Dynamics of Protein-Protein Interactions: A Probabilistic Model Toward Protein Function

Amir Vajdi

Computer Science Department
University of Massachusetts Boston

PhD Dissertation Defense,
November 28, 2018

Committee Members

- Prof. Nurit Haspel (Advisor)
- Prof. Kourosh Zarringhalam (Mathematics Department)
- Prof. Dan Simovici
- Prof. Ming Ouyang

My research projects

- Clustering co-expressed genes using time series data (IEEE BIBM 2015)
- Chromosomal structural variation detection using Jaccard distance (IEEE BIBM 2017)
- Computational biomarker discovery for cancer data based on RNA-Seq profiles(2017)
- Identifying significant TFs in Toxoplasma gondii cell cycle (2018-now)
- Human Gait Database (2017-now)
- Learning structural information as a penalty for Protein-Protein interface prediction (2017-2018)
- Simulation of protein trajectory between open and closed conformations using Monte Carlo tree search method (2016-2017)
- Clustering protein conformations changes (BICOB 2016)

1. Biology Background
    - Protein Structure
    - Protein Function
2. Protein-Protein Interaction Interface Prediction
    - Research Problem and Related Work
    - Probabilistic Graphical Model
    - Our New Proposed Method
3. Simulating Trajectories of Conformational Changes in Proteins and Identifying Intermediate Clusters
    - Research Problem
    - Monte Carlo Tree Search Method for Simulation of Conformational Changes
    - Clustering Coformational Changes using Geometric-Based Distance Function
4. Questions and Answers

1. Biology Background
    - Protein Structure
    - Protein Function
2. Protein-Protein Interaction Interface Prediction
    - Research Problem and Related Work
    - Probabilistic Graphical Model
    - Our New Proposed Method
3. Simulating Trajectories of Conformational Changes in Proteins and Identifying Intermediate Clusters
    - Research Problem
    - Monte Carlo Tree Search Method for Simulation of Conformational Changes
    - Clustering Coformational Changes using Geometric-Based Distance Function
4. Questions and Answers

Central Dogma of molecular biology

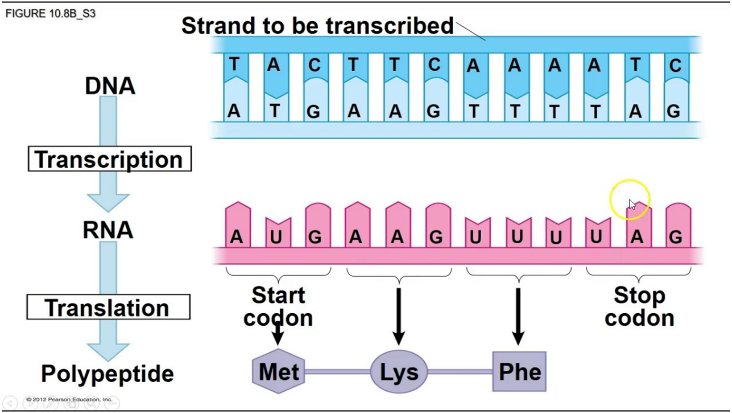
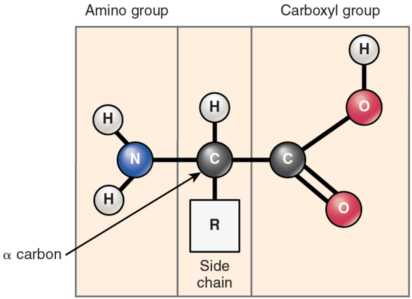


FIGURE 10.8B_S3

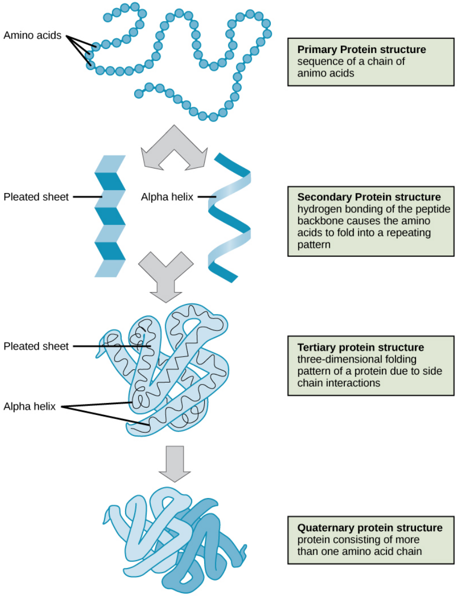
© 2012 Pearson Education, Inc.

Molecular structure of an Amino Acid



- Every Amino Acid has Amino group, C-α, and Carboxyl group
- Amino Acids are different in side chain

Four main representations of a protein



Primary Protein structure
sequence of a chain of animo acids

Secondary Protein structure
hydrogen bonding of the peptide backbone causes the amino acids to fold into a repeating pattern

Tertiary protein structure
three-dimensional folding pattern of a protein due to side chain interactions

Quaternary protein structure
protein consisting of more than one amino acid chain

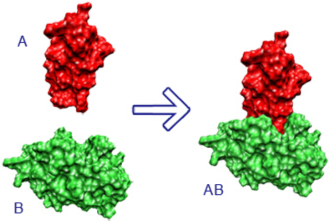
1. Biology Background
   - Protein Structure
   - Protein Function
2. Protein-Protein Interaction Interface Prediction
   - Research Problem and Related Work
   - Probabilistic Graphical Model
   - Our New Proposed Method
3. Simulating Trajectories of Conformational Changes in Proteins and Identifying Intermediate Clusters
   - Research Problem
   - Monte Carlo Tree Search Method for Simulation of Conformational Changes
   - Clustering Coformational Changes using Geometric-Based Distance Function
4. Questions and Answers

Research problem

Given two protein A and B, what are residues from protein A interacting with residues from protein B?
Two residues are contacting if the distance between them are less than n $\AA$

