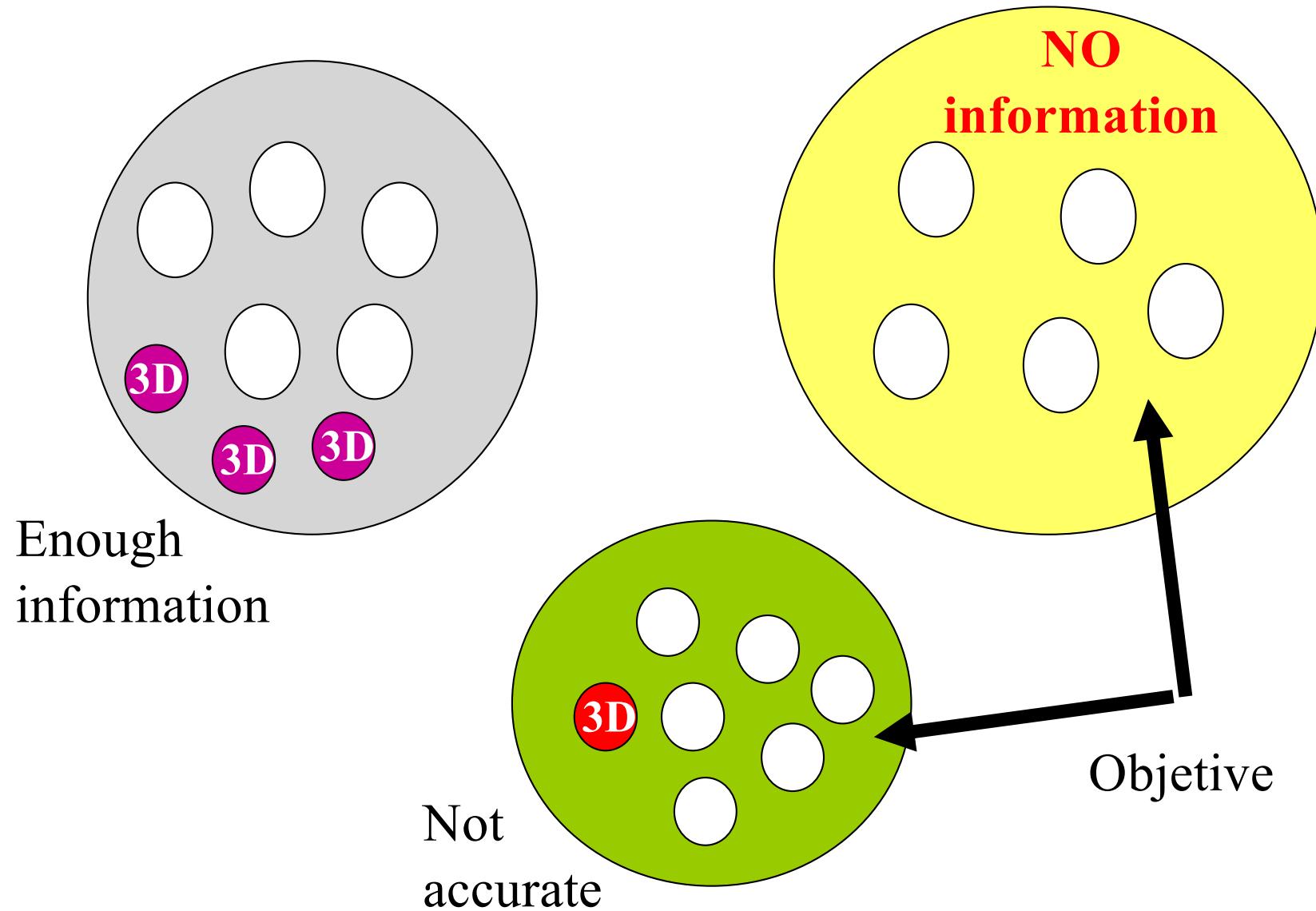
structure

1. Basic concepts of Homology Modeling Structural Genomics



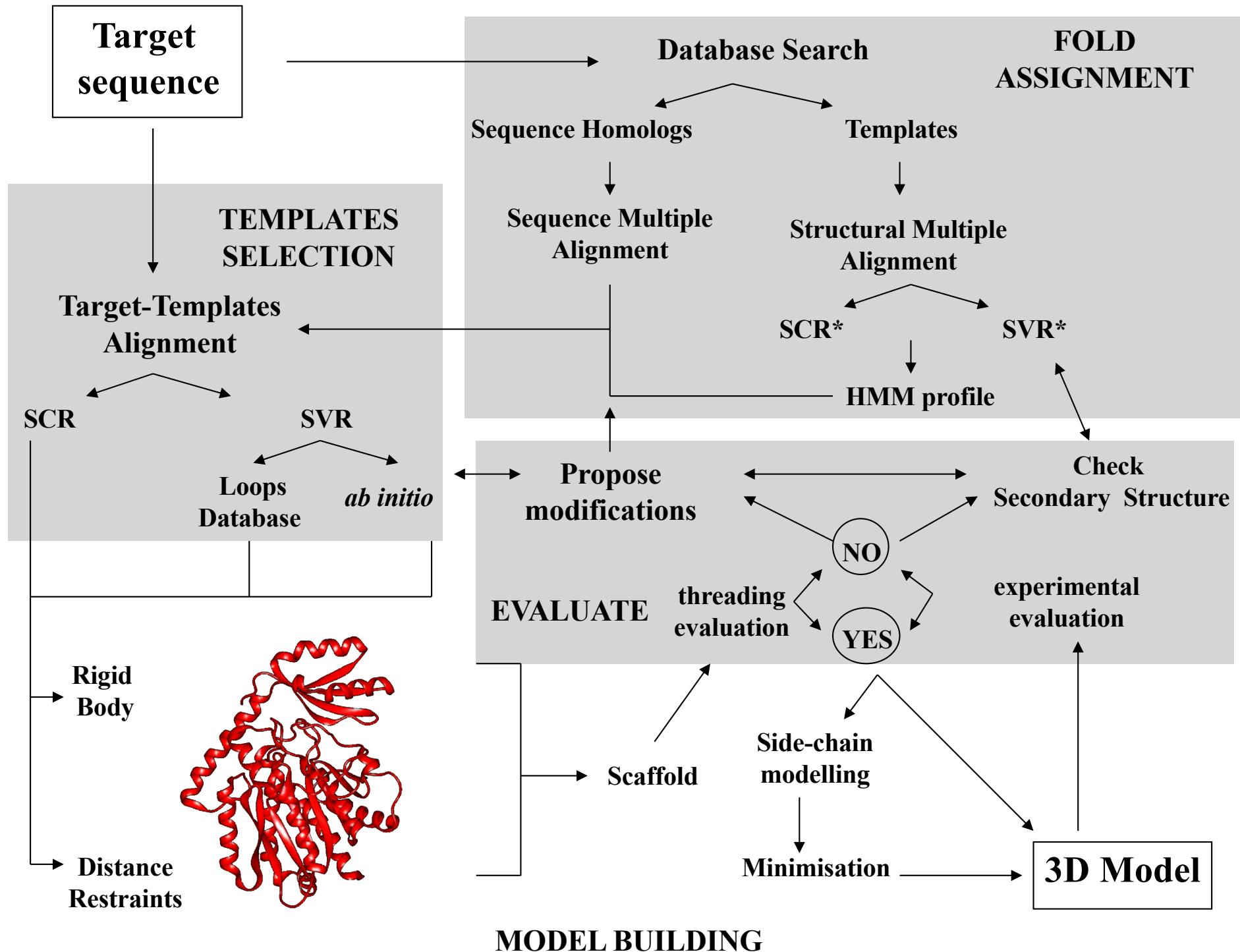
1. Basic concepts of Homology Modeling

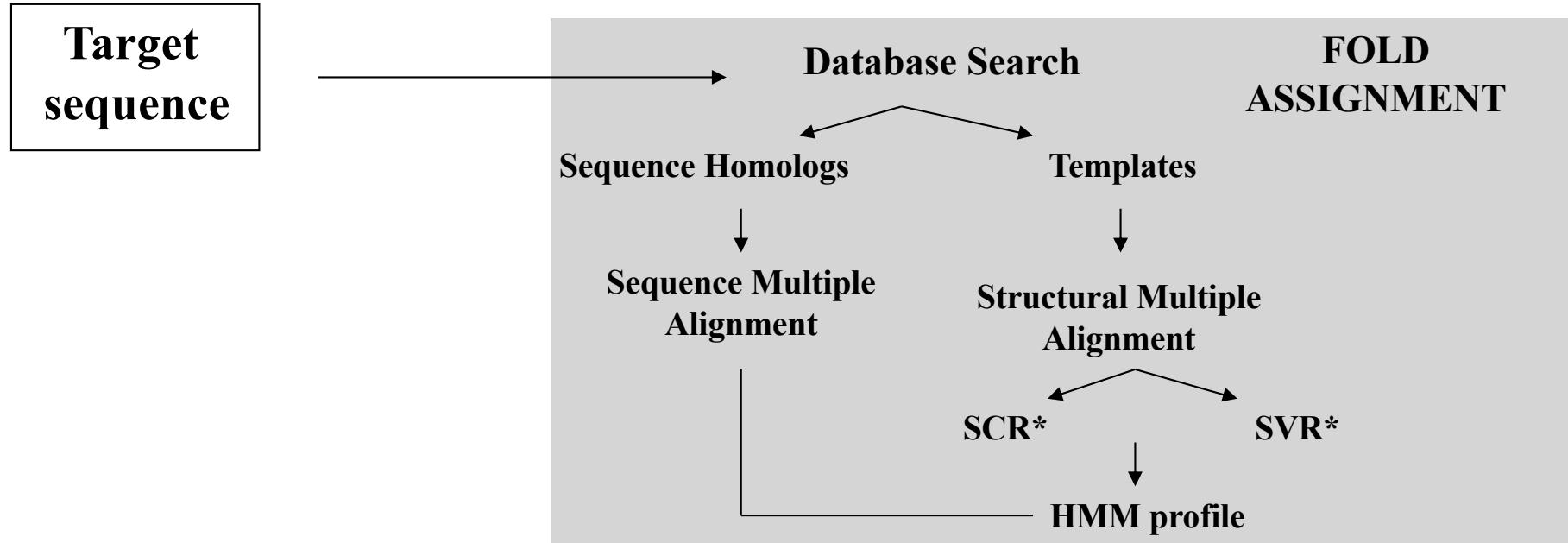
Structural Genomics



2. Schema of the method

- 1.Fold assignment**
- 2.Template selection**
- 3.Model building**
- 4.Evaluation**
- 5.Improvement**



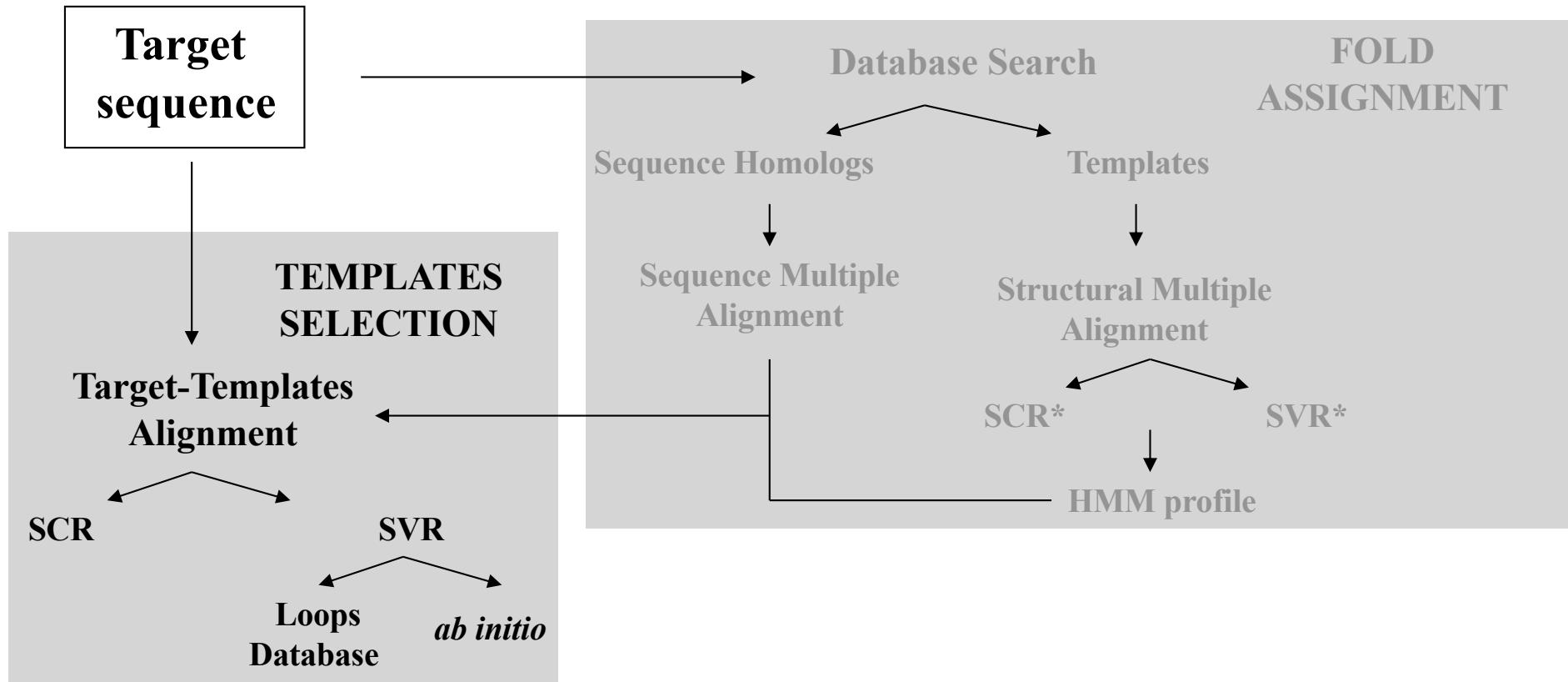


2. Schema of the method

1. Fold assignment

Sequence search with the target

1. Compares the sequence of the target with a set of sequences with known structure
2. Ranking the comparisons by scores.
3. Scores are related to P-values or E-values (high score implies low P-value). P-value is the probability of obtaining the same alignment by chance.
4. Scores are calculated using a residue-substitution matrix:
 1. PAM: based on the alignment of sequences of homologs
 2. BLOSUM: based on the alignment of blocs of similar sequences
5. One sequence can have more than one domain, therefore we can obtain the best scores for partial parts of the target.
6. Methods (see practice)
 1. BLAST algorithm, matches words from a pre-calculated and indexed set and joints them into sentences (forming the sequence)
 2. FastA: Smith & Waterman algorithm
 3. Scanning PFAM: algorithm of Hidden Markov Models



