Networks and protein assembly

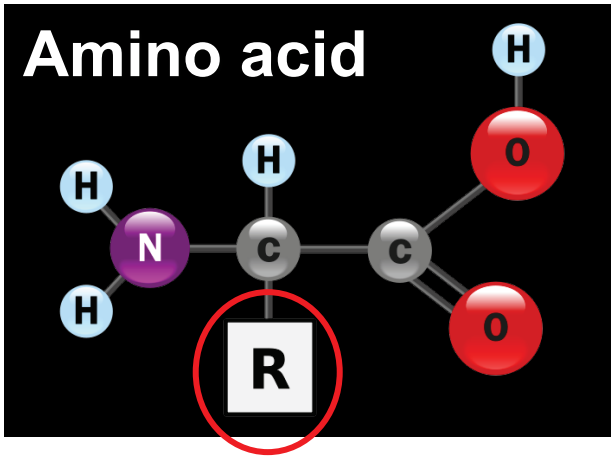
Claire Lesieur, Mounia Achoch, Kave Salamatian, Laurent Vuillon

Proteins

- Biological functions
- Broad shape diversity
- Shape/function relationship
  - Perturbation of local information: evolution/disease

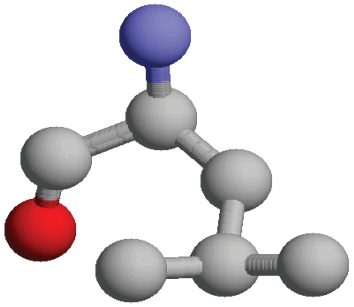
Local information: amino acids

- Amino acids = CHON
- Many different flavors, versatile properties
- Each one has a broad combinatory of possible interactions