Challenges:

- Large search space
- Interface between two proteins is a small fraction of their surface
- Binding site has a complex behavior and it is vary across different complexes



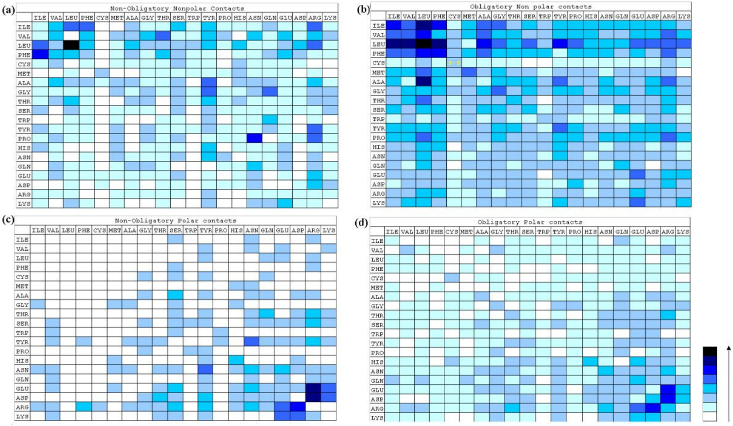
Protein binding site properties

Proteins binding site predictive features are:

- Binding sites are located on surface of the protein (Accessible Surface Area)
- Amino Acid interaction propensity
- Stability of native complex
- Conformational Changes between open and closed structure
- Formed as a set of patches
- Conservation of center of patch among homologous proteins
- Co-evolution of neighbour residues to center of patch among homologous proteins
- Secondary structure (α-Helix and β-Sheet)

There is no general rule. Protein types behave differently from each other.

Amino Acid interaction propensity is different among complex types


(a) Non-Obligatory Nonpolar Contacts

(b) Obligatory Non polar contacts

(c) Non-Obligatory Polar contacts

(d) Obligatory Polar contacts

De, Subhajyoti, et al. "Interaction preferences across protein-protein interfaces of obligatory and non-obligatory components are different." BMC Structural Biology (2005)

Related work

PSICOV

Jones, David T., et al. ”PSICOV: precise structural contact prediction using sparse inverse covariance estimation on large multiple sequence alignments.” Bioinformatics (2012)

GREMLIN

Ovchinnikov, Sergey, et al. ”Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information.” Elife (2014)

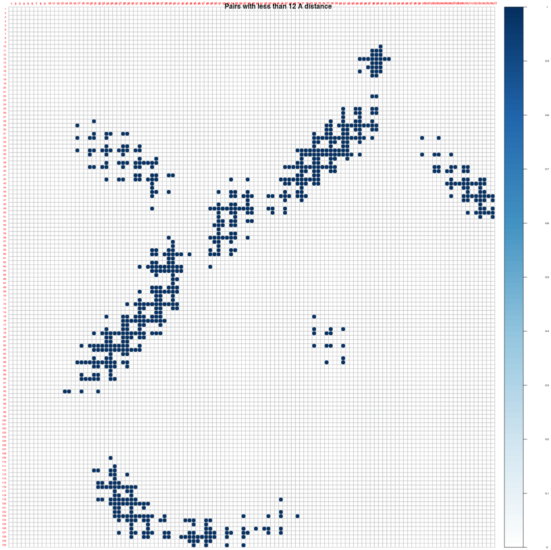
Meta-PSICOV

Jones, David T., et al. ”MetaPSICOV: combining coevolution methods for accurate prediction of contacts and long range hydrogen bonding in proteins.” Bioinformatics (2015)

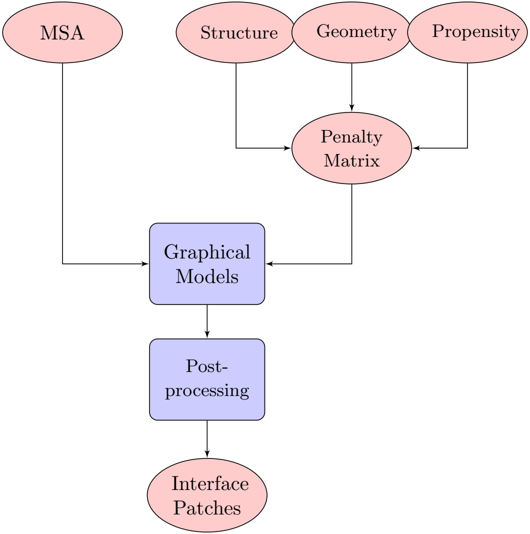
ComplexContact (RaptorX)

Zeng, Hong, et al. ”ComplexContact: a web server for inter-protein contact prediction using deep learning.” Nucleic acids research (2018).‘

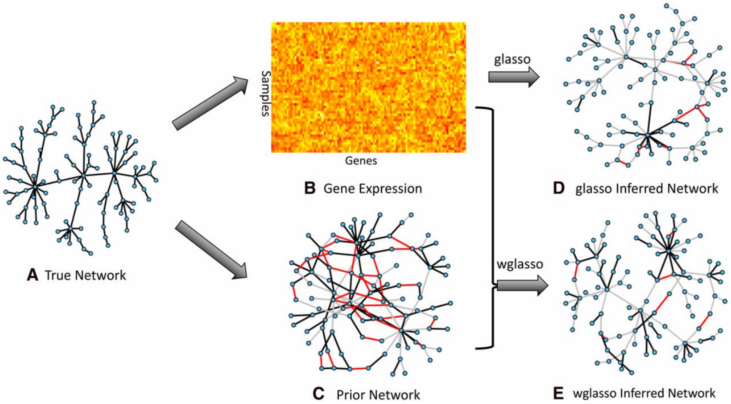
An example of contact map between two proteins