2. Schema of the method

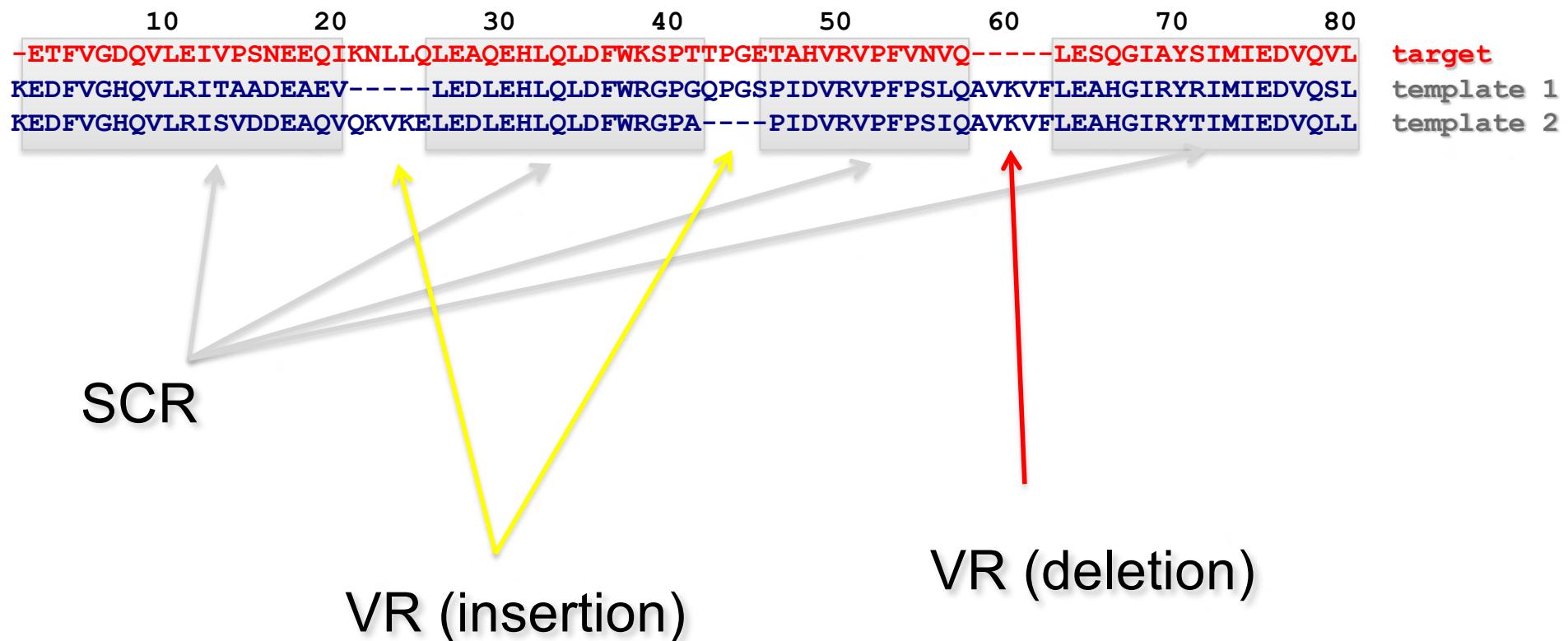
2. Template selection

Selecting the best target-alignment template

1. The template(s) should be the closest homolog(s) to the target
2. Small number of templates to avoid stress on model building
3. Multi-domain proteins require the use of at least one template with the largest coverage of sequence (containing the largest number of domains)
4. Structural alignment of homologs gives the information on position-specific substitutions
5. Detection of structurally conserved regions (SCR) and variable regions (VR)
6. Aligning the target sequence and template sequences using a multiple sequence profile helps to avoid misalignments
7. Methods (see practice)
 1. ClustalW
 2. T-coffee
 3. HMMER
 1. alignment with a known family profile (PFAM)
 2. Alignment with a profile built with the structure of homologs

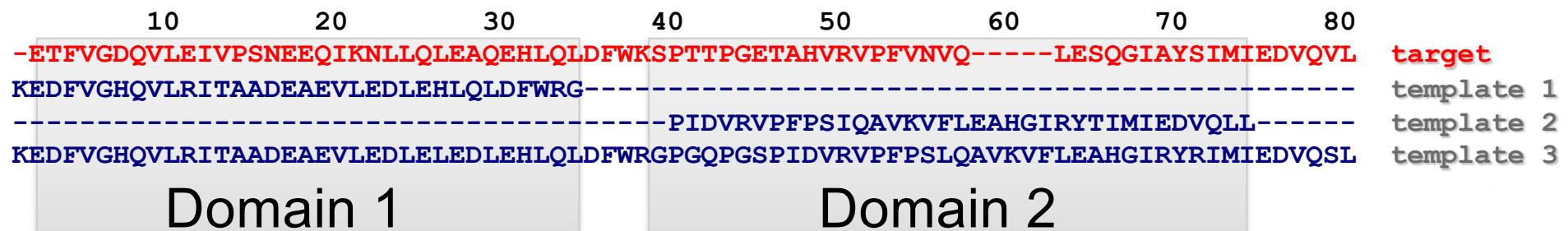
2. Schema of the method

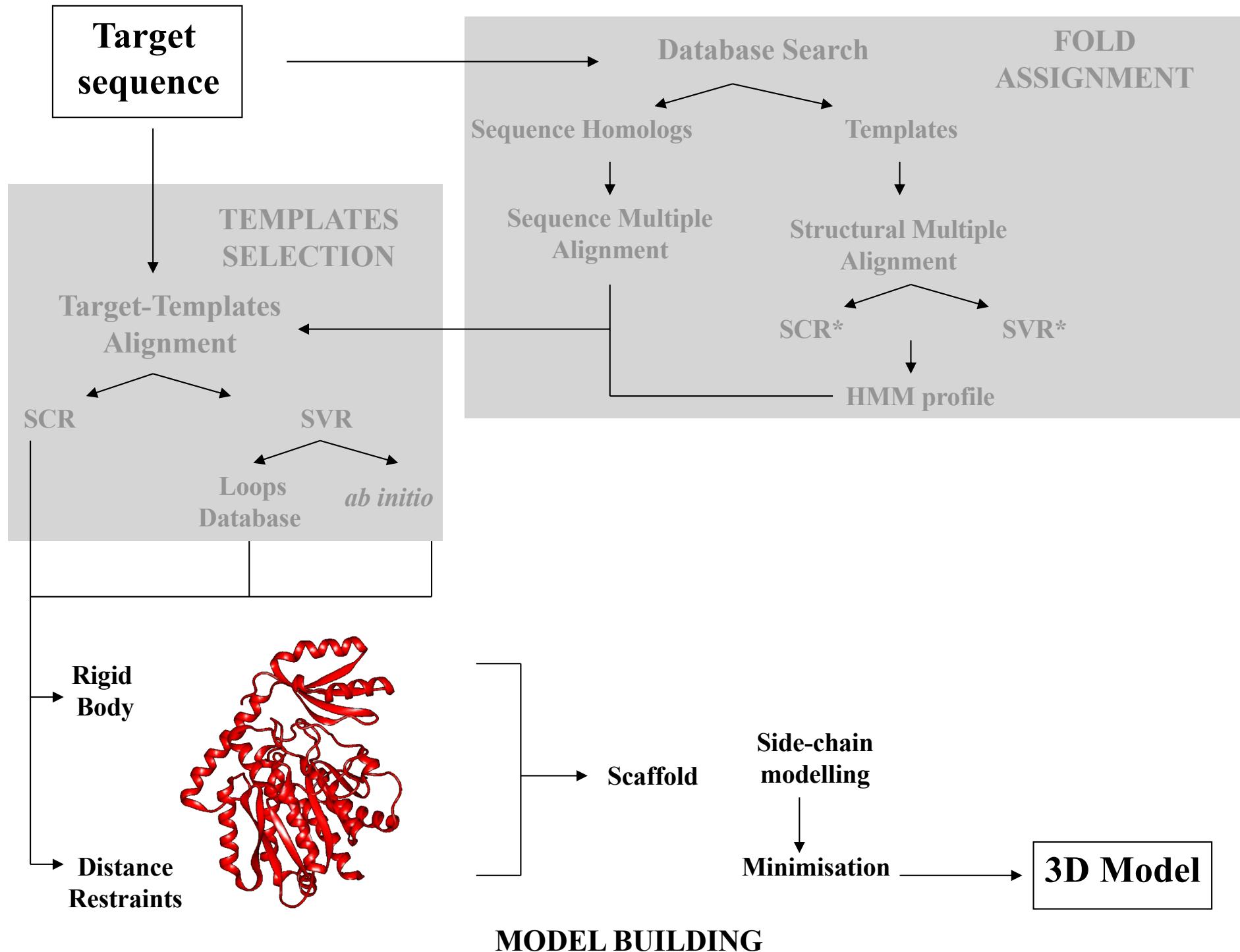
2. Template selection



2. Schema of the method

2. Template selection





2. Schema of the method

3. Model building

1. Rigid Body Assembly

- 1. Core framework (SCR)**
- 2. Loop modeling (VR)**
- 3. Energy minimization**

2. Spatial restraints

- 1. Probability Density Functions (PDF)**
- 2. Distance restraints**
- 3. Simulated Annealing**
- 4. Loop modeling**

3. Side-chain modeling

- 1. Back-bone dependent rotamer libraries**
- 2. Energetic and packing criteria**

2. Schema of the method

3. Model building: Rigid Body Assembling (core framework)



- Averaging core template backbone atoms
(weighted by local sequence similarity with the target sequence)
- Leave non-conserved regions (loops) for later

2. Schema of the method

3. Model building: Rigid Body Assembling (loop modeling)

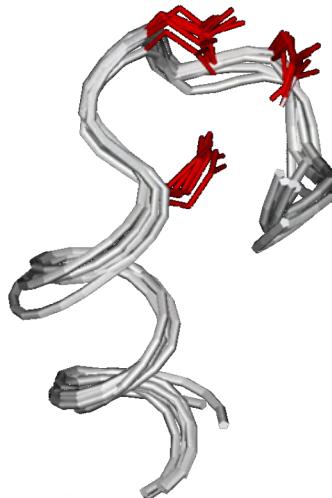
1. Use the “spare part” algorithm to find compatible fragments in a Loop-Database
2. “*ab-initio*” rebuilding of loops (Monte Carlo, molecular dynamics, genetic algorithms, etc.)



2. Schema of the method

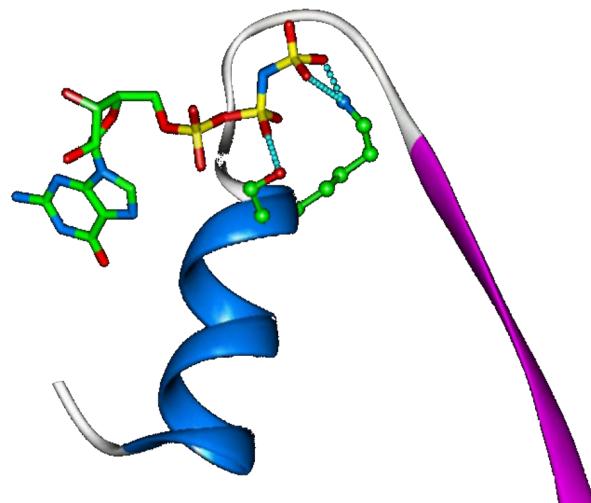
3. Model building: Rigid Body Assembling (loop modeling)