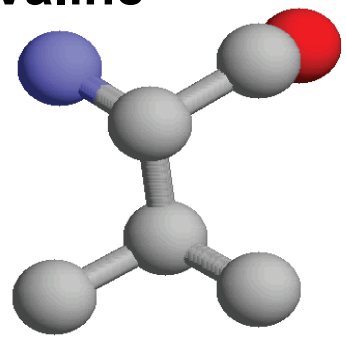


- Size
- Atom number
- Chemistry
- Geometry

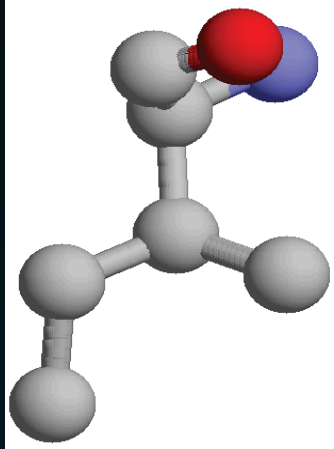
Leucine



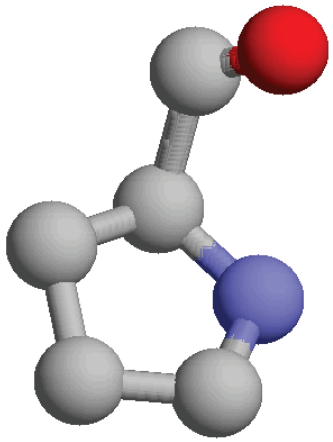
Valine



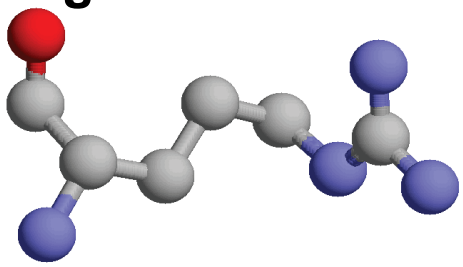
Isoleucine



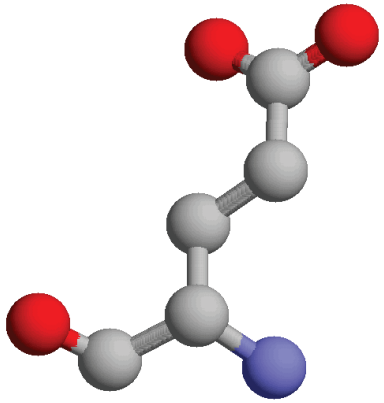
Proline



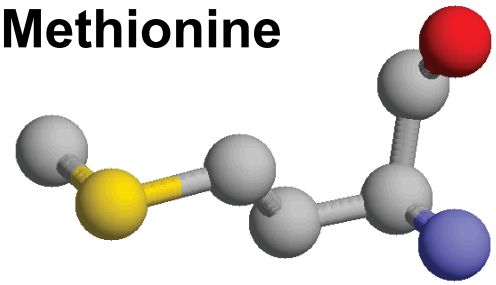
Arginine



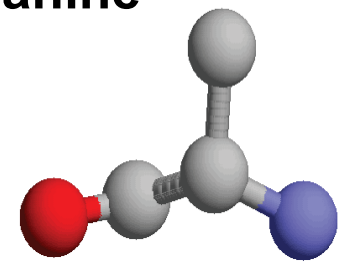
Glutamic acid



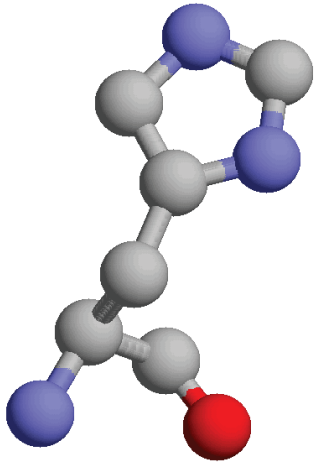
Methionine



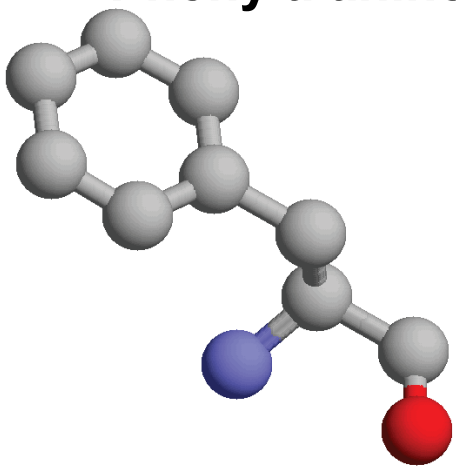
Alanine



Histidine

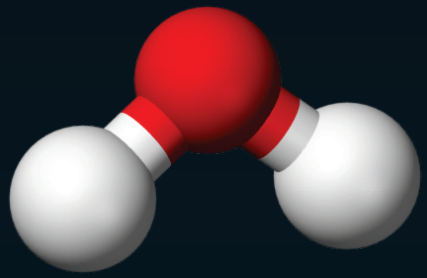


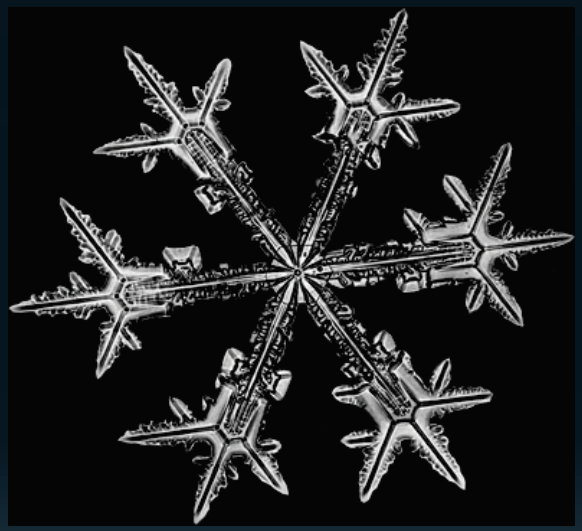
Phenylalanine

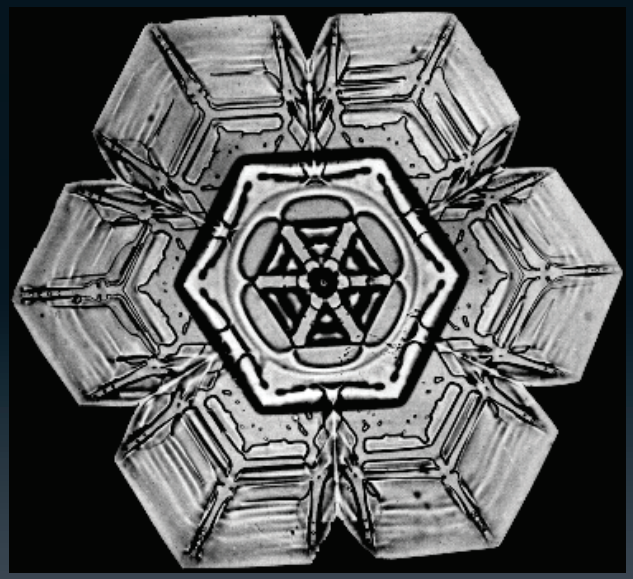


Local information yields global complex structures

- Snow flakes





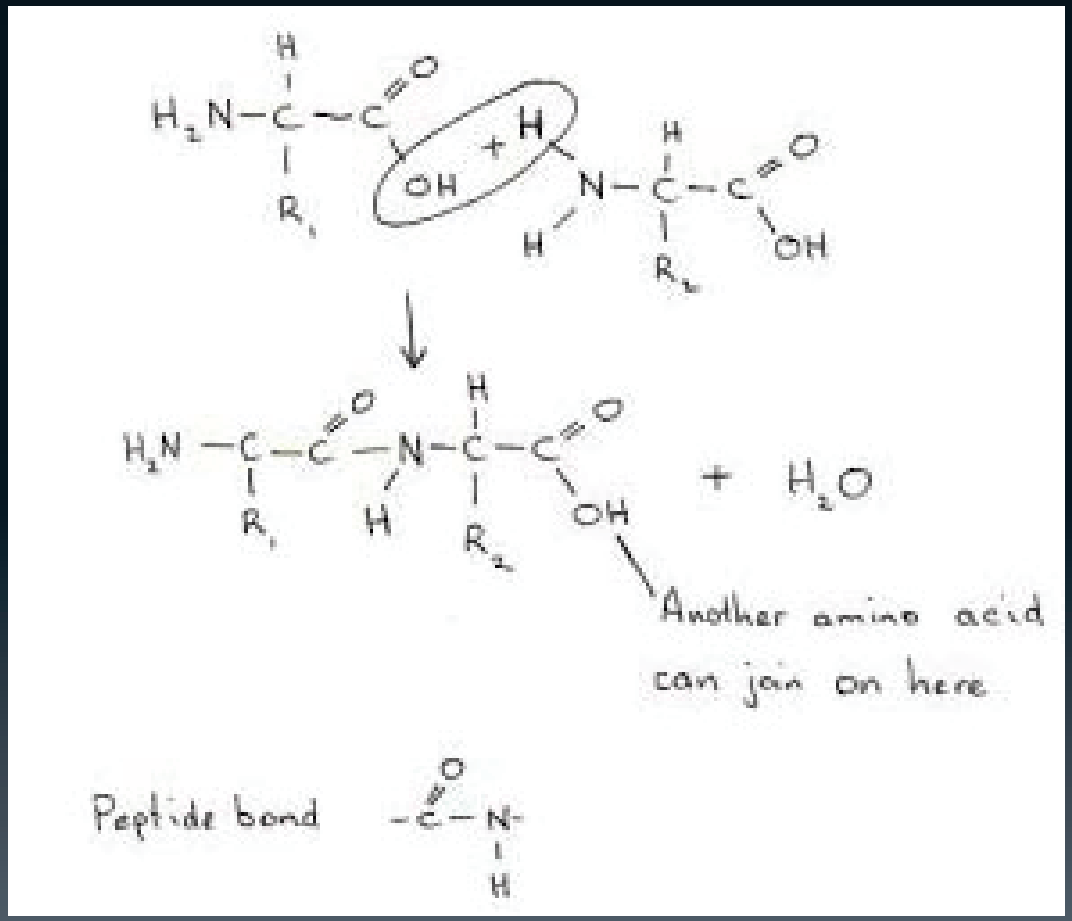


Wilson Bentley

http://www.its.caltech.edu/~atomic/snowcrystals/

Protein interactions and shape

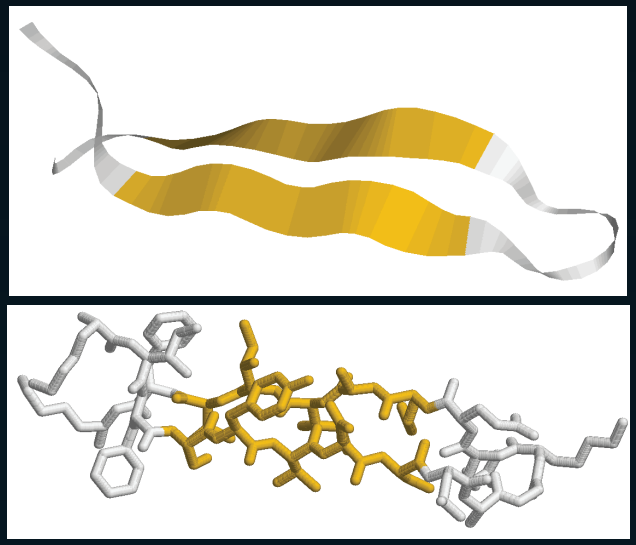
- **Strong Covalent interactions (110-50 kcal/mol): peptide bond :**

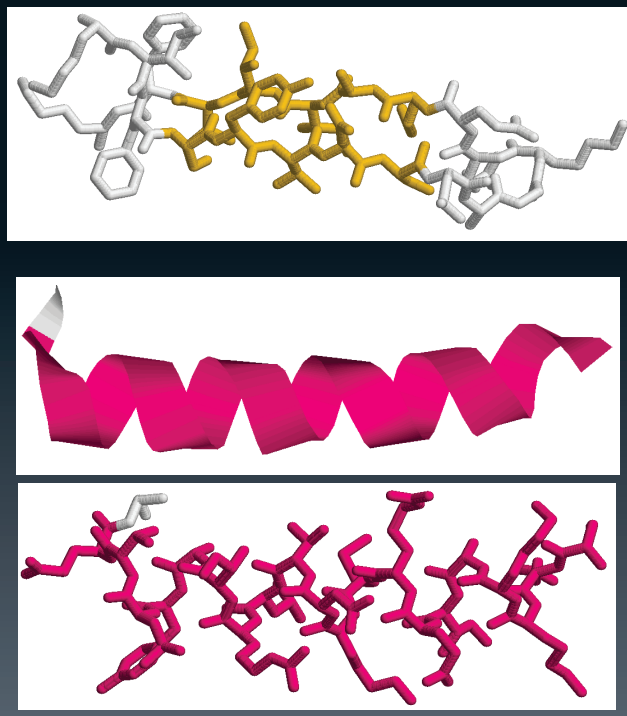


Sequence: ordered information

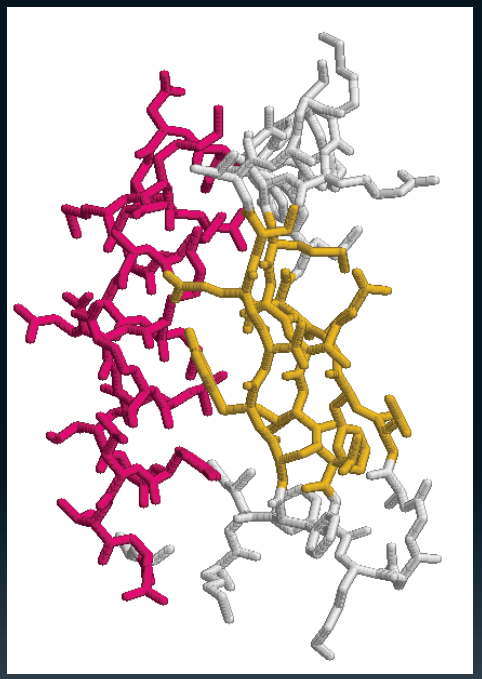
Weak interactions (1-7 kcal/mol):