Flowchart of our proposed method



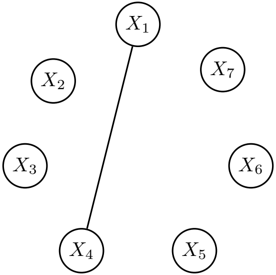
Probabilistic Graphical Models



Li, Yupeng, and Scott A. Jackson. "Gene network reconstruction by integration of prior biological knowledge." G3: Genes, Genomes, Genetics (2015)

Graphical interpretation

X_1	X_2	X_3	X_4	X_5	X_6	X_7
S	Y	C	H	M	F	L
F	Y	P	W	A	R	A
S	Y	K	H	G	R	Q
S	Y	G	H	Q	F	Q
F	Y	N	W	Q	R	M
S	Y	R	H	Q	R	M
F	Y	K	W	A	F	L
F	Y	R	W	R	F	L



Gaussian Graphical Model (GGM)

Probability density function of sequence X

$$f_{\mu,\sum}(x) = (2\pi)^{\frac{-L}{2}}(det\sum)^{\frac{-1}{2}} exp(\frac{-1}{2}(x-\mu)^T(\sum)^{-1}(x-\mu)\Big), x \in R^L$$

by taking trace inner product from above

$$f_{\mu,\sum}(x) = exp\Big(\mu^T\theta x - \langle\theta, \frac{1}{2}xx^T\rangle - \frac{L}{2}log(2\pi) + \frac{1}{2}log(det(\theta)) - \frac{1}{2}\mu^T\theta\mu\Big)$$

where $\theta = (\sum)^{-1}$ is the inverse of covariacne matrix

Objective function of GGM

Maximum Likelihood estimation based on S

S is empirical (sample) covariance matrix.

$$S = \frac{1}{n}\sum_{i=1}^{n}(X^{(i)} - \bar{X})(X^{(i)} - \bar{X})^T \quad where \quad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X^{(i)}$$

$$\ell L(\mu, \sum) \propto -\frac{n}{2} log(det(\sum)) - \frac{n}{2} tr(S(\sum)^{-1}) - \frac{n}{2}(\bar{X} - \mu)^T(\sum)^{-1}(\bar{X} - \mu)$$

$$\max_{\hat{\theta}} \quad logdet(\hat{\theta}) - tr(S\hat{\theta}) \tag{1}$$

by adding L_1 penalty to above

$$\max_{\theta} log(det\theta) - tr(S\theta) - \Lambda||\theta||_1 \tag{2}$$

Blockwise coordinate descent

The objective function is solved using Graphical Lasso (GLasso) method by applying coordinate descent approach.

$$\omega = \begin{pmatrix} \omega_{11} & \hat{\omega}_{12} \\ \hat{\omega}_{12}^T & \omega_{22} \end{pmatrix}, S = \begin{pmatrix} S_{11} & \hat{s}_{12} \\ \hat{s}_{12}^T & s_{22} \end{pmatrix}$$

Where $\omega_{11}, S_{11} \in R^{(L-1)\times(L-1)}$, $\hat{\omega}_{12}, \hat{s}_{12}$ are vectors of size $L-1$, and ω_{22}, s_{22} are scalars.

Start with $\omega = S + \Lambda I$ and update ω iteratively.

$$\hat{\omega}_{12} = \min_y \{y^T \omega_{11}^{-1} y : ||y - \hat{s}_{12}||_\infty \leq \Lambda\}$$

Solution of $\hat{\omega}_{12}$ satisfies the above function is same as the solution of β in the following Lasso problem, since $\hat{\omega}_{12} = \omega_{11}\beta$

$$\min_\beta \{\frac{1}{2}||\omega_{11}^{\frac{1}{2}}\beta - b||^2 + \Lambda||\beta||_1\}, \quad where \quad b = \omega_{11}^{\frac{-1}{2}} \hat{s}_{12}$$

FRIEDMAN,J.H.and et all, Sparse inverse covariance estimation with the graphical lasso. Biostatistics (2008)

Structural based prediction features

Intpred performs interface prediction based on following features:

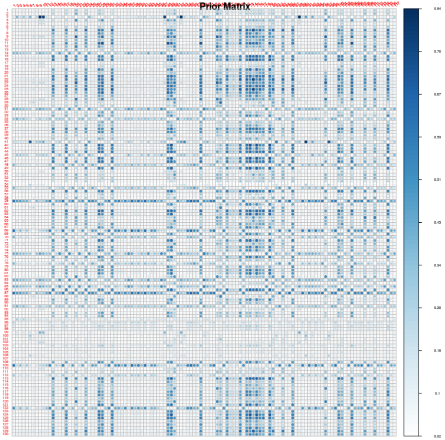
Feature	Description	Source
Hydrophobicity	Kyte and Doolittle hydrophobicity scale	Sequence
Homology	Homology Conservation Score Based on Valader01 Score	Sequence
Conservation	FEP Score for finding functionally equivalent orthologues	Sequence
Propensity	Residue Propensity based on position and type	Sequence and Structure
Disulfide Bonds	Disulfide Bridge with in 2.2 $\mathring{A}$ Distance + 10% tolerance	Structure
Hydrogen Bonds	Binary Score if exist any H Bonds	Structure
α-Helix	if percentages of α-Helix $>$0.2 and β-Sheet$\leq$0.2	Structure
β-Sheet	if percentages of α-Helix $\leq$ 0.2 and β-Sheet$>$0.2	Structure
mix	if percentages of α-Helix $>$0.2 and β-Sheet$>$0.2	Structure
Coil	if percentages of α-Helix $\leq$ 0.2 and β-Sheet $\leq$0.2	Structure
Planarity	RMSD of all atoms in a patch from best fitted Plane	Structure

Using random forest to predict the interface from non-interface residues and return probability of a residue is interface.