1. Use the “spare part” algorithm to find compatible fragments in a Loop-Database



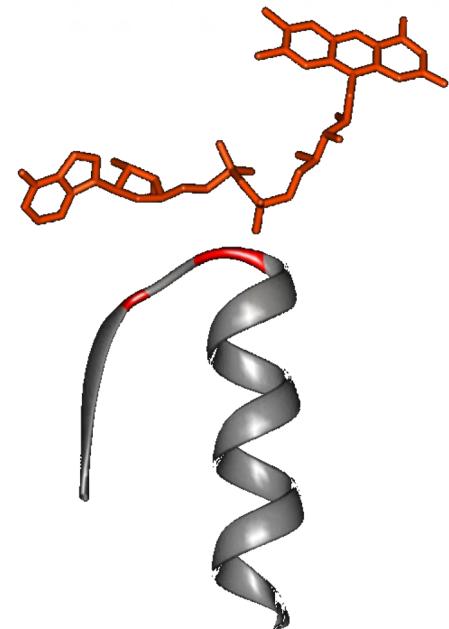
**EF-Hand
Calcium binding**

aa{baalal}bb
Xh{DXDpDG}Xh



P-loop GTP binding

bb{eppgag}aa
hh{GhXXpG}Kp



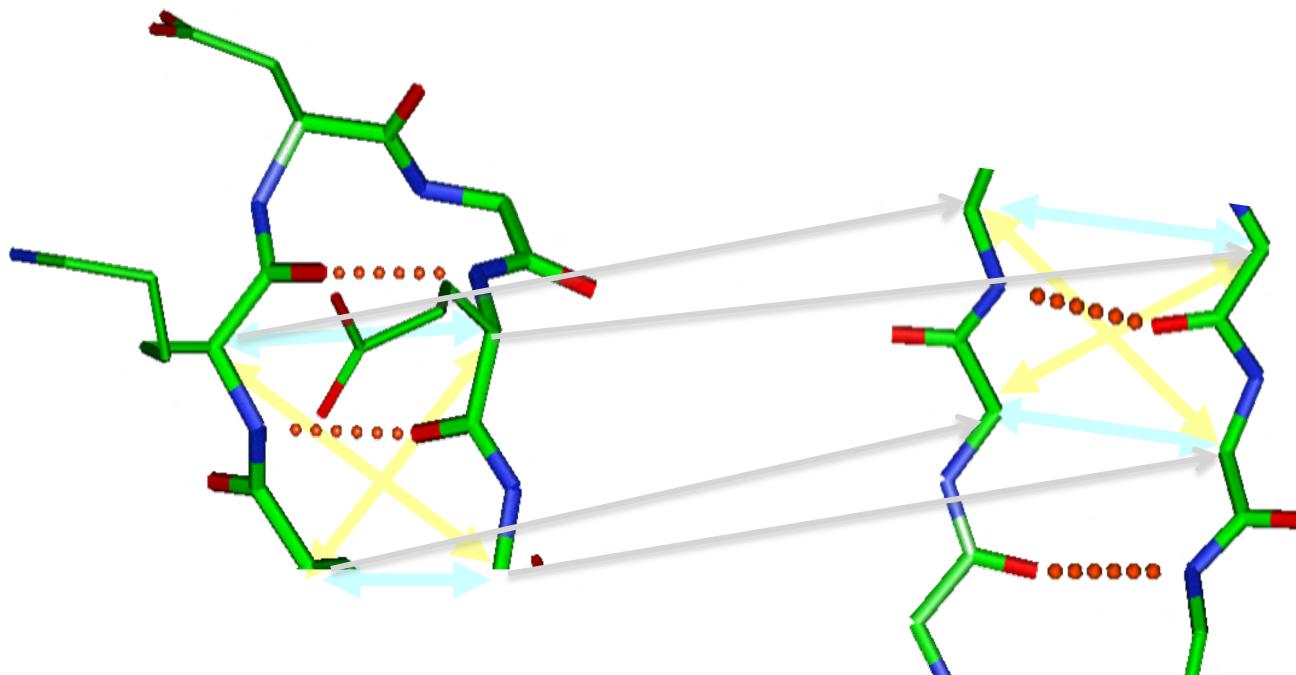
**NAD(P)/FAD
binding**

bb{eab}aa
hh{GhG}hX

2. Schema of the method

3. Model building: Rigid Body Assembling (loop modeling)

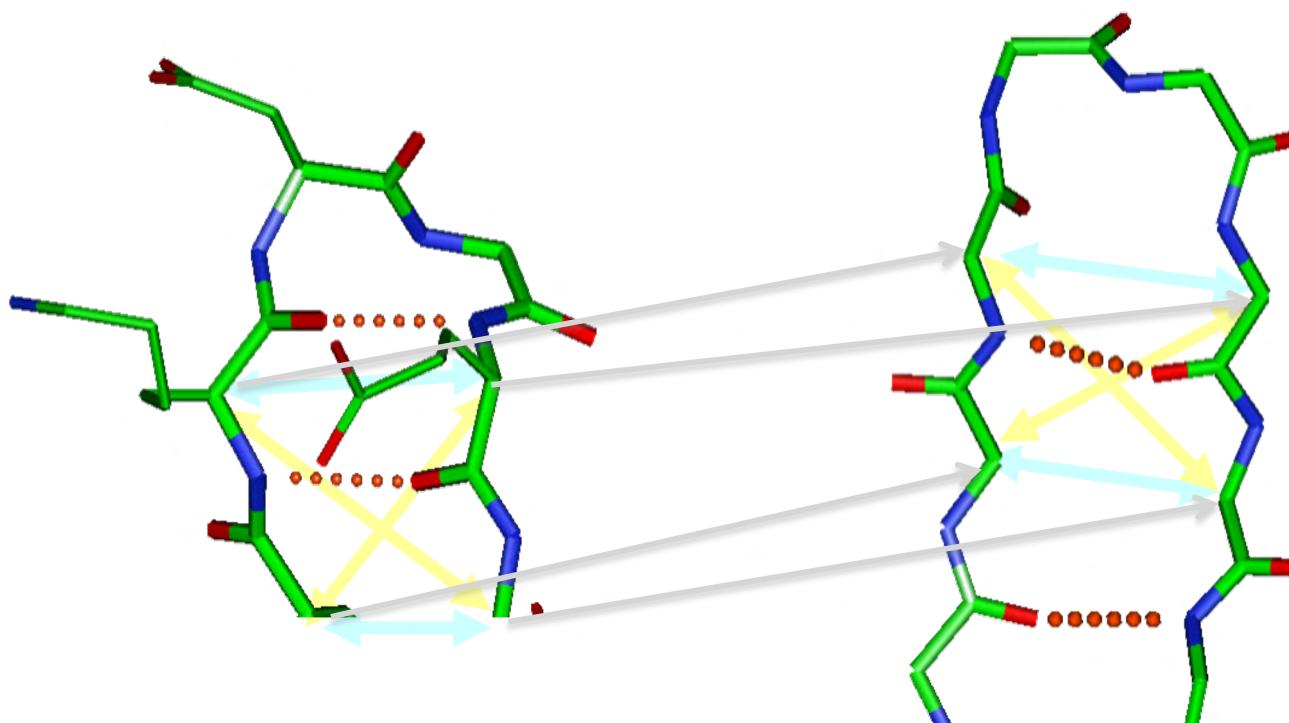
1. Use the “spare part” algorithm to find compatible fragments in a Loop-Database



2. Schema of the method

3. Model building: Rigid Body Assembling (loop modeling)

1. Use the “spare part” algorithm to find compatible fragments in a Loop-Database



2. Schema of the method

3. Model building: Rigid Body Assembling (Energy minimization)

$$E_{bonding} = \sum_{bonds} \frac{1}{2} k_i (d_i - d_i^0)^2 + \sum_{angles} \frac{1}{2} k_j (\alpha_j - \alpha_j^0)^2 + \sum_{improper dihedral} \frac{1}{2} k_n (\omega_n - \omega_n^0)^2 + \sum_{angles} E_m \cos(\omega_m \phi_m + \varphi_m)^2$$

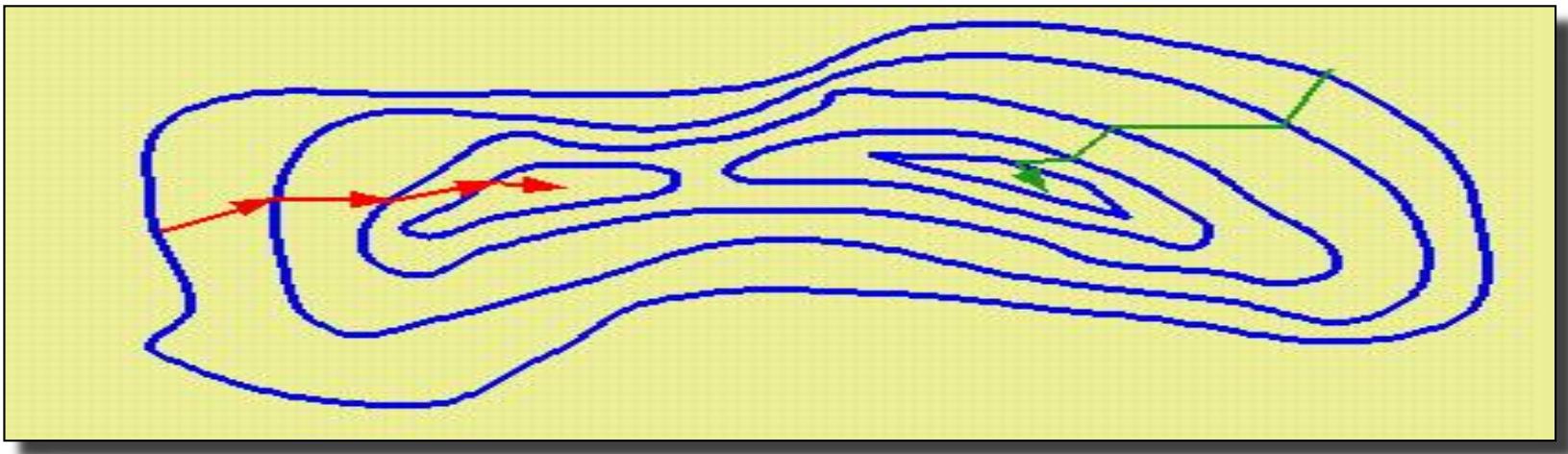
$$E_{non-bonding} = \frac{1}{4\pi\epsilon_0} \sum_i \sum_{j>i} \frac{q_i q_j}{r_{ij}} + \sum_i \sum_{j>i} \frac{C_6^{ij}}{r_{ij}^6} - \frac{C_{12}^{ij}}{r_{ij}^{12}}$$

$$E = E_{bonding} + E_{non-bonding}$$

- modeling will produce unfavorable contacts and bonds: idealization of local bond and angle geometry
- extensive energy minimization will move coordinates away: keep it to a minimum
- Methods: Newton Rapson; Steepest Descent; Conjugate Gradient

2. Schema of the method

3. Model building: Rigid Body Assembling (Energy minimization)



$$x_{i+1} = x_i + \lambda \nabla E$$

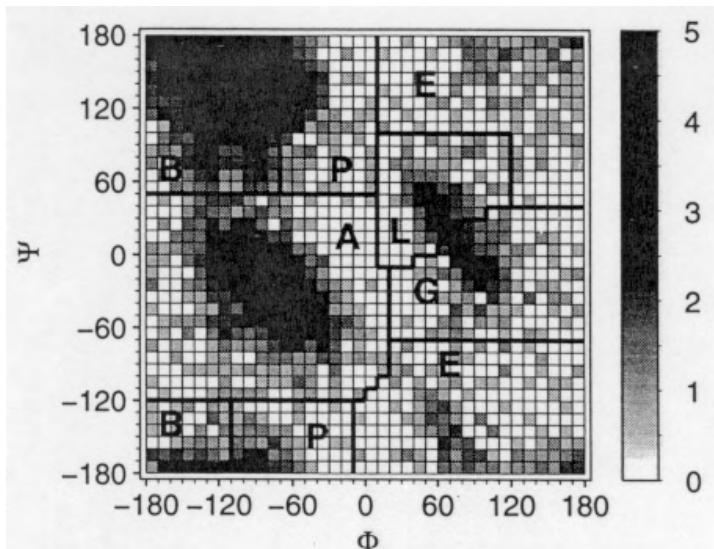
$$\lambda = \begin{cases} E(x_{i+1}) < E(x_i) \Rightarrow \lambda = \lambda + \varepsilon \\ E(x_{i+1}) > E(x_i) \Rightarrow \lambda = \lambda/2 \\ \lambda < \lambda_{\max} \\ E(x_{i+1}) \approx E(x_i) \Rightarrow STOP \end{cases}$$

2. Schema of the method

3. Model building: Spatial restraints (Probability Density Functions)

Feature properties can be associated with

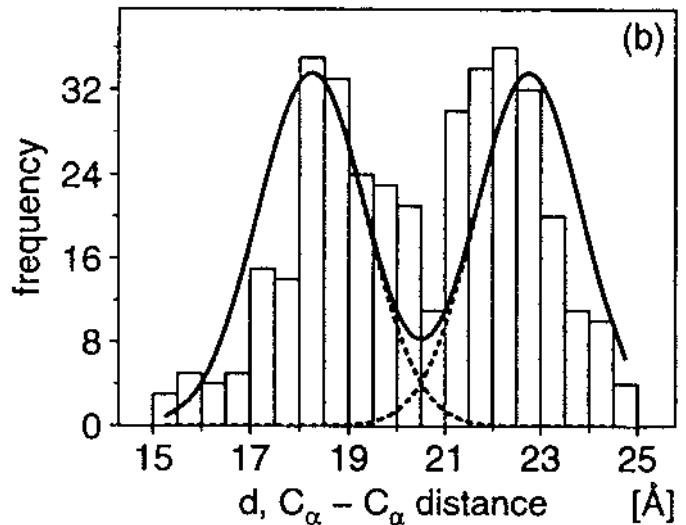
- a protein (e.g. X-ray resolution)
- residues (e.g. solvent accessibility)
- pairs of residues (e.g. C_a - C_a distance)
- other features (e.g. main chain classes)



Example: Ramachandran Plot
Distribution of (ϕ, ψ) angles

2. Schema of the method

3. Model building: Spatial restraints (Probability Density Functions)



Example:
Distribution of C_α-C_α distances

**How can we derive modeling
restraints from this data?**

A restraint is defined as probability density function (*pdf*), p(x):

$$p(x_1 \leq x < x_2) = \int_{x_2}^{x_1} p(x) dx$$

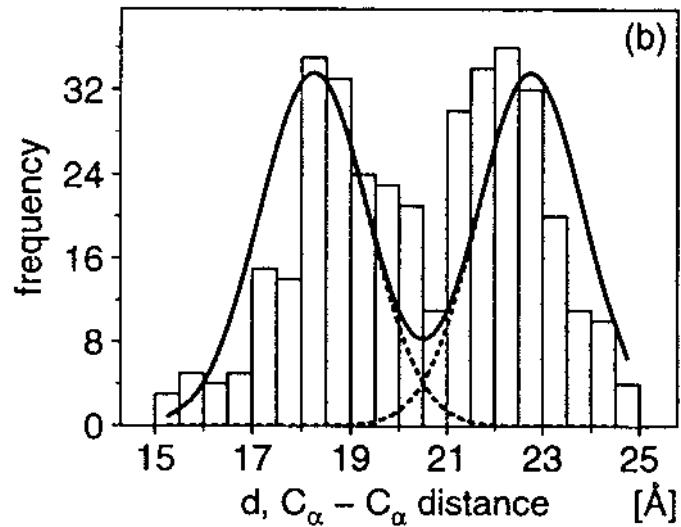
with

$$p(x) > 0$$

$$\int p(x) dx = 1$$

2. Schema of the method

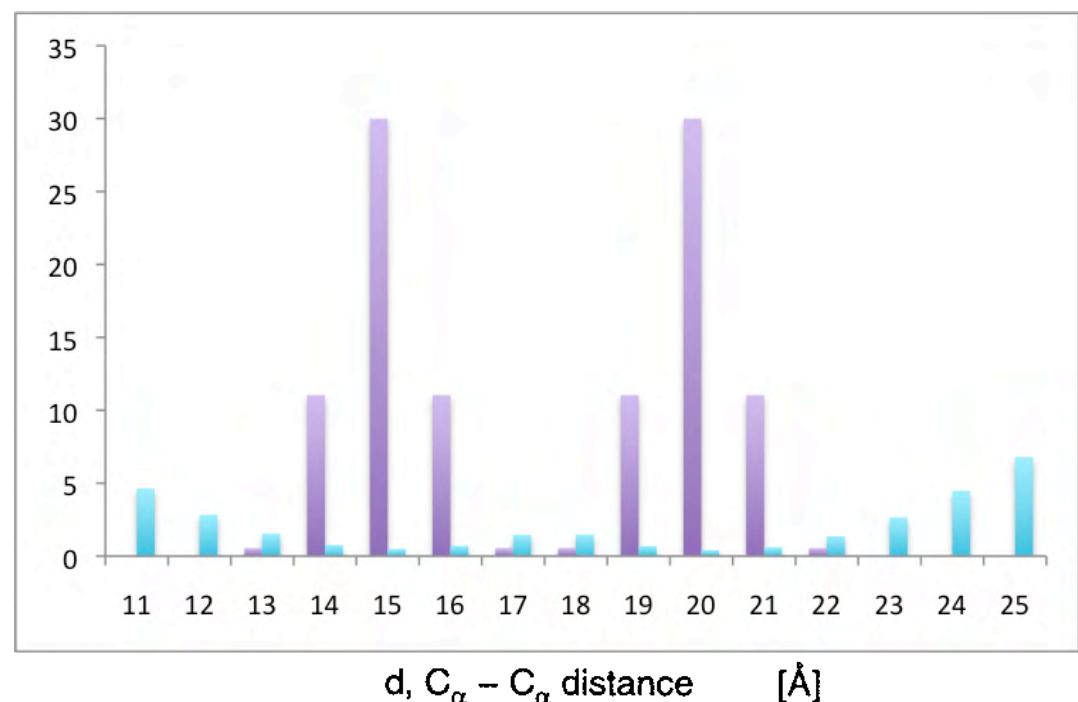
3. Model building: Spatial restraints (Probability Density Functions)