SHORT RANGE: backbone atom interactions



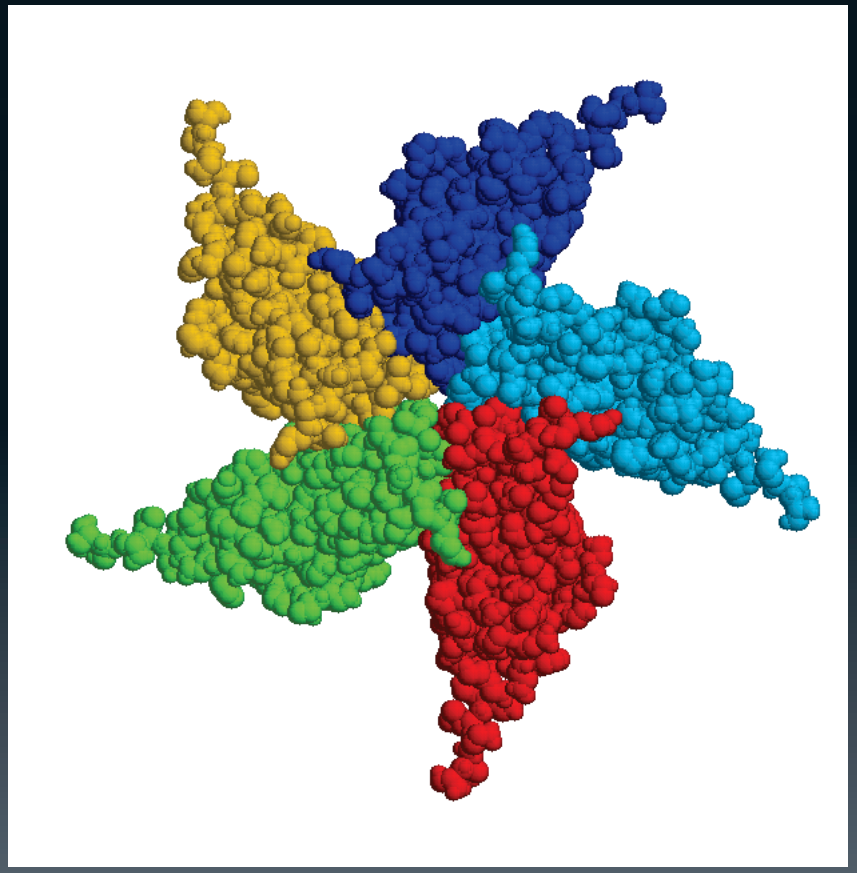


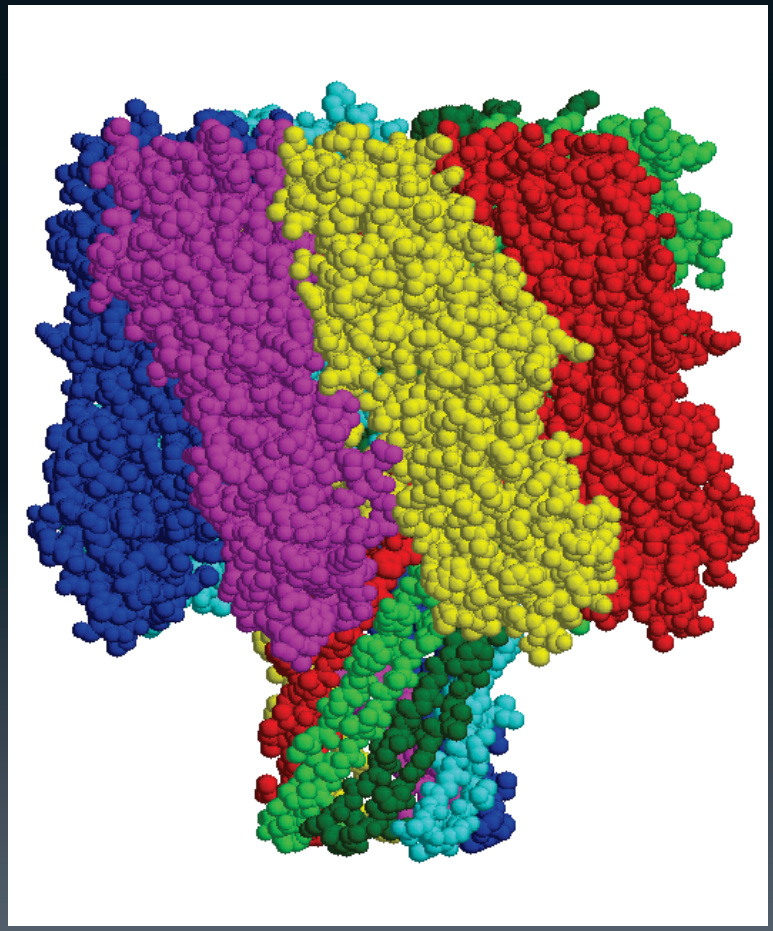
LONG RANGE: Side chain atom interactions