Northey, et all. "IntPred: a structure-based predictor of protein-protein interaction sites." Bioinformatics (2017)

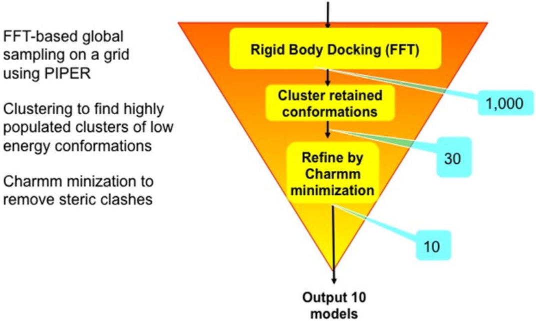
Building joint probability based on structural information

The joint probability matrix is $M^1 \in R^{n \times m}$ where n and m are number of residues in protein A and B, respectively. $P''_{i,j} = P_i \times P'_j$.



ClusPro docking algorithm

- Fast Fourier Transform (FFT) based search. One protein is placed on a fixed grid and the other on a moveable grid, and the search is conducted based on geometric and energetic constraints.
- Clustering the resulting conformations based on Interface RMSD.
- Filtering and refinement to remove steric clashes.


FFT-based global sampling on a grid using PIPER
Clustering to find highly populated clusters of low energy conformations
Charmm minization to remove steric clashes
Rigid Body Docking (FFT)
Cluster retained conformations
Refine by Charmm minimization
1,000
30
10
Output 10 models

Vajda S, et al. The ClusPro web server for protein-protein docking. Nature Protocols. 2017

Converting docking result to probability

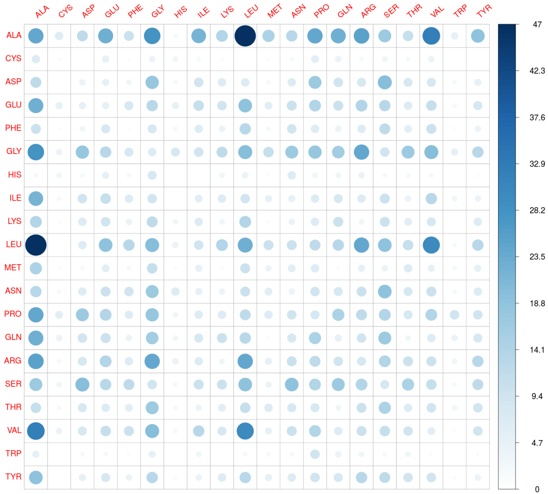
Probability matrix $M^2 \in R^{n \times m}$ is constructed where n, m are the number of residues for proteins A and B, respectively.

- set $M^2 = 0$
- for each predicted complex from ClusPro do the following
- calculate distance for every two residues between protein A and B.
- if distance between residues $i, j < 8\AA$ then $M^2 = M^2 + 1$

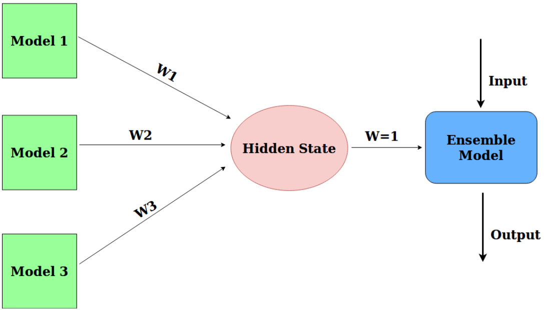
Perform Gaussian filter with kernel of size 3×3 for smoothing.

Normalize smooth matrix by diving every element by the maximum value of the matrix.

Amino Acid propensity in E.Coli proteins

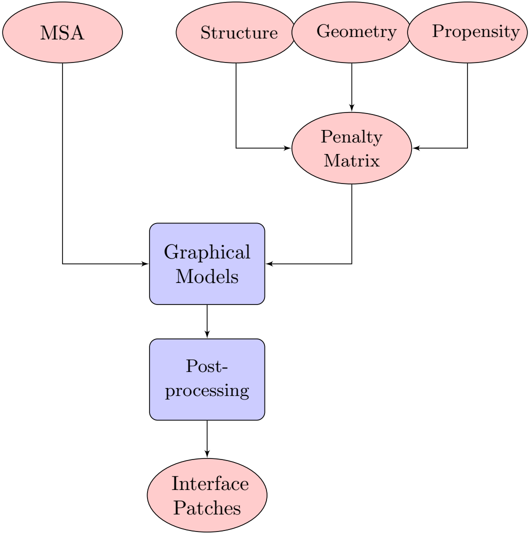


Ensemble average method for learning coefficients and turn it to penalty matrix



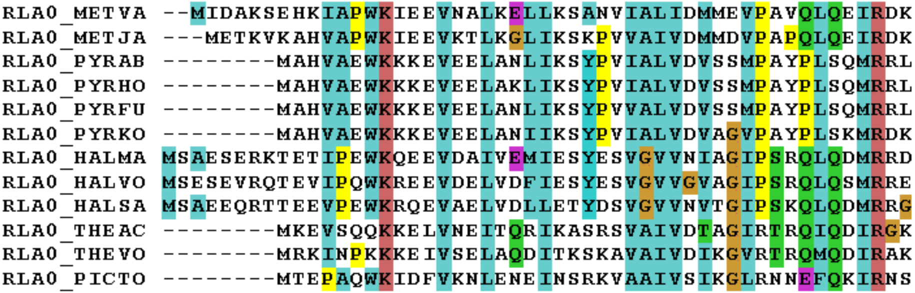
- M=$w_1 \times M^1 + w_2 \times M^2 + w_3 \times M^3$
- $\Lambda_{j,i} = \Lambda_{i,j} = \lambda_{max}$, where i, j belong to only one protein
- $\Lambda_{j,i} = \Lambda_{i,j} = \lambda_{min} + C \times \lambda_{min}(1 - \frac{M_{i,j} - min(M)}{max(M) - min(M)})$, where, i and j belong to protein A and B, respectively. $C = \frac{\lambda_{max}}{\lambda_{min}}$ is constant that obtained from training set.