Example:
Distribution of C_α-C_α distances

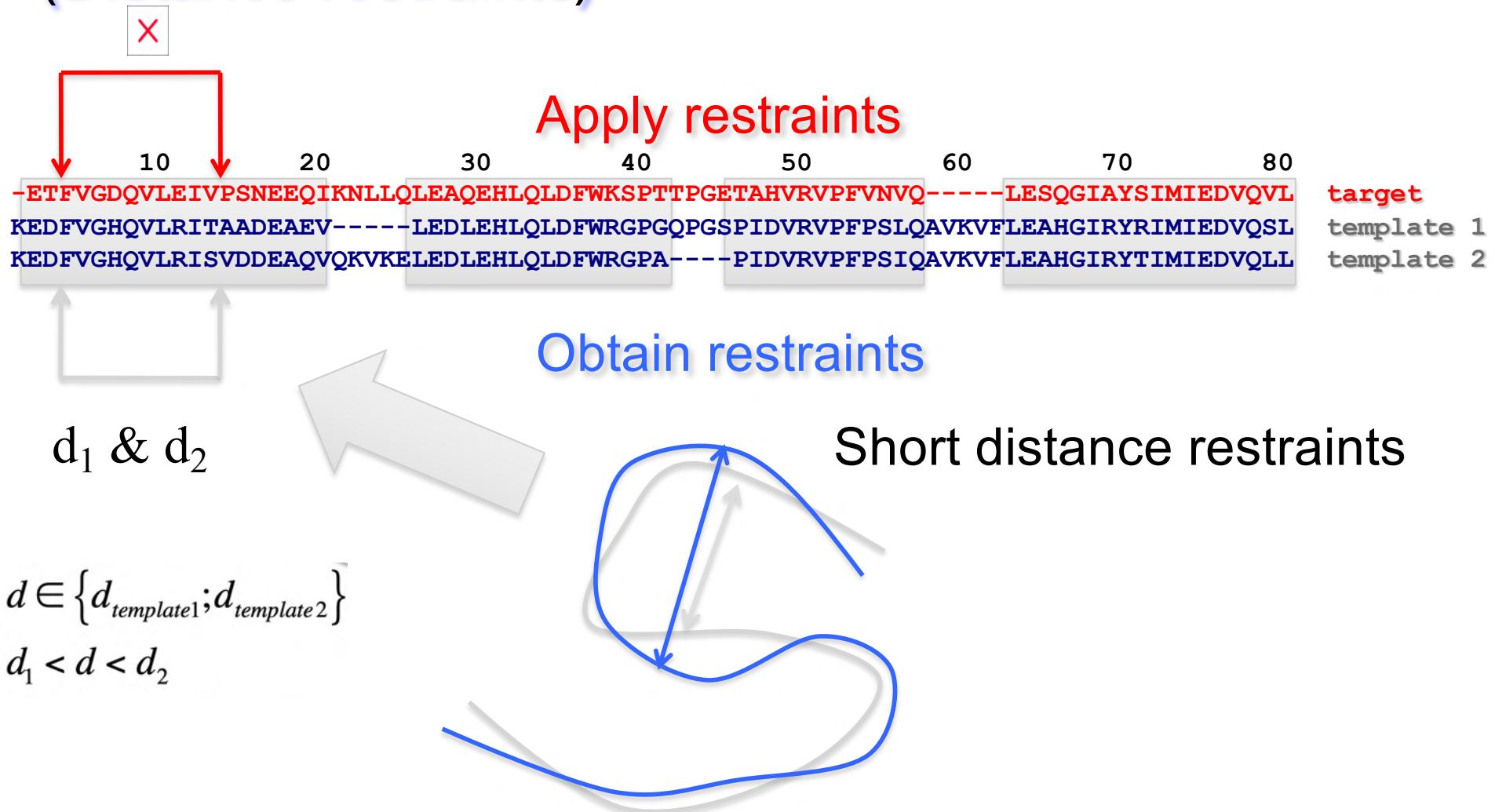
How can we derive modeling
restraints from this data?

$$E_{pdf}(x) = -RT \log(p(x))$$



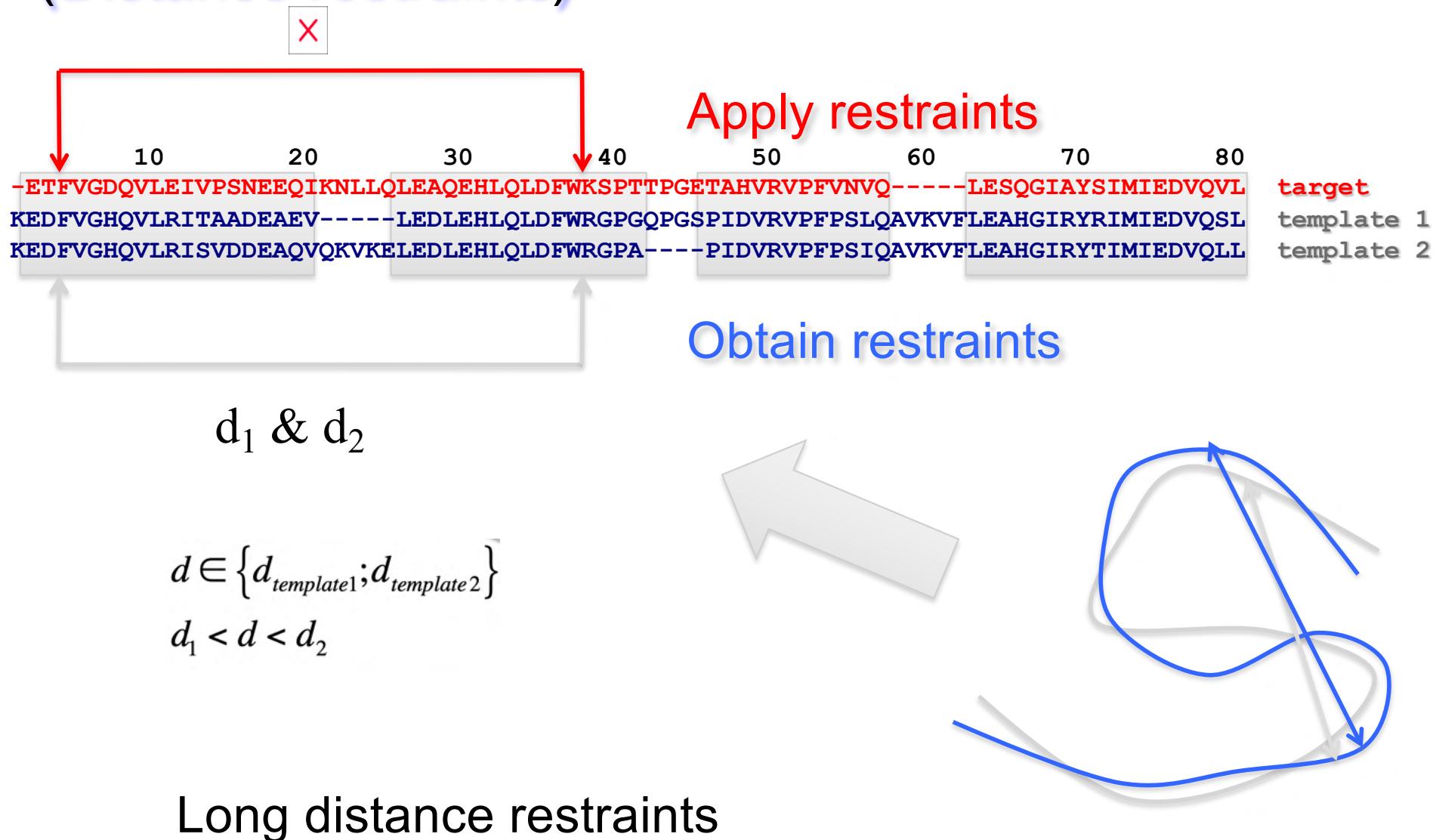
2. Schema of the method

3. Model building: Spatial restraints (Distance restraints)



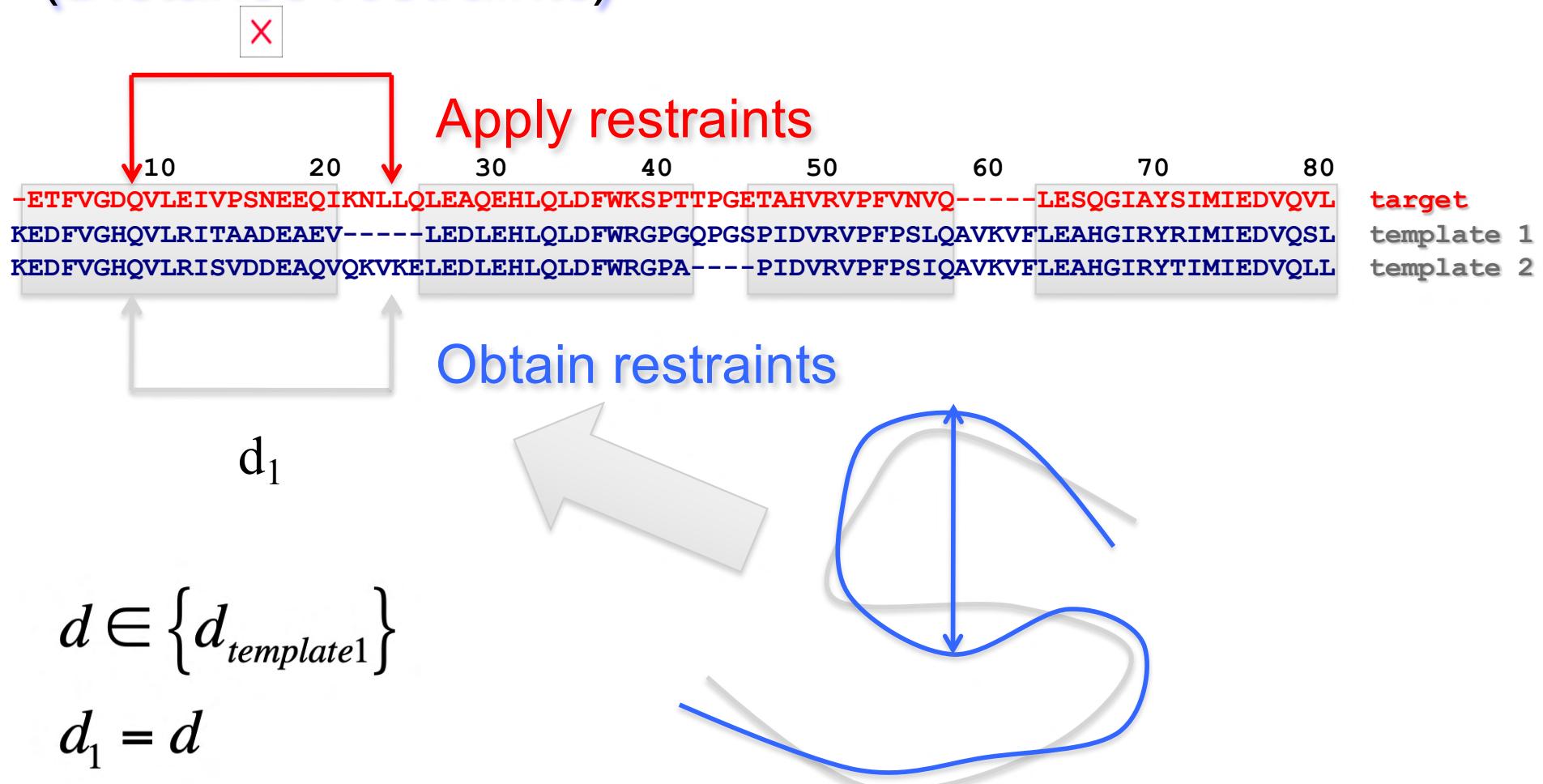
2. Schema of the method

3. Model building: Spatial restraints (Distance restraints)



2. Schema of the method

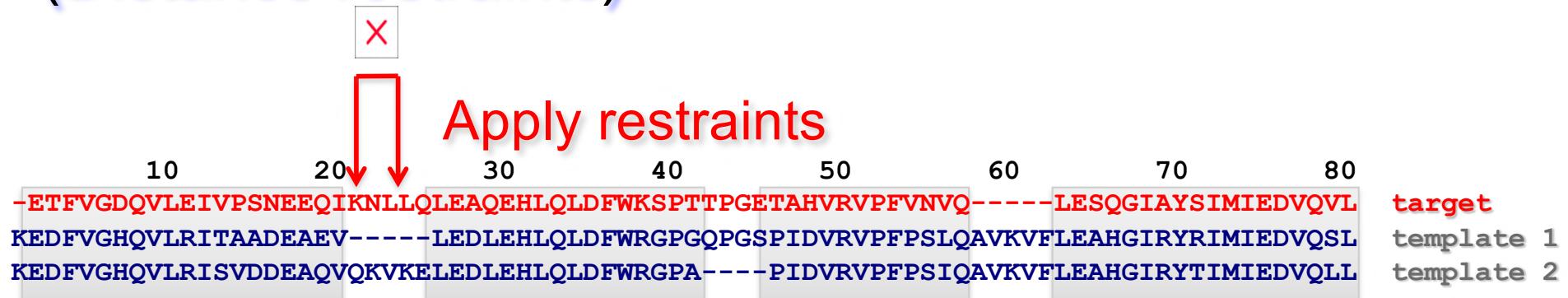
3. Model building: Spatial restraints (Distance restraints)



Distance restraints between Aa in SCR & VR
(required to locate the conformation of the VR)

2. Schema of the method

3. Model building: Spatial restraints (Distance restraints)

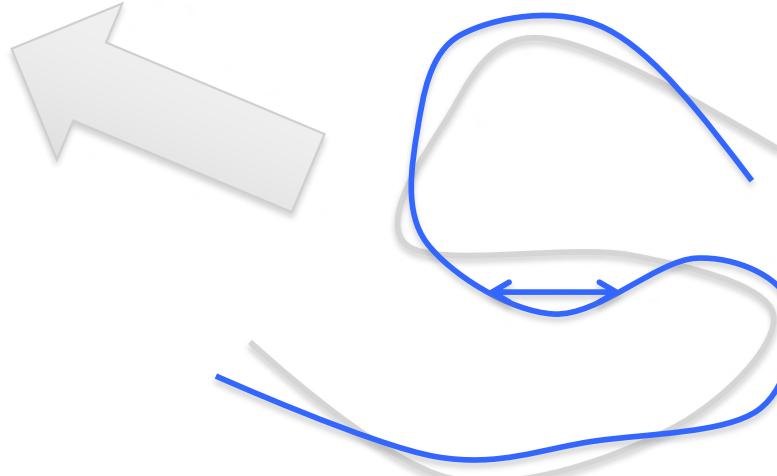


Obtain restraints

d_1

$$d \in \{d_{\text{template1}}\}$$

$$d_1 = d$$



Distance restraints between Aa in VR & VR
(required to obtain the conformation of the VR)

2. Schema of the method

3. Model building: Spatial restraints (Simulated annealing)

Optimizing a target function:

1. Start with e.g. a random conformation model and use only local restraints
2. Minimize some steps using a conjugate gradient optimization and molecular dynamics steps
3. Repeat, introducing more and more long range restraints until all restraints are used

$$E_{bonding} = \sum_{bonds} \frac{1}{2} k_i (d_i - d_i^0)^2 + \sum_{angles} \frac{1}{2} k_j (\alpha_j - \alpha_j^0)^2 + \sum_{improper \ dihedral} \frac{1}{2} k_n (\omega_n - \omega_n^0)^2 + \sum_{angles} E_m \cos(\omega_m \phi_m + \varphi_m)$$

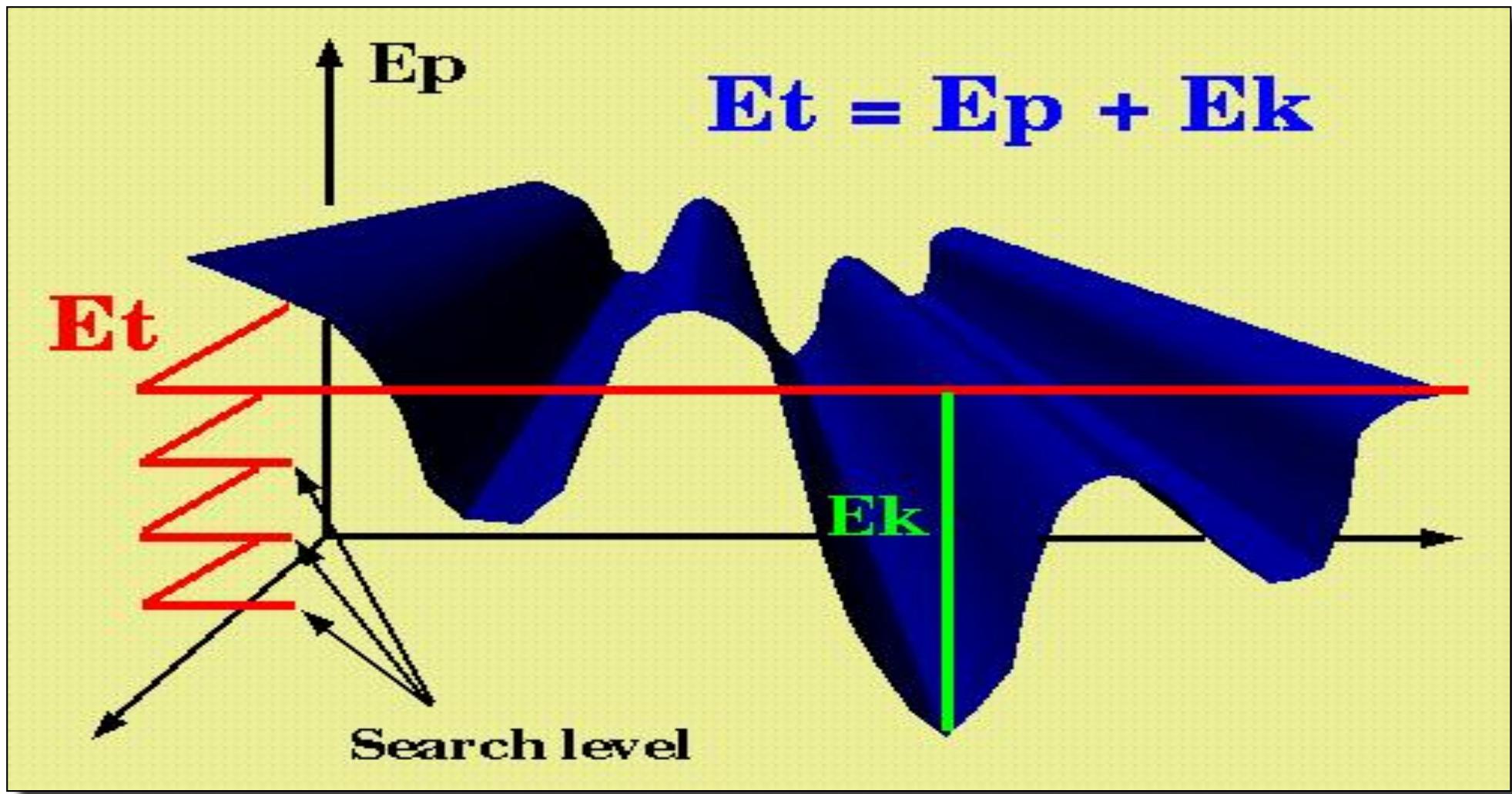
$$E_{non-bonding} = \frac{1}{4\pi\epsilon_0} \sum_i \sum_{j>i} \frac{q_i q_j}{r_{ij}} + \sum_i \sum_{j>i} \frac{C_6^{ij}}{r_{ij}^6} - \frac{C_{12}^{ij}}{r_{ij}^{12}}$$

$$E_{dist} = \sum_{rest} \frac{1}{2} k_r (d_r - \langle d_r^0 \rangle)^2$$

$$E = E_{bonding} + E_{non-bonding} + E_{pdf} + E_{dist}$$

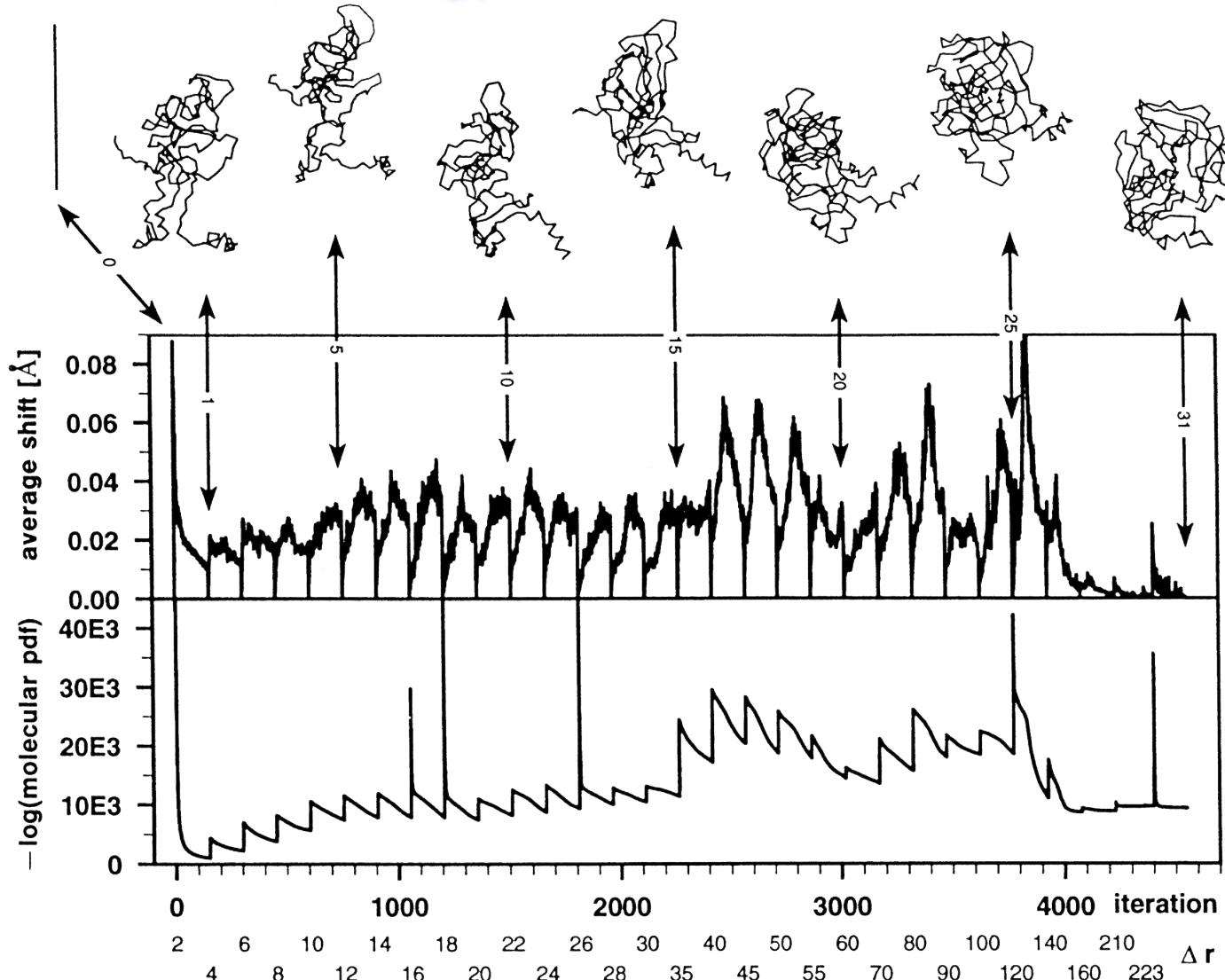
2. Schema of the method

3. Model building: Spatial restraints (Simulated annealing)



2. Schema of the method

3. Model building: Spatial restraints (Simulated annealing)



2. Schema of the method

3. Model building: Spatial restraints (Loop modeling using a database of loops)

Apply restraints

10 20 30 40 50 60 70 80

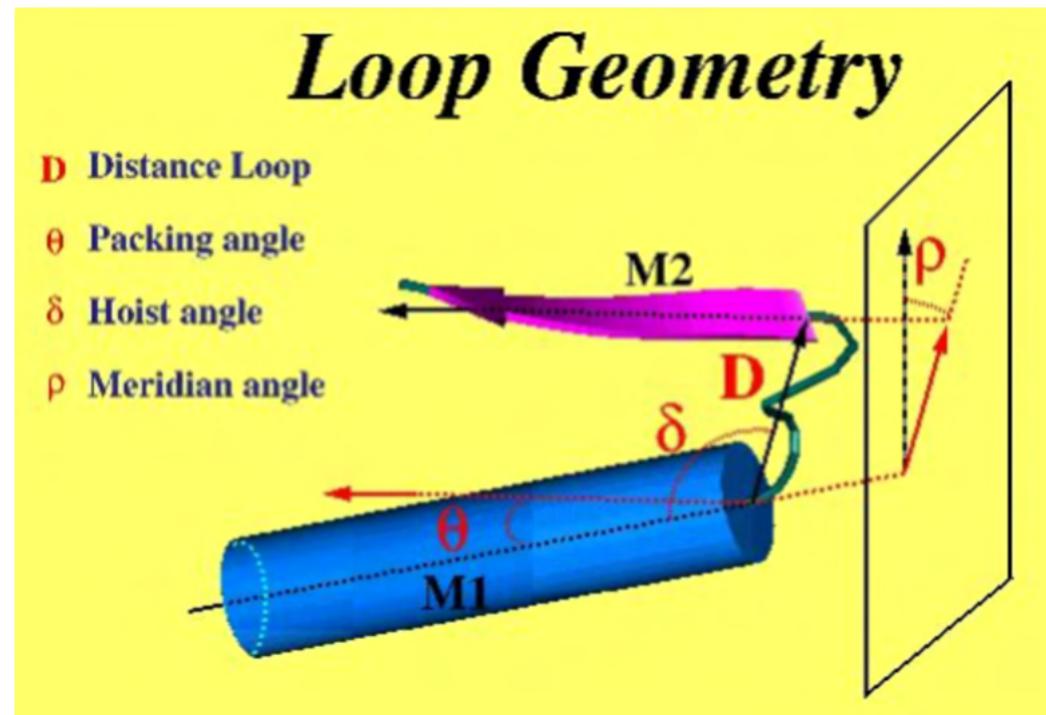
-ETFVG DQVLEI VPSNEE QIKNLL QLEA QEH LQLDF WKSP TPG ETA HVRV PFVN VQ ----- LESQ GIAY SIMI EDVQ VL

KEDF VGHQ VL RITA ADEAEV ----- LED LEHL QL DFWR GPG QPG SPID VR VP FPSL QAV KV FLEAH GIRY RIMI EDVQ SL

----- VD DEAQ VQ KV KELED LEH -----

target
template 1
template 2

Obtain restraints



2. Schema of the method

3. Model building: Spatial restraints (Loop modeling using a database of loops)



Using the structure of a known loop:

1. The C-tail and N-tail of the loop (template 2) when superposed with the core of the main template (template 1) produce a low RMSD
2. The selection of the loop follow two criteria: similar sequence profile with the target and similar anchoring geometry of the loop with the main template

2. Schema of the method

3. Model building: Spatial restraints (Loop modeling *ab initio*)

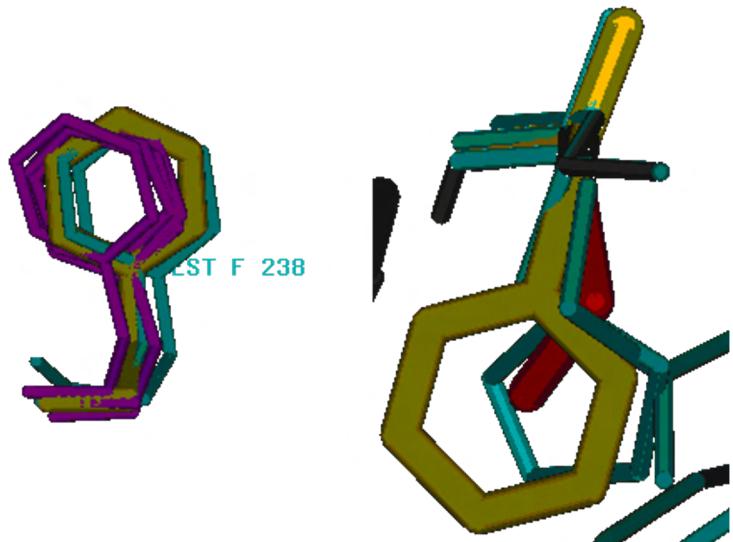
Using PDF of loops and minimization methods:

1. Calculate specific PDF residue properties of loops
2. Minimize by simulated annealing the loops
3. Extract main motion from normal modes on templates and apply them as restrictions on the conformational changes of the model
4. Methods:
 1. Loop-model from MODELLER
 2. ArchPred
 3. Rosetta

2. Schema of the method

3. Model building: Side-chains

Let be a rotamer library, we define the probability of side-chain “i” in conformation “k” as $\text{CM}(i,k)$. Initially $\text{CM}(i,k)=1/K_i$, where K_i is the total of rotamers of residue “i”.