Protein shape plasticity

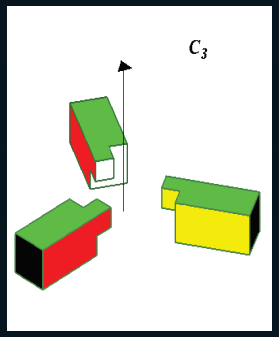
Different sequences

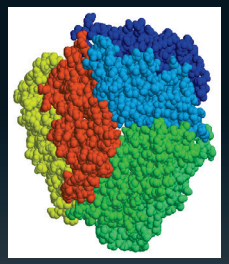




Different sequences

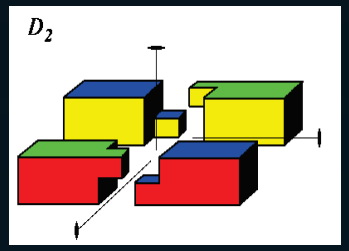
Cyclic

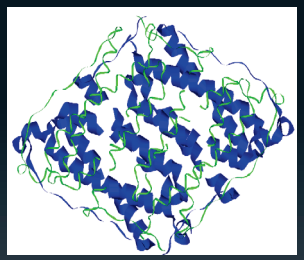




2F7N

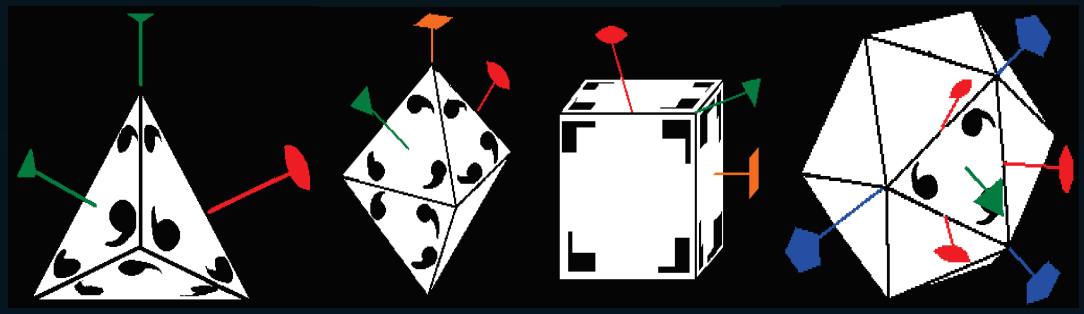
Dihedral



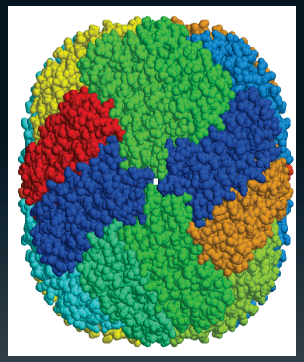


2RBD

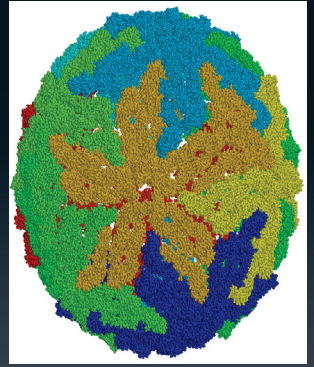
Cubic



1DPS

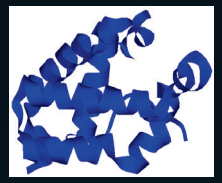


1LB3

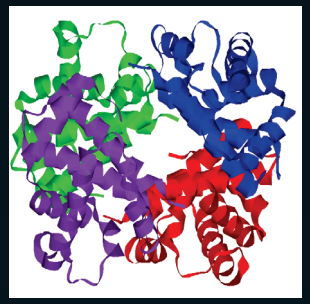


1K4R

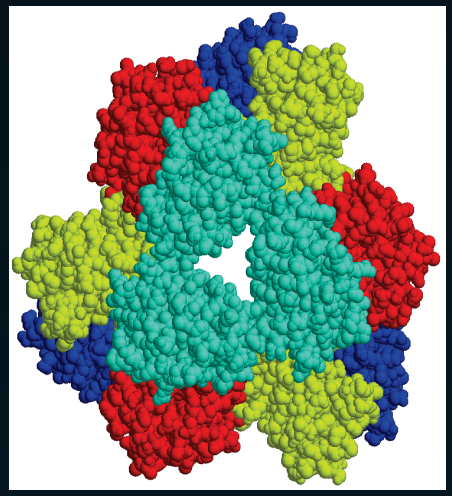
1S69	2HHB	2ZS1	2GTL	PDB
C1	C2	D3	D6	Symmetry