Flowchart of our proposed method



Multiple Sequence Alignment (MSA) of Homologous proteins

For proteins A and B, the homologous proteins are identified and then by concatenating them with respect to species.
Then we represent each position in a MSA with binary vector of size 21.



```
RLA0_METVA  --MIDAKSEHKIAPWKIEEVNALKELLKSANVIALIDMMEVPAVQLQEIRDK
RLA0_METJA  ---METKVKAHVAPWKIEEVKTLKGLIKSKPVVAIVDMMDVPAPQLQEIRDK
RLA0_PYRAB  --------MAHVAEWKKKEVEELANLIKSYPVIALVDVSSMPAYPLSQMRRL
RLA0_PYRHO  --------MAHVAEWKKKEVEELAKLIKSYPVIALVDVSSMPAYPLSQMRRL
RLA0_PYRFU  --------MAHVAEWKKKEVEELANLIKSYPVVALVDVSSMPAYPLSQMRRL
RLA0_PYRKO  --------MAHVAEWKKKEVEELANIIKSYPVIALVDVAGVPAYPLSKMRDK
RLA0_HALMA  MSAESERKTETIPEWKQEEVDAIVEMIESYESVGVVNIAGIPSRQLQDMRRD
RLA0_HALVO  MSESEVRQTEVIPQWKREEVDELVDFIESYESVGVVGVAGIPSRQLQSMRRE
RLA0_HALSA  MSAEEQRTTEEVPEWKRQEVAELVDLLETYDSVGVVNVTGIPSKQLQDMRRG
RLA0_THEAC  --------MKEVSQQKKELVNEITQRIKASRSVAIVDTAGIRTRQIQDIRGK
RLA0_THEVO  --------MRKINPKKKEIVSELAQDITKSKAVAIVDIKGVRTRQMQDIRAK
RLA0_PICTO  --------MTEPAQWKIDFVKNLENEINSRKVAAIVSIKGLRNNEFQKIRNS
```

Post-processing of θ matrix with Average Product Correction

- In order to overcome to phylogenetics tree biases during building MSA and also homologous searching

  $Q_{ij} = (\sum_{i=1}^{20} \sum_{j=1}^{20} \theta_{ij})^{\frac{1}{2}}$

  $$\hat{Q}_{ij} = Q_{ij} - \frac{Q_{i.} \times Q_{.j}}{Q_{..}}$$

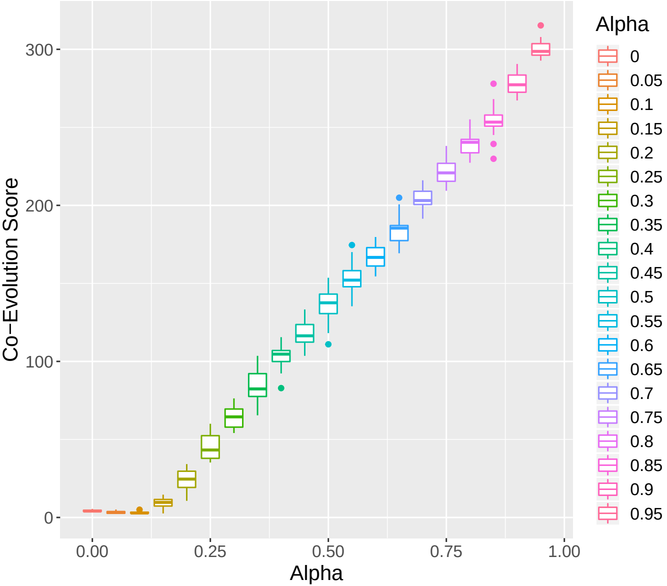
  $Q_{i.} = (\sum_{k=1}^{L} Q_{ik})$
  $Q_{.j} = (\sum_{k=1}^{L} Q_{kj})$
  $Q_{..} = (\sum_{k=1}^{L} \sum_{y=1}^{L} Q_{ky})$ And we sort pairs based on $\hat{Q}_{ij}$ score.

Generating simulated MSAs

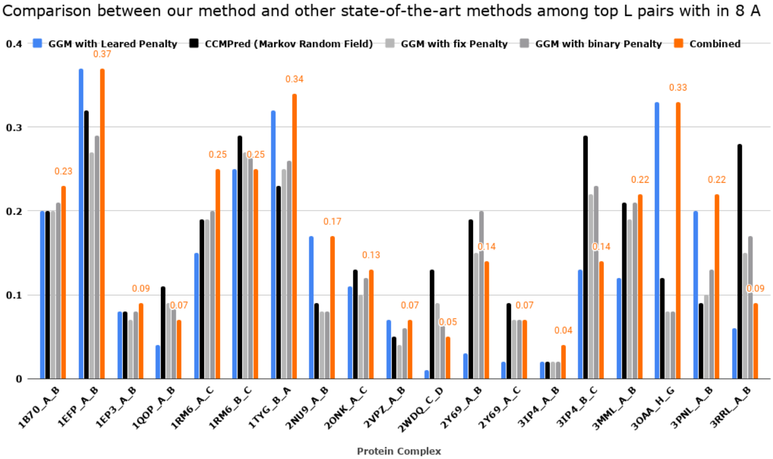
1^{st} order Hidden Markov Model (HMM) is used to generate multiple MSAs with different degree of co-evolution based on BLOSUM62 matrix. 3 parameters are used to generate MSAs with size of 1000×200 as following:

- Co-evolution parameter α: A a score between 0 and 1, controlling the transition probability of a 21^2 states HMM, with 0 corresponding to no co-evolution and 1 corresponding to maximum co-evolution.
- Conservation parameter C: The rate of Amino Acid change from one type to another. 0 means that we expect to see no conservation, and 1 represents that co-evolution occurs between 2 Amino Acid types.
- Bias control b: We have a fair bias which represents an original PAM matrix.