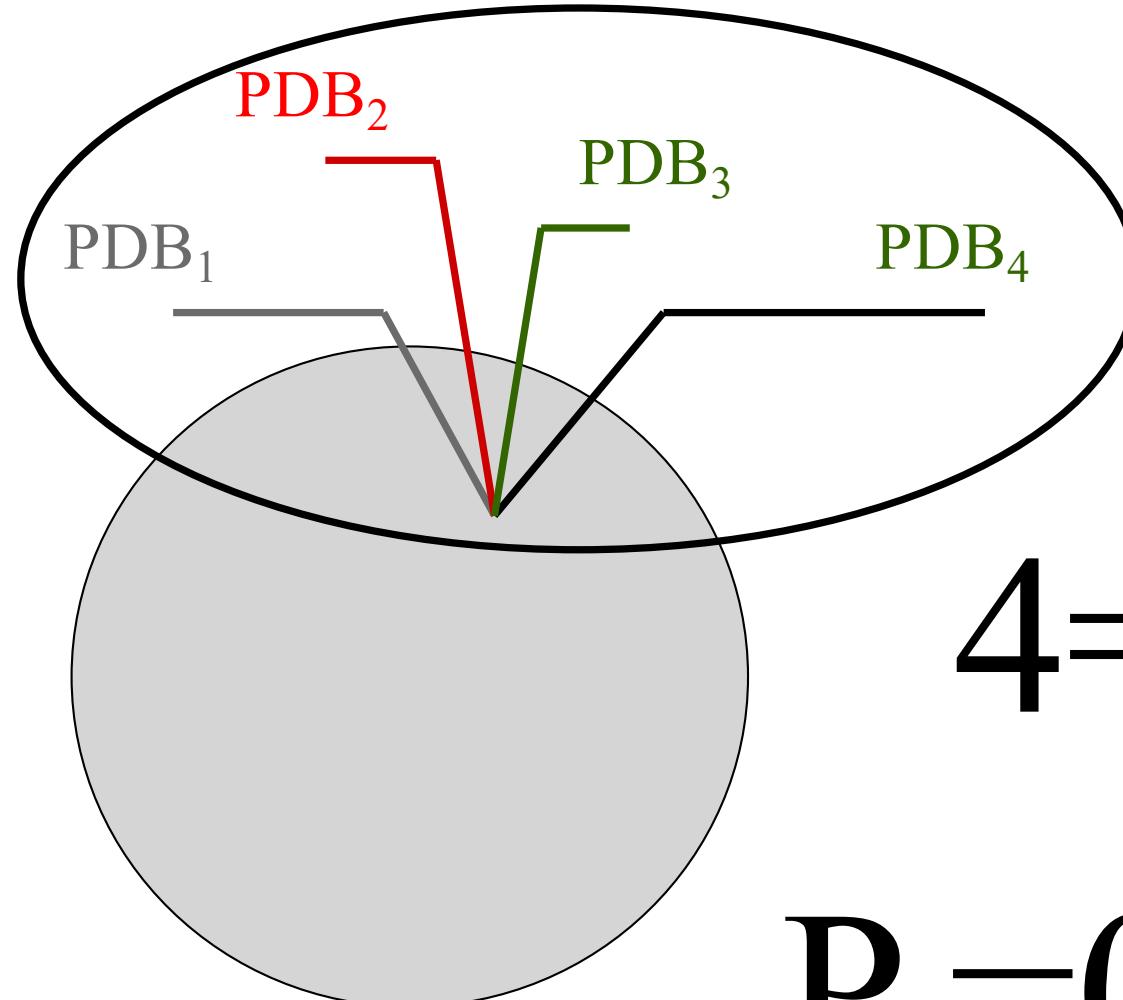


Multi-copy.

Koehl and Delarue *J.Mol. Biol.* (1994) **239**, 249-275

2. Schema of the method

3. Model building: Side-chains



$$4=K_i$$

$$P_i=0.25$$

2. Schema of the method

3. Model building: Side-chains

Given U the total potential energy of the protein and its environment, we define the effective potential of rotamer “ k ” of residue “ i ” as $E(i,k)$, where:

$$E(i,k) = \int \rho(x)U(i,k,X)dX;$$

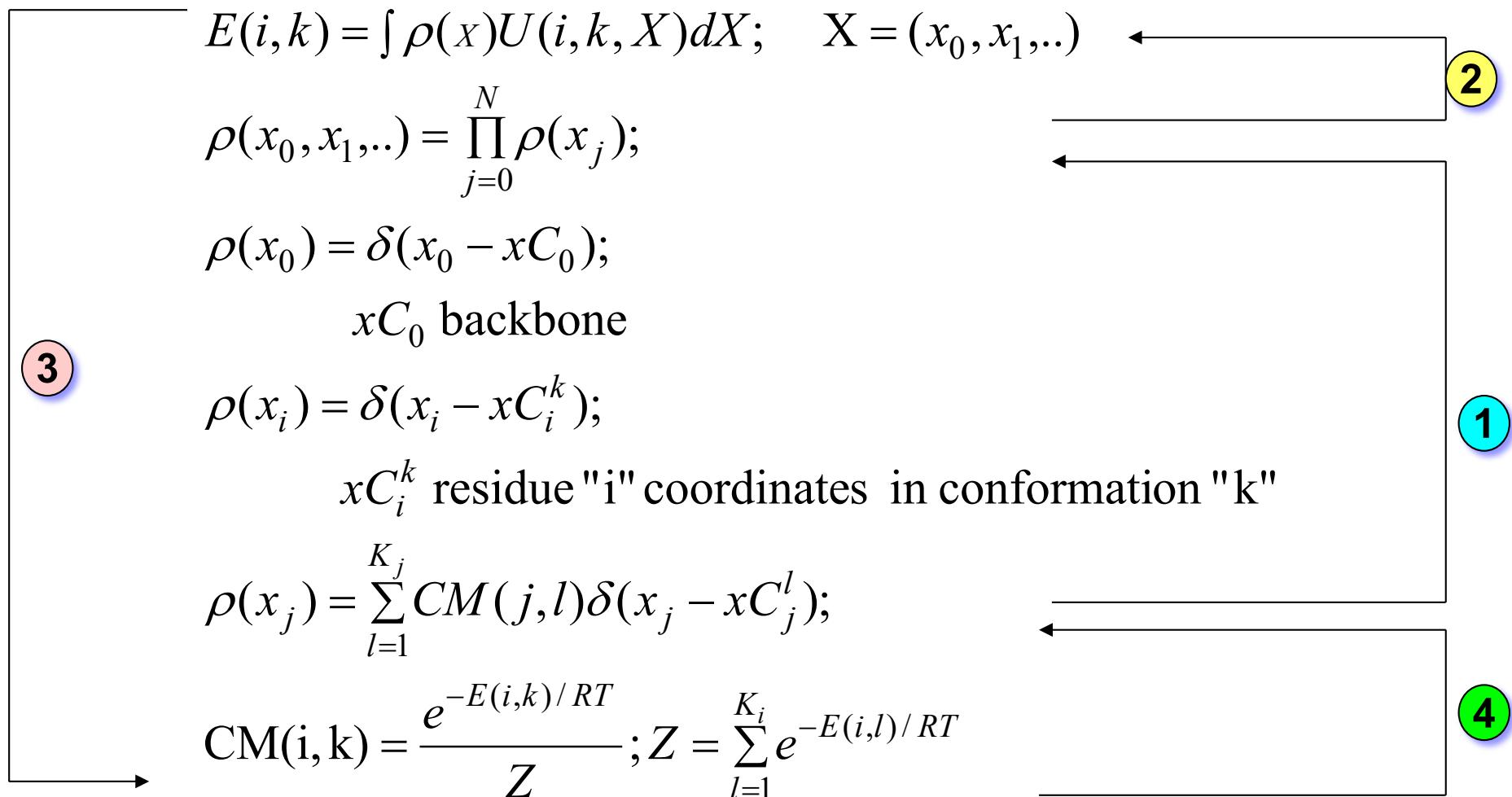
$$X = (x_0, x_1, \dots)$$

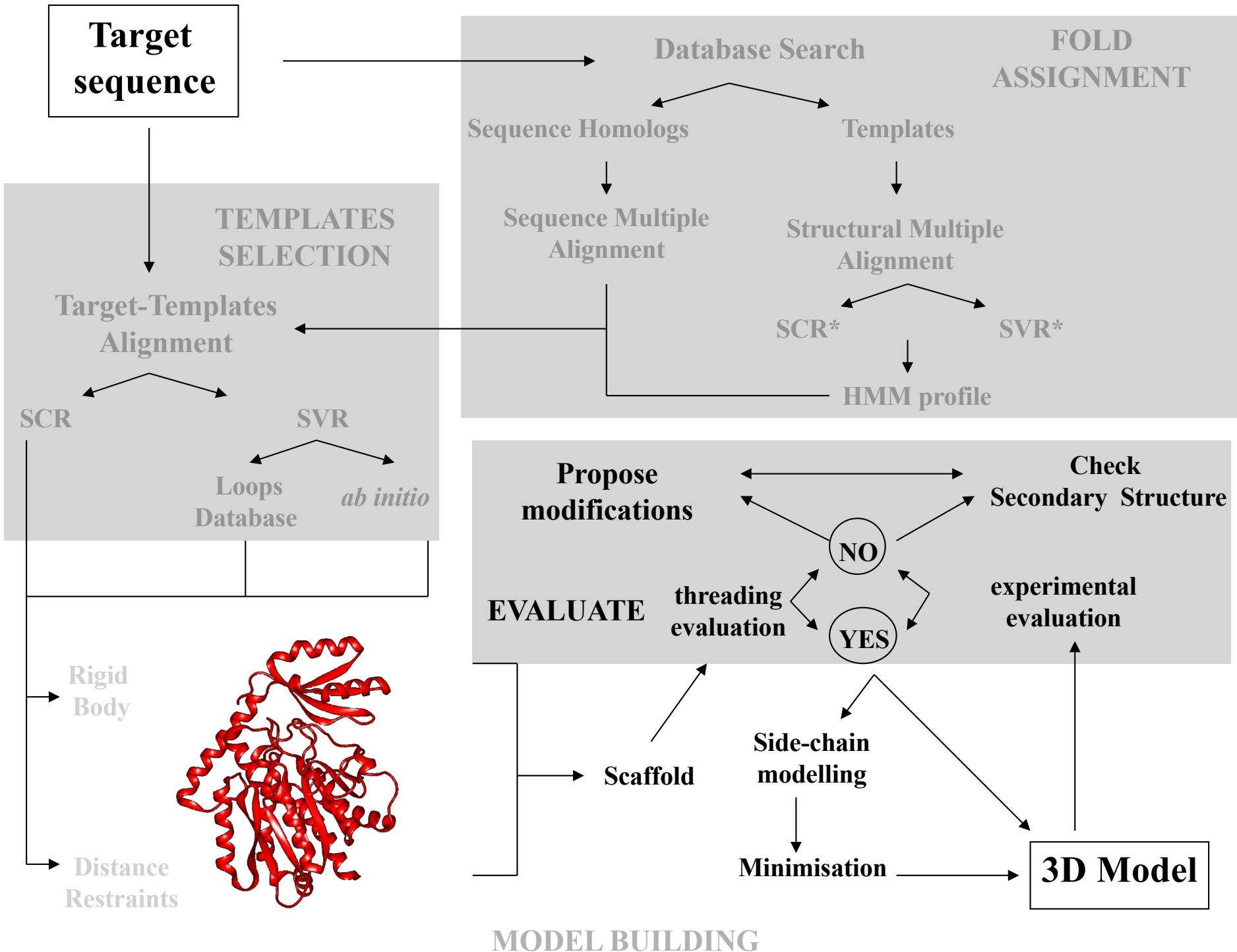
U is obtained with $E_{\text{non-bonding}}$, E_{bonding} on a system that includes the protein and water molecules

2. Schema of the method

3. Model building: Side-chains

Iterative optimization





2. Schema of the method

4. Evaluation

Types of Errors

1. *Errors in side-chain packing .*
2. *Shifts of correctly aligned residues .*
3. *Regions without template .*
4. *Errors due to misalignments .*
5. *Errors produced by incorrect templates .*

2. Schema of the method

4. Evaluation

Shifts of correctly aligned residues

HHHHHHHHH HHH .HHC

GARFIELD THE .CAT

GARFIELD THE CCAT

Solution

HHHHHHHHH HHH HHC.

GARFIELD THE CAT.

GARFIELD THE CCAT

2. Schema of the method

4. Evaluation

Errors due to misalignments .

GARFIELD THE CAT ...

GARFIELD THE FAT CAT

Solution

GARFIELD THE ... CAT

GARFIELD THE FAT CAT

2. Schema of the method

4. Evaluation

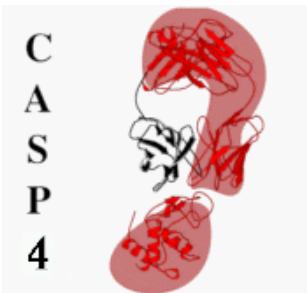
How to test the model?

1. Compare the RMSD between the model and the real structure
2. Check that secondary structures are correctly aligned
3. Calculate the percentage of residues that are closer than a threshold after superposing the model and the real structure
4. Calculate the percentage of identical residues aligned when superposing the real structure and the model.
5. Check the energy of threading to compare the real structure and the model (see next chapter)

2. Schema of the method

4. Evaluation

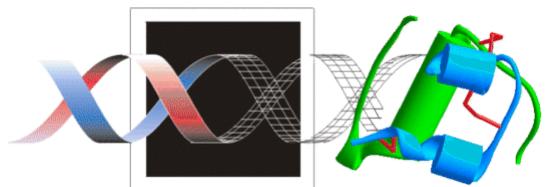
Model Accuracy Evaluation



CASP
Community Wide Experiment on the Critical Assessment of
Techniques for Protein Structure Prediction
<http://PredictionCenter.llnl.gov/casp5/>



EVA
Evaluation of Automatic protein structure prediction
[Burkhard Rost, Andrej Sali, <http://maple.bioc.columbia.edu/eva/>]



3D - Crunch
Very Large Scale Protein Modeling Project

http://www.expasy.org/swissmod/SM_LikelyPrecision.html

2. Schema of the method

5. Improvement

How to detect possible errors in the model if we don't know the solution?