Hemoglobin from *Synechocystis cyanoglobin*

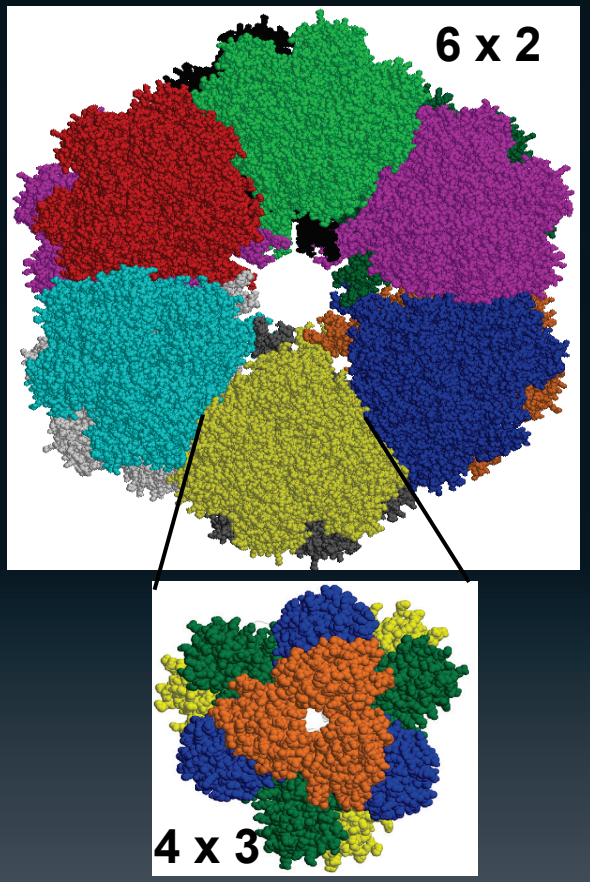


Hemoglobin from *Homo sapiens*



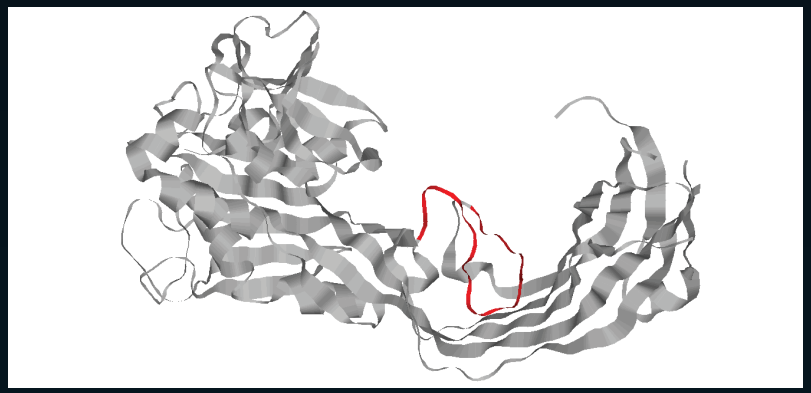
Hemoglobin from *Oligobrachia mashikoi*

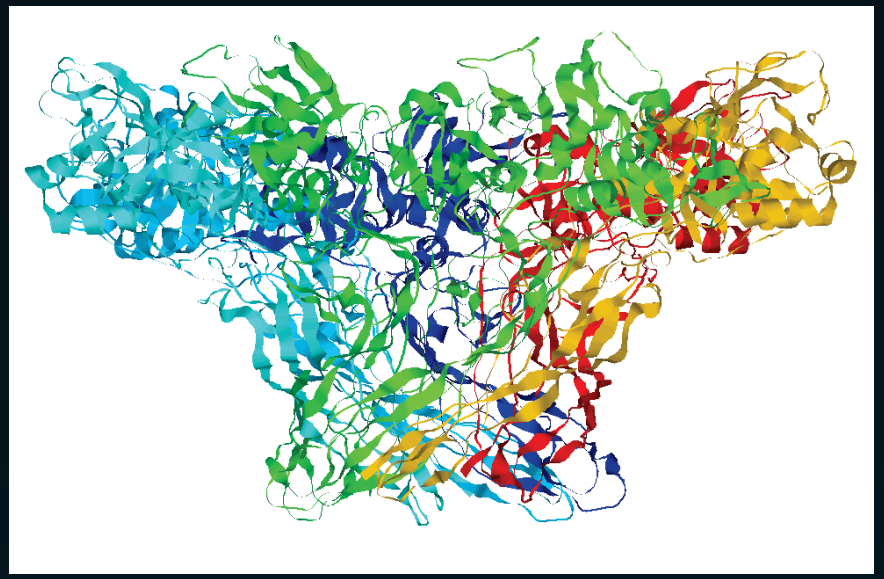
Same sequences

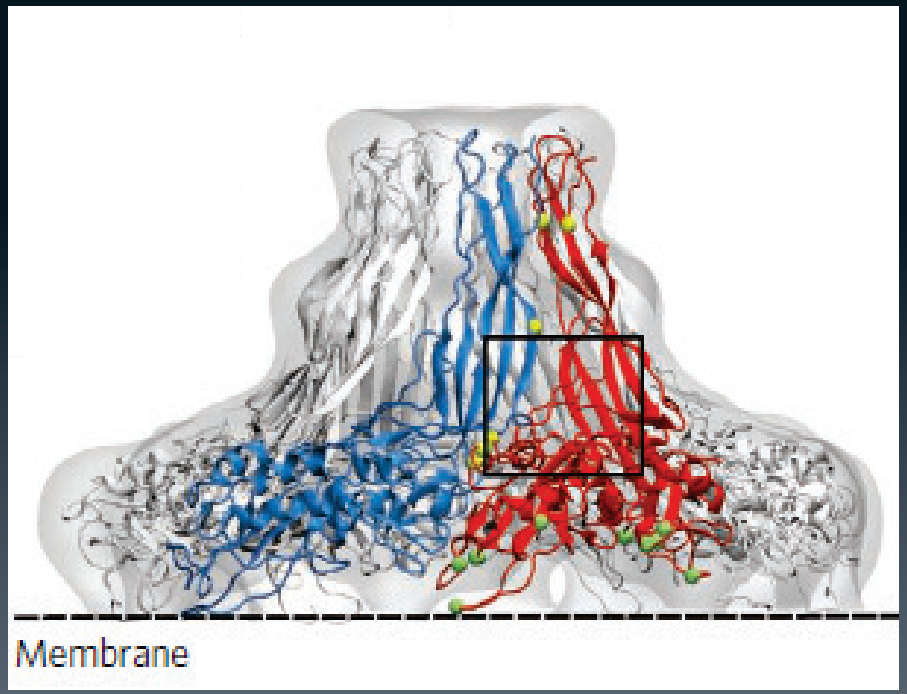


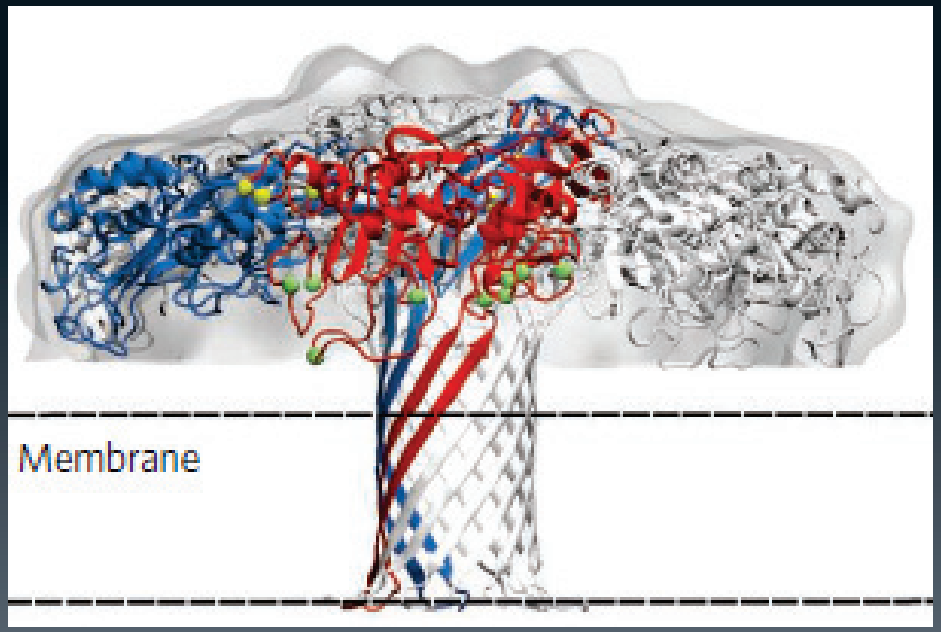
Giant earthworm hemoglobi
(144 chains: 4x3x6x2)