Simulated data result

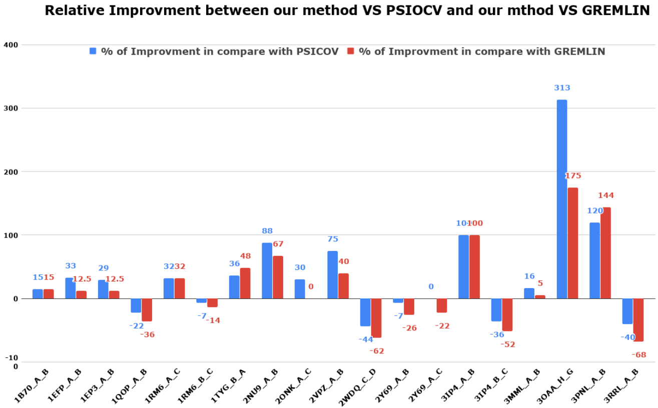


Precision comparison between our method and other state-of-the-art methods



Comparison between our method and other state-of-the-art methods among top L pairs with in 8 A

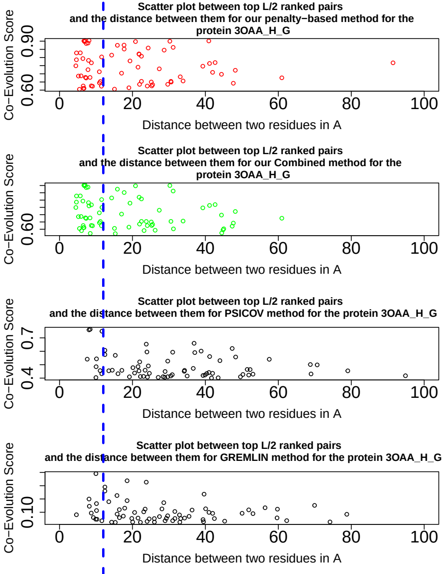
Relative precision improvement



Relative Improvment between our method VS PSIOCV and our mthod VS GREMLIN

Our method performs 40% and 20% better in compare with PSICOV and GREMLIN, respectively.

An example of top $L/2$ predicted pairs for 3OAA proteins between chains H and G



Scatter plot between top L/2 ranked pairs and the distance between them for our penalty−based method for the protein 3OAA_H_G

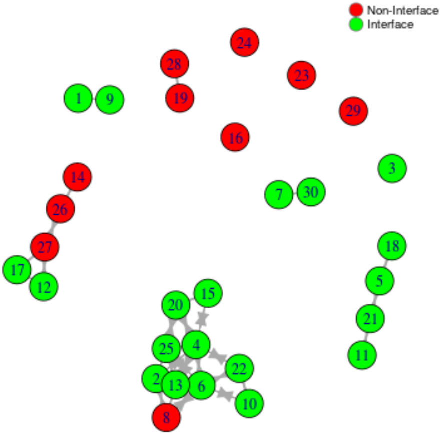
Scatter plot between top L/2 ranked pairs and the distance between them for our Combined method for the protein 3OAA_H_G

Scatter plot between top L/2 ranked pairs and the distance between them for PSICOV method for the protein 3OAA_H_G

Scatter plot between top L/2 ranked pairs and the distance between them for GREMLIN method for the protein 3OAA_H_G

Future work

Let us consider pair of residues i and j that is among top L ranked pairs. $Patch(r_i)$ is built for residue i in protein A, where every residue in $Patch(r_i)$ is within $6\mathring{A}$ form residue i. Jaccard Distance is calculted between every two patches.



Conclusion

- We found an upper bound for penalty in GLasso model.
- Learning structural information and imposing that as a penalty for GGM can significantly improve the performance of predicting binding site between two proteins.
- Structural information reveals new set of co-evolving pairs.
- Propensity matrix needs to calculated for each species differently from others.
- We are releasing parallel version of our method along with a package for more stable version of GLasso.

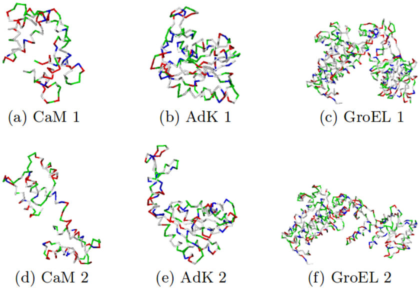
1. Biology Background
    - Protein Structure
    - Protein Function
2. Protein-Protein Interaction Interface Prediction
    - Research Problem and Related Work
    - Probabilistic Graphical Model
    - Our New Proposed Method
3. Simulating Trajectories of Conformational Changes in Proteins and Identifying Intermediate Clusters
    - Research Problem
    - Monte Carlo Tree Search Method for Simulation of Conformational Changes
    - Clustering Coformational Changes using Geometric-Based Distance Function
4. Questions and Answers

Relationship between binding site and conformational changes

Motivation: What is the association between the structure-dynamics of a protein and its function?

These can be studied in two steps:

- Identifying relationship between conformational space and protein function
- Identifying relationship between Highly populated region and local minima



(a) CaM 1

(b) AdK 1

(c) GroEL 1

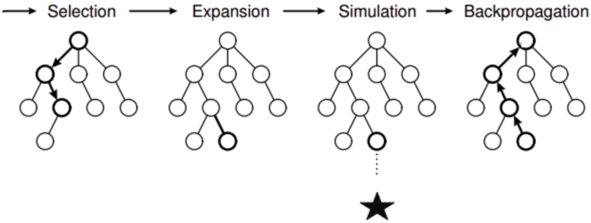
(d) CaM 2

(e) AdK 2

(f) GroEL 2

Monte Carlo Tree Search Method for Simulation of Conformational Changes

Given two open and closed conformationals, simulate the path that it takes to move from one conformation to another one.