- 1. Compare the model and all the templates**
- 2. Check that secondary structures are not broken**
- 3. Check if the prediction of secondary structure agrees with the secondary structure of the model**
- 4. Check if the loops of the target are similar to some loops in the database of loops and they agree in sequence and anchoring geometry**
- 5. Check the capping of helices**
- 6. Check pseudo-energies of threading and compare the model with the templates.**

2. Schema of the method

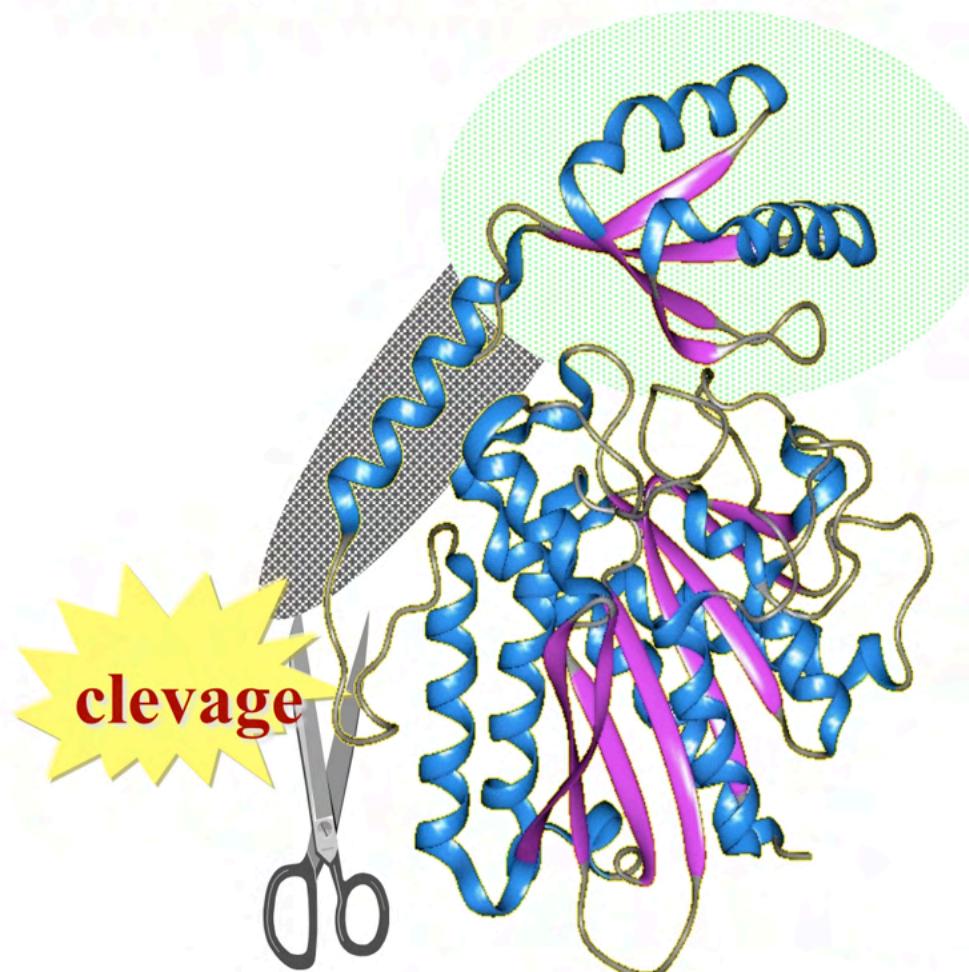
5. Improvement

How to improve the model?

1. Decide the changes in the alignment according to the secondary structure prediction or the structure of the templates and recalculate the model
2. Change the main template and recalculate the model
3. Include new templates
4. Calculate the main motion of normal modes from the templates of the homologous family and optimize by molecular dynamics under motion restrictions the conformation
5. Recalculate the pseudo energy profile of the new model and compare with the original model to test the improvement

EXAMPLE

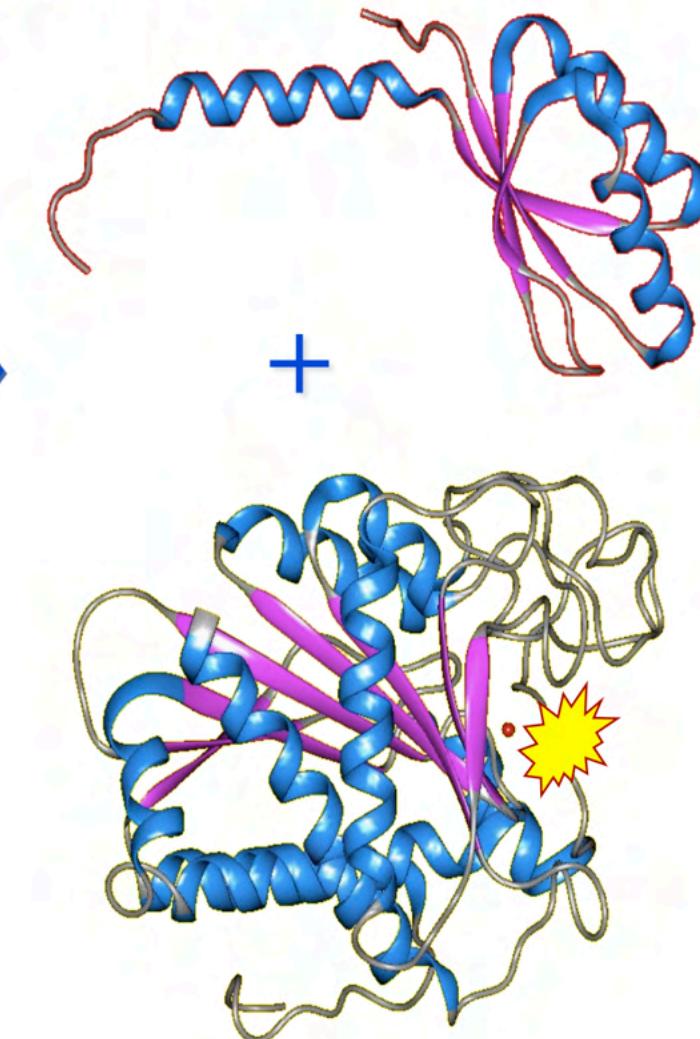
PRO-CARBOXIPEPTIDASA



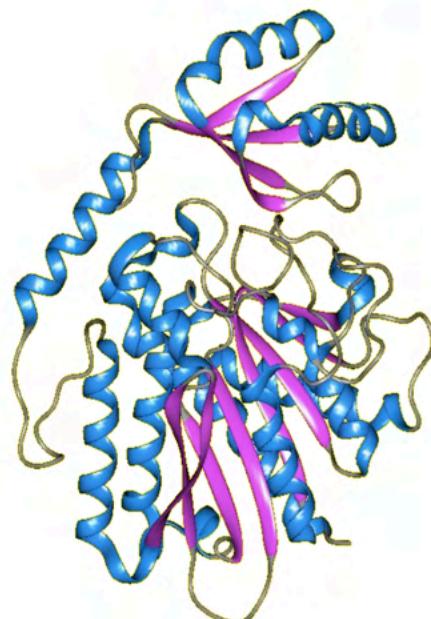
Enzime

activation

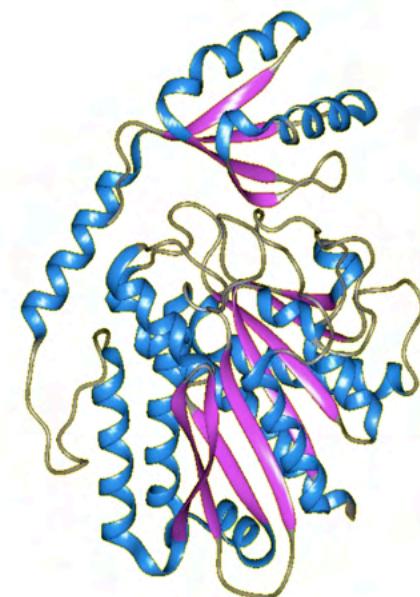
Activation segment



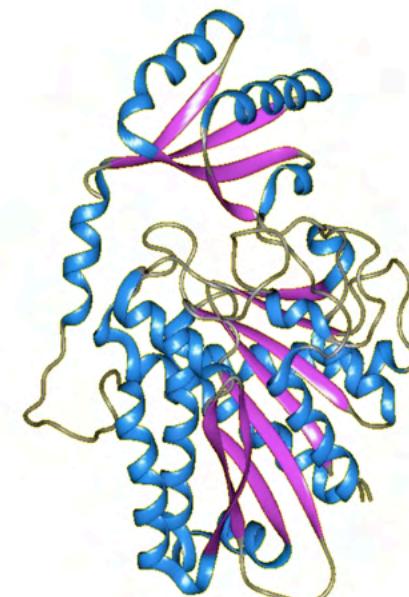
PRO-CARBOXYPEPTIDASES



Bovine
Pro Carboxypeptidase A1
PCPA1b



Porcine
Pro Carboxypeptidase A1
PCPA1p



Porcine
Pro Carboxypeptidase B
PCPBp

SEQUENCE ALIGNMENT OF PRO-CARBOXYPEPTIDASES

10	20	30	40	50	60	70	80		
-ETFVGDQVLEIVPSNEEQIKNLLQLEAQEHLQLDFWKSPTPGETAHVRVPFVNQAVKVFLESQGIAYSIMIEDVQVL								PCPA2h	
KEDFVGHQVLRITAADEAEVQTVKELEDLEHLQLDFWRGPQPGSPIDVRVPFPSIQAQKVFLEAHGIRYRIMIEDVQSL								PCPA1b	
KEDFVGHQVLRISVDDEAQVQKVKELEDLEHLQLDFWRGPARPGFPIDVRVPFPSIQAQKVFLEAHGIRYRIMIEDVQLL								PCPA1p	
90	100	110	120	130	140	150	160		
LDKENEEMLFNRRRERSGN-FNFGAYHTLEEISQEMDNLVAEHPGLVSKVNIGSSFENRPMNVLKFKSTGG-DKPAIWIDA								PCPA2h	
LDEEQEQMFASQSRARSTNTFNYATYHTLDEIYDFMDLLVAEHPQLVSKLQIGRSYEGRPIYVLKFSTGGSNRPAIWIDL								PCPA1b	
LDEEQEQMFASQGRARTTSTFNYATYHTLEEIYDFMDILVAEHPALVSKLQIGRSYEGRPIYVLKFSTGGSNRPAIWIDS								PCPA1p	
170	180	190	200	210	220	230	240		
GIHAREWVTQATALWTANKIVSDYGKDPSITSILDALDI	FLL	PVTNP	DGYVFSQTKNRMWRKTRSKVGSCLCVGVDPNRN					PCPA2h	
GIHSREWITQATGVWFACKFTEDYQDPSFTAILDSMDIFL	EIVT	NPDGF	AFTHSQNRLWRKTRSVSSLCVGVDANRN					PCPA1b	
GIXSRXWITQASGVWFACKITENYGQNSFTAILDSMDIFL	EIVT	NPNGF	AFTHSDNRLWRKTRSKASGSLCVGSDSNRN					PCPA1p	
250	260	270	280	290	300	310	320		
WDAGFGGPASSNPCSDSYHGPSANSEVEVKSIVDFIKSHGKVKA	FI	I	LHSYSQ	LLMFPYGYKCTK	LDDF	DELSEVAQKA		PCPA2h	
WDAGFGKAGASSSPCSETYHGKYANSEVEVKSIVDFVKDHGNFK	AFLSI	HSYSQ	LLL	YPYGYTTQSIPDKTEL	NQVAKSA			PCPA1b	
WDAGFGGAGASSSPCAETYHGKP	NSEVEVKSITDFVKNNGNI	KA	FI	SIXSYSQ	LLL	YPYGYKTQSPADKSELNQIAKSA		PCPA1p	
330	340	350	360	370	380	390	400		
AQSLRSLHGT	KVGPICSVIYQASGGSIDWSYDYG	IKYSFA	ELRTG	RYGFLLPARQ	I	LP	TAEETWLGLKAI	MEHVRD	PCPA2h
VEALKSLYGT	SYKYGSITT	IYQASGGSIDWSYNQ	GIKYSFTF	ELRTG	RYGFLLP	AQETWL	GVLTIMEHTLN		PCPA1b
VAALKSLYGT	SYKYGSITVIYQASGGVIDWTY	NQGIKYSFS	FELRTG	RRGFLLP	AQETWL	ALL	TIMEHTLN		PCPA1p

Secondary Structure Prediction and Multiple Alignment of Pro-Carboxypeptidases

10	20	30	40	50	60	70	80	
EEEEEE HHHHHHHHHHHHHHH EEEEEE EEEEEE HHHHHHHHHHH EEEEHHHHHHHH PCPA2h PHD								
-ETFVGDQVLEIVPSNEEQIKNLLQLEAQEHQLQDFWKSPTTPGETAHVRVPFVNQAVKVFLESQGIAYSIMIEDVQVL PCPA2h								
KEDFVGHQVLRITAADEAEVQTVKELEDLEHLQDFWRGPQPGSPIDVRVPFFPSLQAVKVFLEAHGIRYRIMIEDVQSL PCPA1b								
KEDFVGHQVLRISVDDEAQVKVKELEDLEHLQDFWRGPARGFPIDVRVPFFPSIQAVKVFLEAHGIRYTIMIEDVQLL PCPA1p								
90	100	110	120	130	140	150	160	
HHHHHHHHHHHHHHH PCPA2h PHD								
LDKENEEMLFNRRRERSGN-FNFGAYHTLEEISQEMDNLVAEHPGLVSKVNIGSSFENRPMNVLKFKSTGG-DKPAIWDA PCPA2h								
LDEEQEQMFASQSRARSTNTFNYATYHTLDEIYDFMDLLVAEHPQLVSKLQIGRSYEGRPIYVILFKSTGGSNRPAIWIDL PCPA1b								
LDEEQEQMFASQGRARTTSTFNYATYHTLEEIYDFMDILVAEHPALVSKLQIGRSYEGRPIYVILFKSTGGSNRPAIWIDS PCPA1p								
170	180	190	200	210	220	230	240	
GIHAREWVTQATALWTANKIVSDYGDPSITSILDALDIFLPVTNPDGVSQTKNRMRKTRSKVSGSLCVGVDPNRN PCPA2h								
GIHSREWITQATGVWFACKFTEDYQDPSFTAILEDSDMIFLEIVTNPDGFATHSQNRLWRKTRSVTSSLCVGVDANRN PCPA1b								
GIXSRXWITQASGVWFACKKITEINYQNSSFTAILEDSDMIFLEIVTNPNGFAFTHSDNRLWRKTRSKASGSLCVGSDSNRN PCPA1p								
250	260	270	280	290	300	310	320	
WDAGFGGPGASSNPCSDSYHGPSANSEVEVKSIVDFIKSHGKVKAFIGIILHSQLLMFPYGYKCTKLDDFELSEVAQKA PCPA2h								
WDAGFGKAGASSSPCSETYHGKYANSEVEVKSIVDFVKDHGNFKAFLSIHSYSQLLLYPYGYTTQSIPDKTELNQVAKSA PCPA1b								
WDAGFGGAGASSSPCAETYHGKYNSEVEVKSITDFVKNNNGNIKAFISIXSYSQLLLYPYGYKTQSPADKSELNQIAKSA PCPA1p								
330	340	350	360	370	380	390	400	
AQSLRSLHGTKYKGPICSVYQASGGSIDWSYDYGIKYSFAFELRTGRYGFLLPARQILPTAETWLGLKAIMEHVRD PCPA2h								
VEALKSLYGTSYKGSIITTIYQASGGSIDWSYNQGIKYSFTFELRTGRYGFLLPASQIIPTAQETWLGVLTIMEHTLN PCPA1b								
VAALKSLYGTSYKGSIITVIYQASGGVIDWTYNQGIKYSFSFELRTGRRGFLLPASQIIPTAQETWLALLTIMEHTLN PCPA1p								