Plasticity: one sequence

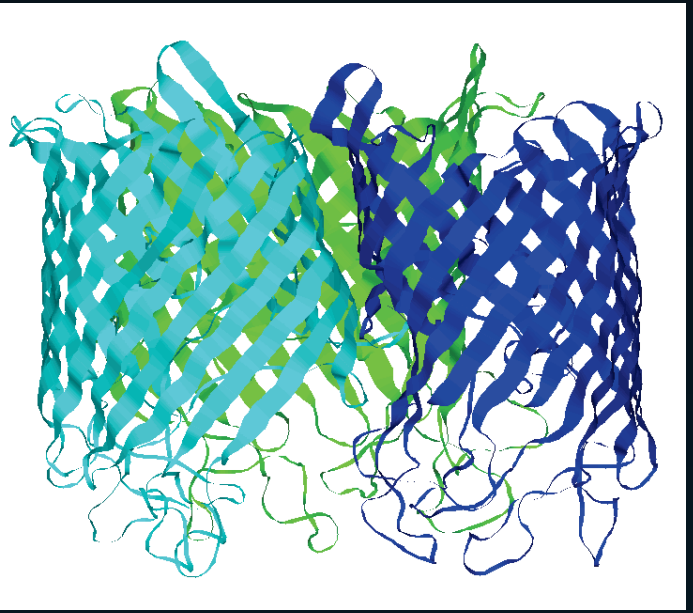


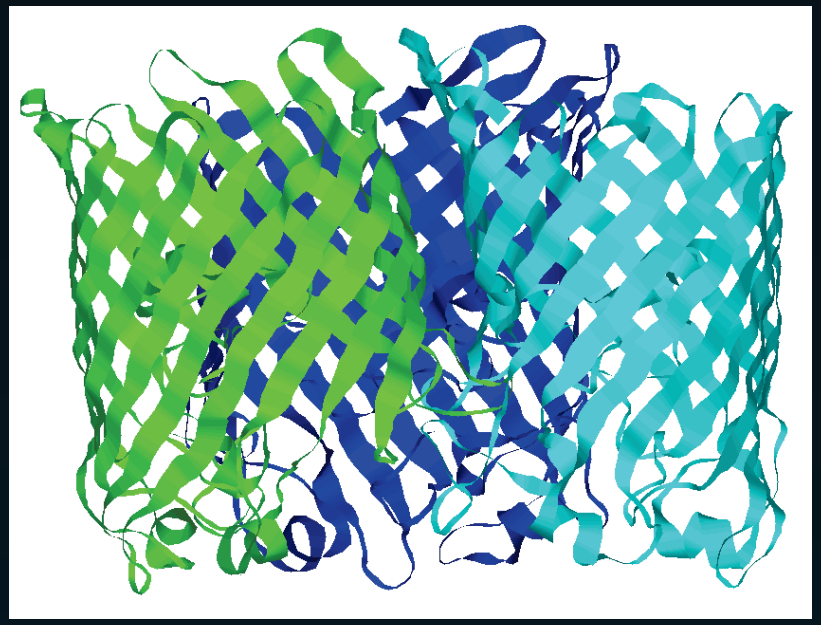




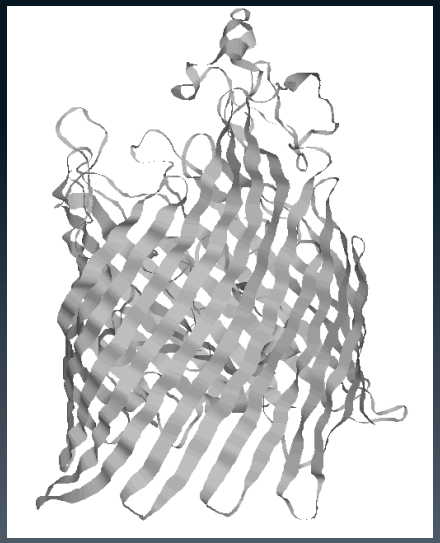


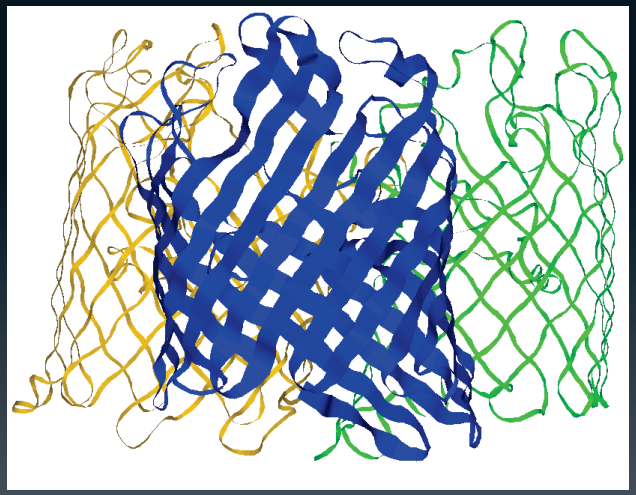
Robustness





Different sequences





Definitions of Robustness and Plasticity

- **Robustness: a change of amino acid (mutation) that does not affect the shape and/or the function of the protein**

- **Plasticity: a change of amino acid (mutation) that affects the shape and/or the function of the protein**

Single amino acid mutation —?→ **Global shape change**

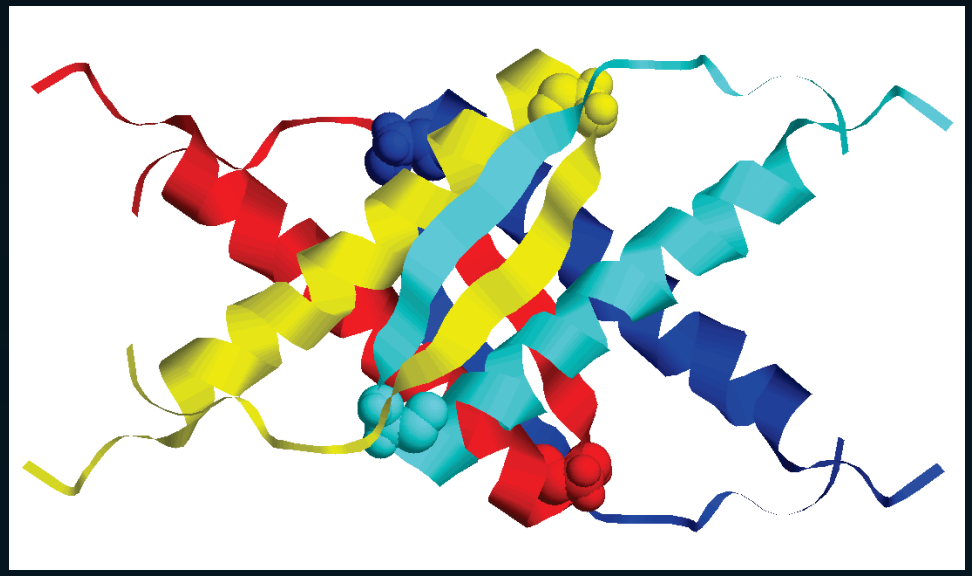
Why use network and graph theory ?

Local perturbation in the network: global effect: propagation

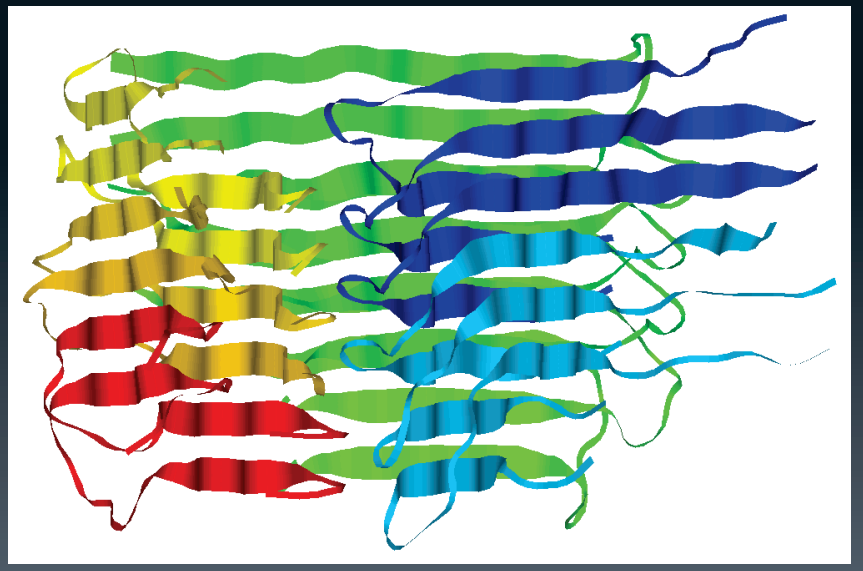
Do we have an equivalent in proteins?

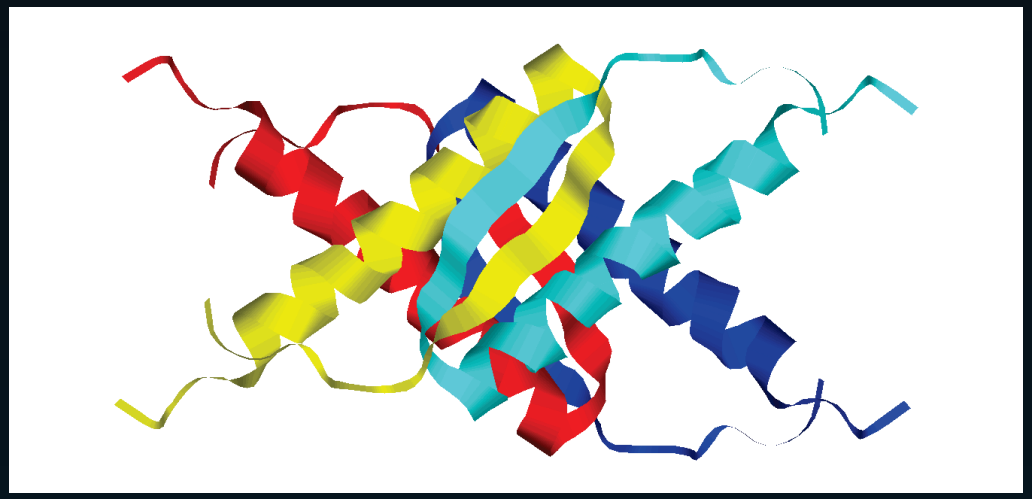
- Hypothesis: an amino acid robustness and/or plasticity is related to network properties