C$-\alpha$ representation

For each protein P with L residues, it is represented with two coarse-grained models
C-α representation: size is L and the energy function is calculated as:

$$E_{total} = \sum_{angles} \frac{1}{2} k_{\theta}(\theta - \theta_0)^2 + \sum_{dihedrals} [A[1 + cos(\phi - \phi_0] + B[1 - cos(\phi + \phi_0)] + C[1 + cos3(\phi + \phi_0)] + D[1 + cos(\phi + \phi_0 + \frac{\pi}{4})]] + \sum_{i,j \geq i+3} 4\epsilon H S_1[\frac{\sigma}{r_{ij}^{12}} - S_2 \frac{\sigma}{r_{ij}^{6}}] + \sum_{HB} E_{HB}$$

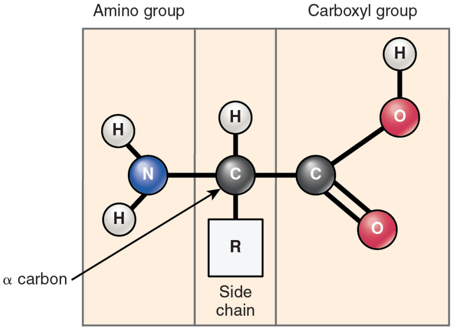
θ is angle defined by 3 consecutive C-α atoms
ϕ is dihedral angle defined by 4 consecutive C-α atoms.
ϵH is hydrophobic strength
$k_{\theta} = \frac{20\epsilon H}{rad^2}$ is bond angle force constant

Yap, EngHui, et all. "A coarsegrained carbon protein model with anisotropic hydrogenbonding." Proteins: Structure, Function, and Bioinformatics (2008)

Backbone representation

Backbone + C-β representation: size is $5 \times L$

$$E_{total} = E_{vdw} + E_{HB} + E_{burial} + E_{water} + E_{bond} + E_{angle}$$



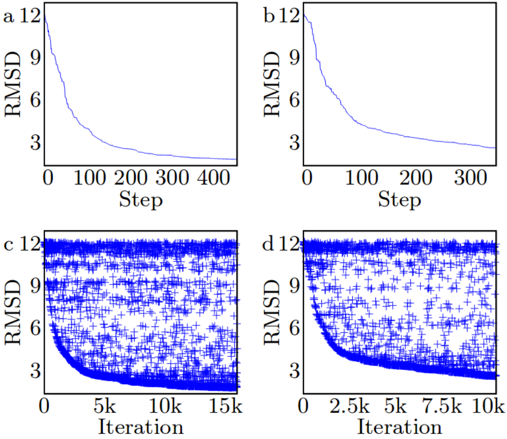
Papoian, Garegin A., et al. "Water in protein structure prediction." Proceedings of the National Academy of Sciences (2004)

Search methodology

Using biased Monte Carlo tree search method.

- start from one conformation and calculate the dihedral angle between every 4 consecutive residues.
- compare dihedral angles between current conformation with endpoint and pick the largest
- perturb the selected angle by $+/$-5 degree and update conformation
- if $RMSD_{child} < RMSD_{parent}$ or $r < e^{-\frac{RMSD_{child} - RMSD_{parent}}{A \times RMSD_{child}}}$ (A is a constant and r is a random number between 0 and 1), add the new conformation to the tree, otherwise start from root

Trajectory of conformational pool



a: C-α best path, b: backbone best path

c: C-α all conformations, d: backbone all conformations

Luo, Dong, and Nurit Haspel. "Multi-resolution rigidity-based sampling of protein conformational paths." Proceedings of the International Conference on Bioinformatics,

Result

Name	RMSD	Residues	PDB	Conformations
AdK	6.95	214	1AKE→4AKE	5,235
			4AKE→1AKE	6,588
Calmodulin	14.72	144	1CLL→1CTR	11,483
			1CTR→1CLL	3,232
GroEL	12.21	525	1SS8→1SX4	1,689
			1SX4→1SS8	1,528

Feature Vector Representation

The goal is to cluster the conformations into a intermediate clusters.

For each conformation C, we can represent every secondary element such as i as:

$$score(C^i) = \sum_{j \in K} \left(|\alpha_{ij} - \alpha'_{ij}| \times w + |d_{ij} - d'_{ij}| \times w' \right).$$

Where, K is all the manipulated secondary structures excluding i,

α_{ij} and $d_{i,j}$ are the angle and distance between i^{th} and j^{th} elements, respectively.

α'_{ij} and $d'_{i,j}$ are the angle and distance between i^{th} and j^{th} elements in goal structure, respectively. w and w' are weights which are equal to 1 and 5, respectively.

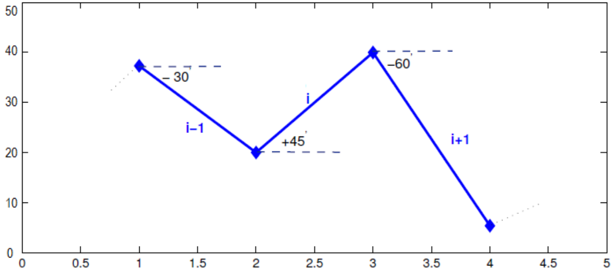
As a result each conformation C is represented in lower dimension:

$$v_C = \langle score(C^1), score(C^2), \ldots, score(C^k) \rangle$$

Haspel, Nurit, et al. "Tracing conformational changes in proteins." BMC structural biology (2010).

Distance metric

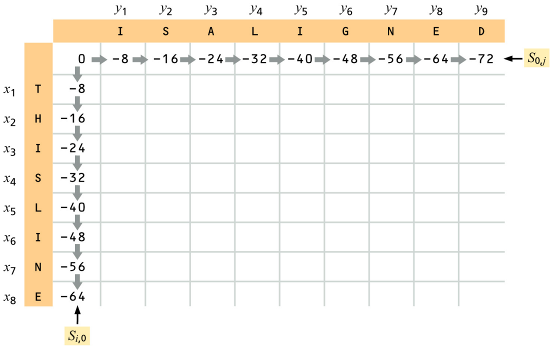
Size of V_C is between 8 and 15 depends on protein which also corresponds to a polygon. Every polygon P_C is represented as $P_C :< (L_1, A_1), (L_2, A_2), ..., (L_{n-1}, A_{n-1}) >$, where L_i and A_i are length of i^{th} feature and angle between i^{th} and $(i+1)^{th}$ features.