α -Helix C-cap Schellman Motif

Schellman Motif Profile
 $C'' \gg C3 / C' \text{ Gly}$

C3	C2	C1	Ccap	C'	C''
L	Q	E	H	G	I
A	R	A	K	G	V
M	K	K	L	G	K

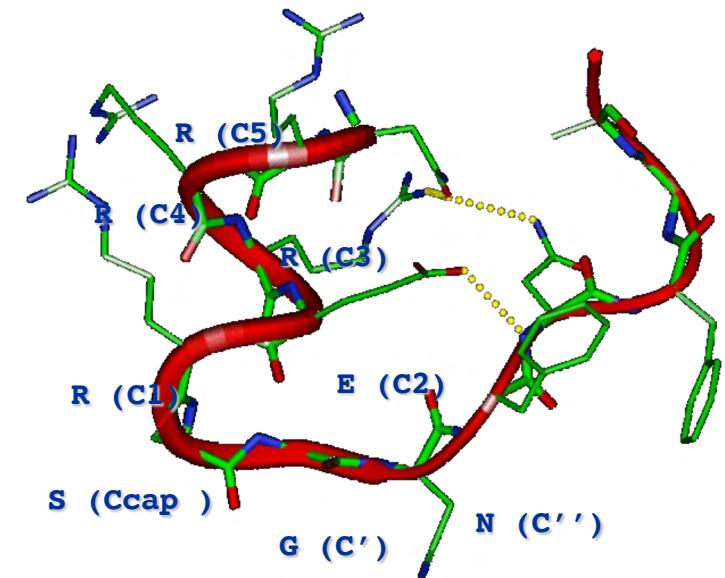
Hbond

hydrophobic

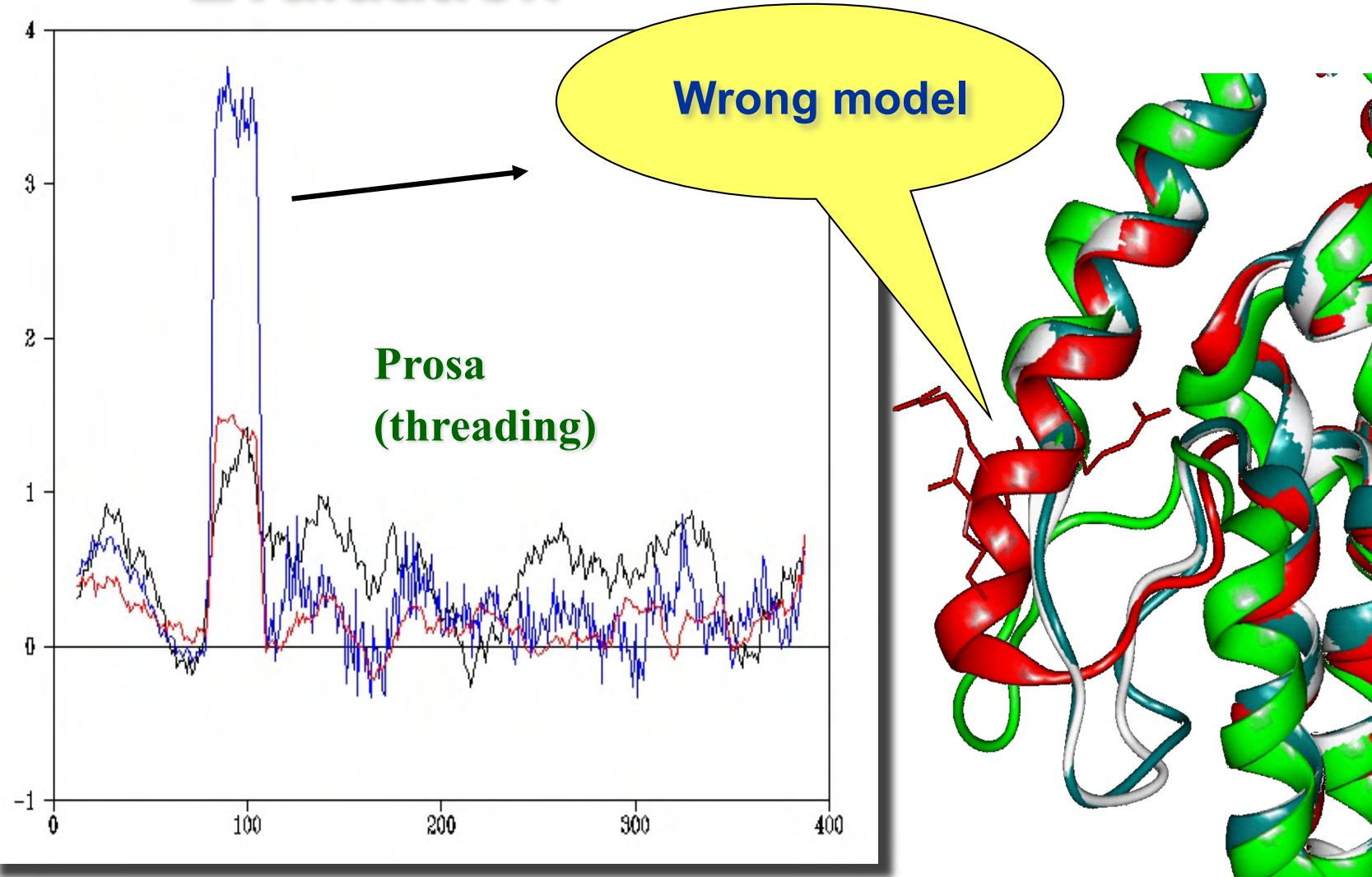
C3 C C' C''
 N R R R E R S G N F N

↑ ↑

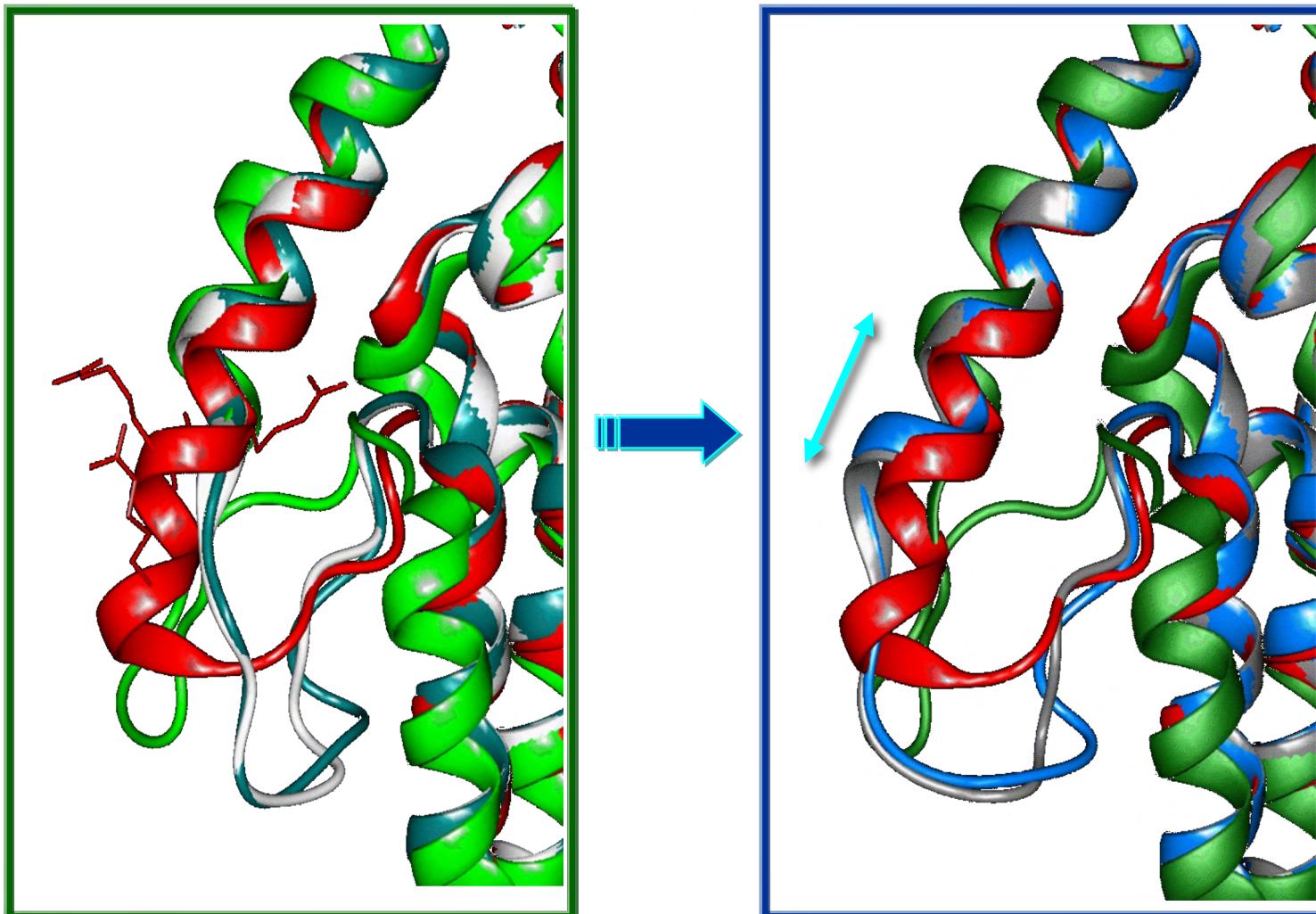
Hbond

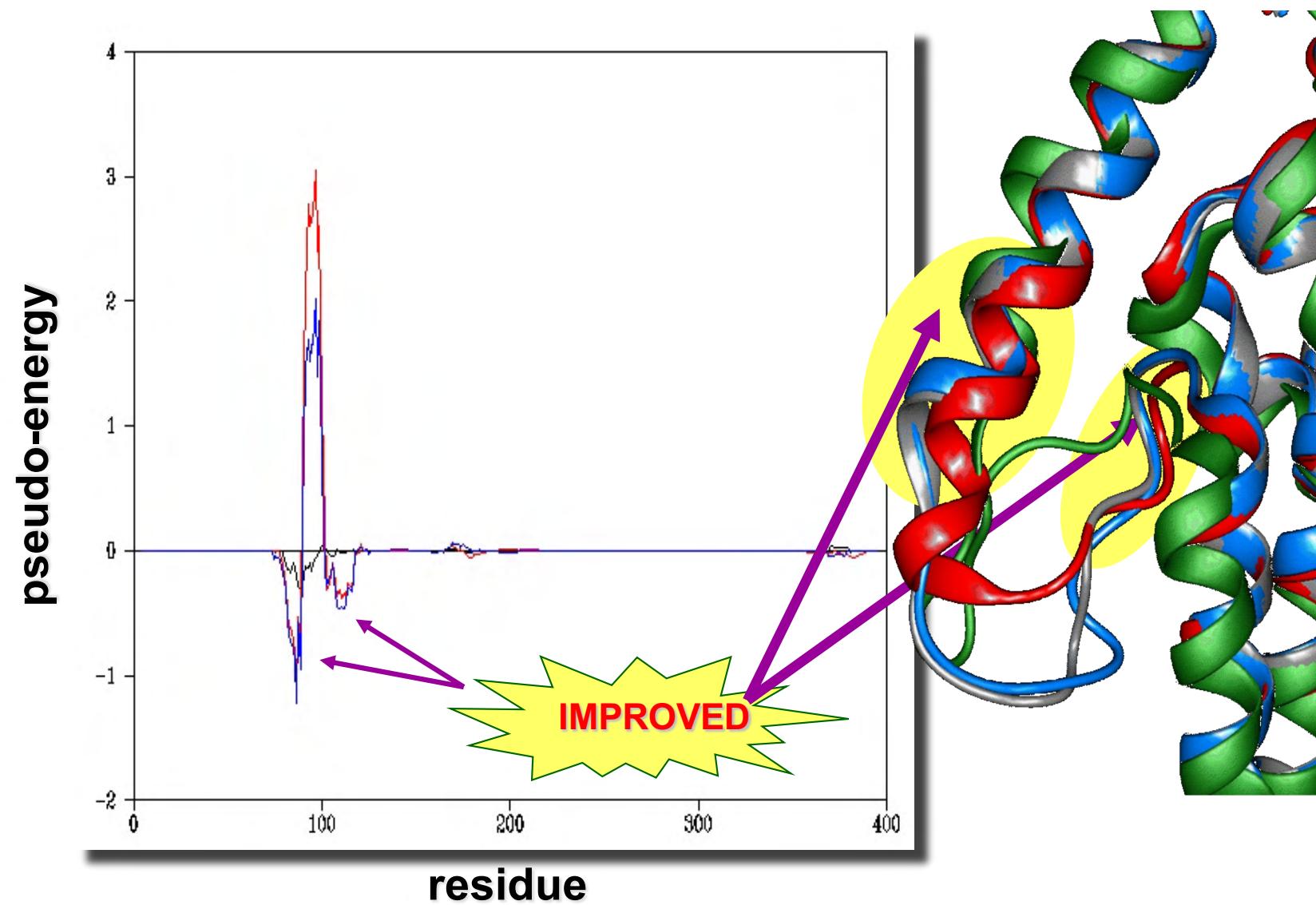


Evaluation



Refinement of the Model





Annex

Protein Structure Resources

PDB <http://www.pdb.org>

PDB - Protein Data Bank of experimentally solved structures (RCSB)

CATH <http://www.biochem.ucl.ac.uk/bsm/cath/>

Hierarchical classification of protein domain structures

SCOP <http://scop.mrc-lmb.cam.ac.uk/scop/>

Alexey Murzin's Structural Classification of proteins

DALI <http://www2.ebi.ac.uk/dali/>

Lisa Holm and Chris Sander's protein structure comparison server

SS-Prediction and Fold Recognition

PHD <http://cubic.bioc.columbia.edu/predictprotein/>

Burkhard Rost's Secondary Structure and Solvent Accessibility Prediction Server

3DPSSM <http://www.sbg.bio.ic.ac.uk/~3dpssm/>

Fold Recognition Server using 1D and 3D Sequence Profiles coupled with Secondary Structure and Solvation Potential Information.

Annex

Protein Homology Modeling Resources

SWISS MODEL: <http://www.expasy.ch/swissmod/>

Deep View - SPDBV:

homepage: <http://www.expasy.ch/spdbv/>

Tutorials <http://www.usm.maine.edu/~rhodes/SPVTut/>
<http://www.bbsrc.ac.uk/molbiol/>

WhatIf <http://www.cmbi.kun.nl/whatif/>

Gert Vriend's protein structure modeling analysis program WhatIf

Modeller: <http://guitar.rockefeller.edu/modeller/>

Andrej Sali's homology protein structure modelling by satisfaction of spatial restraints

FAMS: <http://physchem.pharm.kitasato-u.ac.jp/FAMS/fams.html>

Full Automatic Modelling System (FAMS); Kitasato University; Tokyo, Japan

3D-JIGSAW: <http://www.bmm.icnet.uk/people/paulb/3dj/form.html>

Comparative Modelling Server; Imperial Cancer Research Fund; London, UK

CPHmodels: <http://www.cbs.dtu.dk/services/CPHmodels/>

Centre for Biological Sequence Analysis; The Technical University of Denmark; Denmark

SDSC1: <http://cl.sdsc.edu/hm.html>

SDSC Structure Homology Modelling Server; San Diego Supercomputing Centre