p53 cancer G334V

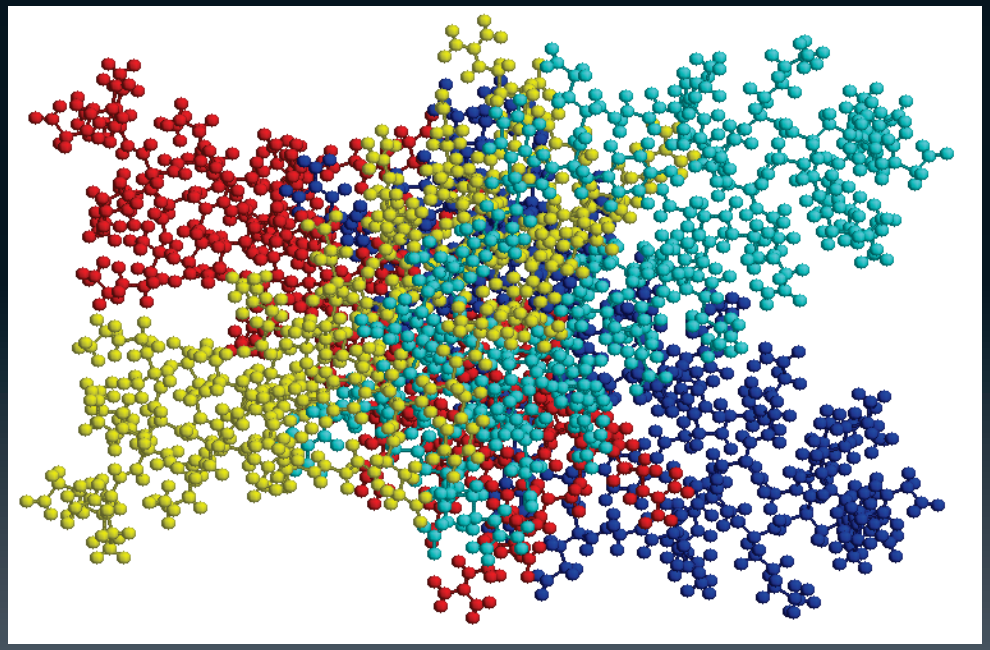


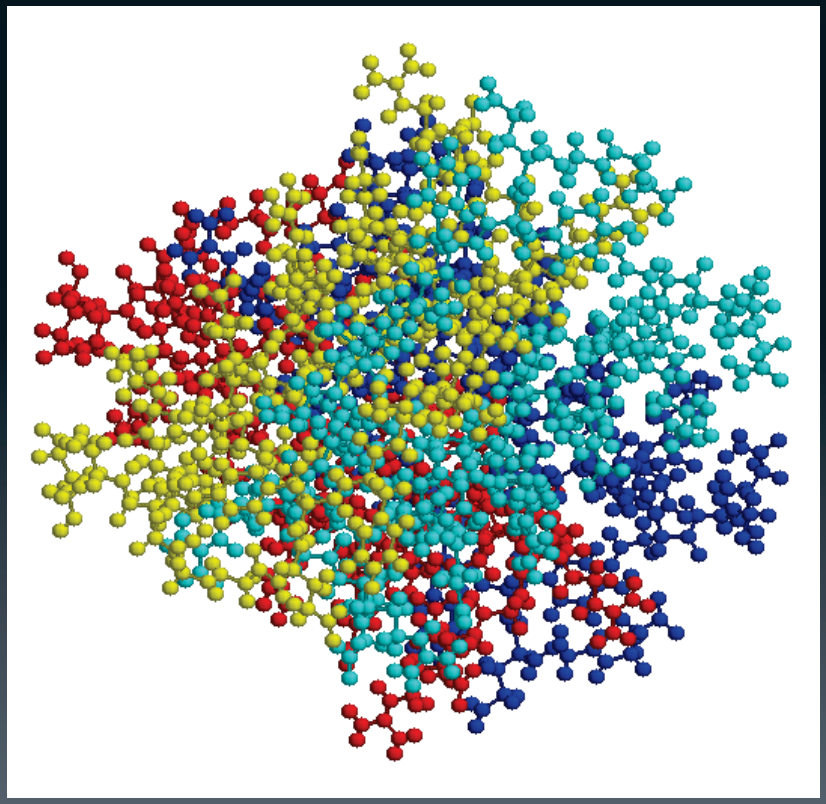
Single mutation



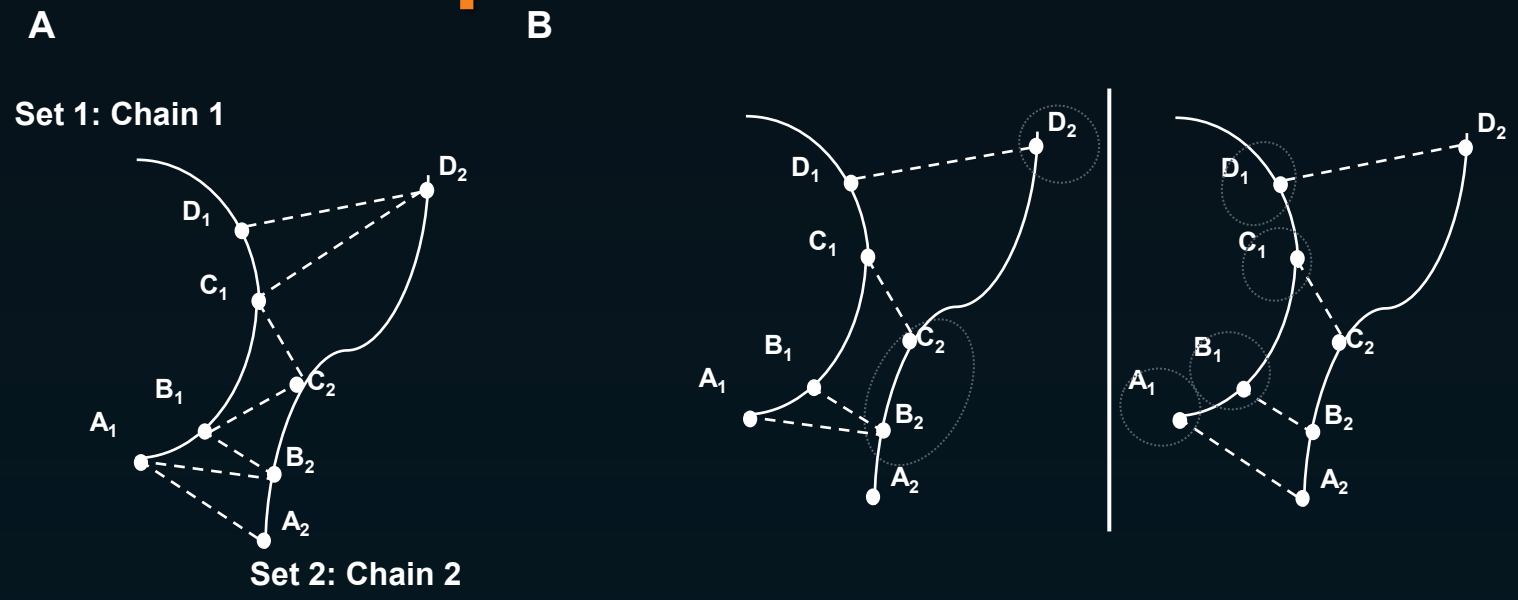


Is a protein a network?