Building Matrix of scores

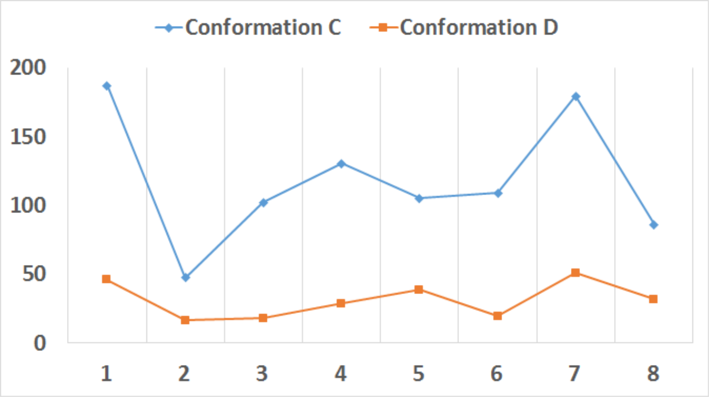
For two given polygons P_{C1} and P_{C2}, the score matrix is built between every two line as:

$$S(P_{C1}[i], P_{C2}[j]) \;=\; \omega_{length} \times (\; 1 - |L(i) - L(j)| \;) \\ + \;\omega_{angle} \times (\; 1/(\theta + |A(i) - A^{(}j)| \;)$$



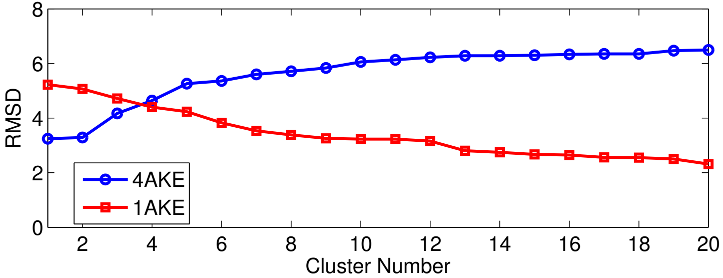
			y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
			I	S	A	L	I	G	N	E	D
		0	-8	-16	-24	-32	-40	-48	-56	-64	-72
x_1	T	-8									
x_2	H	-16									
x_3	I	-24									
x_4	S	-32									
x_5	L	-40									
x_6	I	-48									
x_7	N	-56									
x_8	E	-64									

Example of needleman wunsch

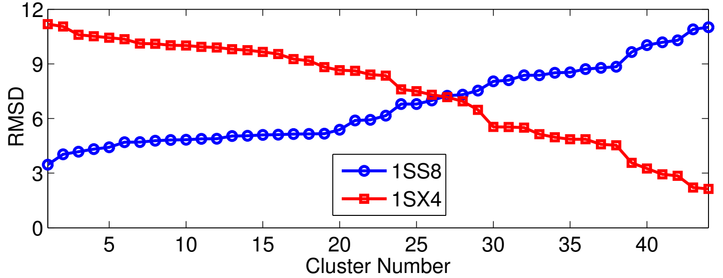


C: 1 2 3 4 G 6 7 8
D: 1 2 3 G 5 6 7 8
Similarity Score: 0.87

Result



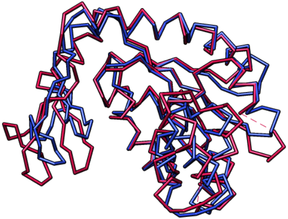
RMSD
Cluster Number
0
2
4
6
8
2
4
6
8
10
12
14
16
18
20
4AKE
1AKE



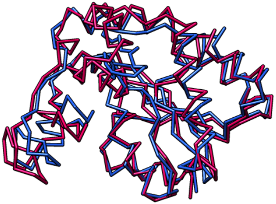
RMSD
Cluster Number
0
3
6
9
12
5
10
15
20
25
30
35
40
1SS8
1SX4

Alignment of simulated goal structure and actual goal structure

AdK



GroEL



PDB	closest cluster (RMSD)
1E4Y	Cluster 8 (2.3Å)
1DVR	Cluster 17 (2.6Å)
2RH5A	Cluster 19 (3.1Å)
2RH5B	Cluster 20 (3.0Å)
2RH5C	Cluster 20 (2.3Å)

Conclusion

- We represented a Monte Carlo based simulation for proteins dynamic
- We represented a clustering method that is compatible with protein 3D structure
- Centers of clusters can be investigated as an interesting conformationals

1. Biology Background
   - Protein Structure
   - Protein Function
2. Protein-Protein Interaction Interface Prediction
   - Research Problem and Related Work
   - Probabilistic Graphical Model
   - Our New Proposed Method
3. Simulating Trajectories of Conformational Changes in Proteins and Identifying Intermediate Clusters
   - Research Problem
   - Monte Carlo Tree Search Method for Simulation of Conformational Changes
   - Clustering Coformational Changes using Geometric-Based Distance Function
4. Questions and Answers

Acknowledgement

- Prof. Nurit Haspel
- Prof. Kourosh Zarringhalam
- Prof. Dan Simovici
- Prof. Ming Ouyang
- Prof. Todd Riley
- Dr. Sergey Ovchinnikov
- Prof. Haspel's Lab: Arpita Joshi
- Prof. Zarringhalam's Lab: Saman Farahmand and Yasaman Rezavani
- Prof. Riley's Lab: Andrew S Judell-Halfpenny
- Hamidreza Mohebbi

Thank you!