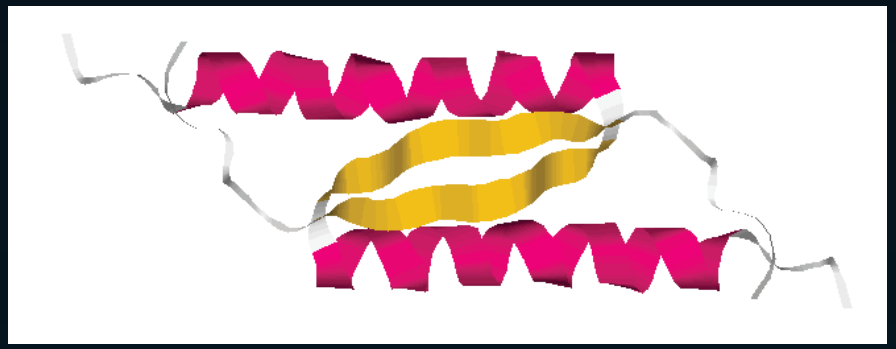


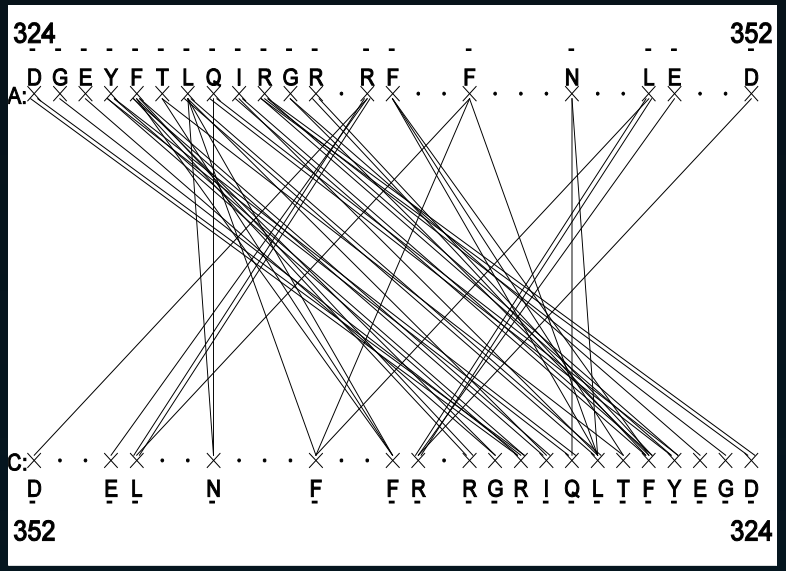
GEMINI: hot spots



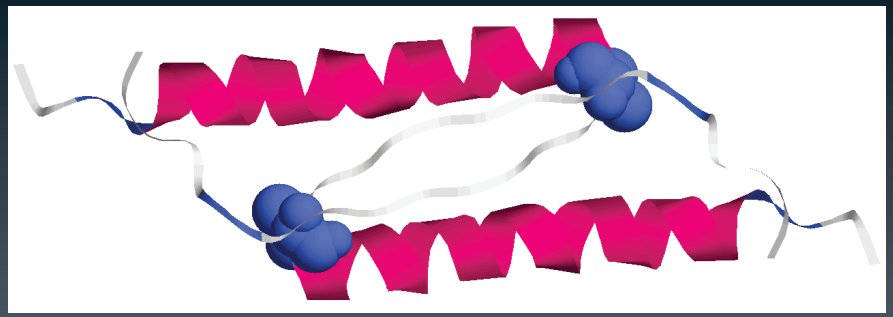
Proteins work like a network ?

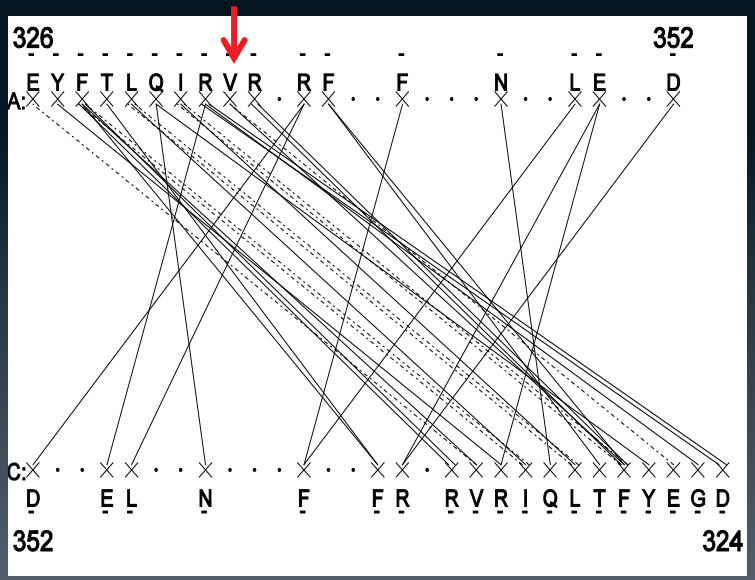
P53 WT





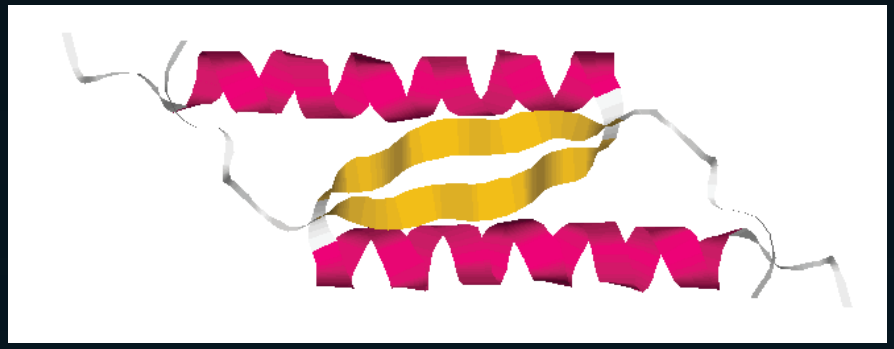
P53 G334V

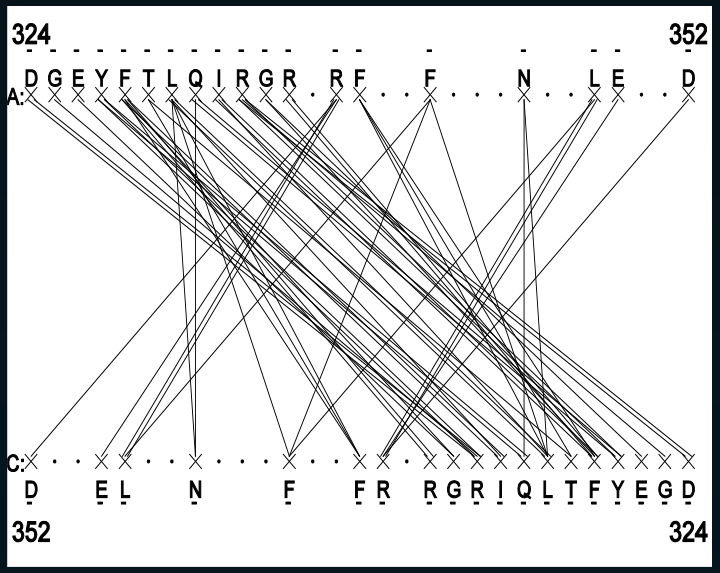




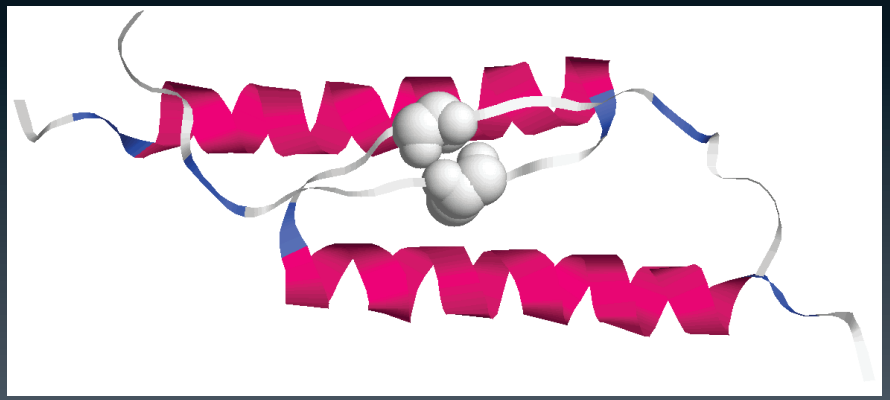
Interpretation ?

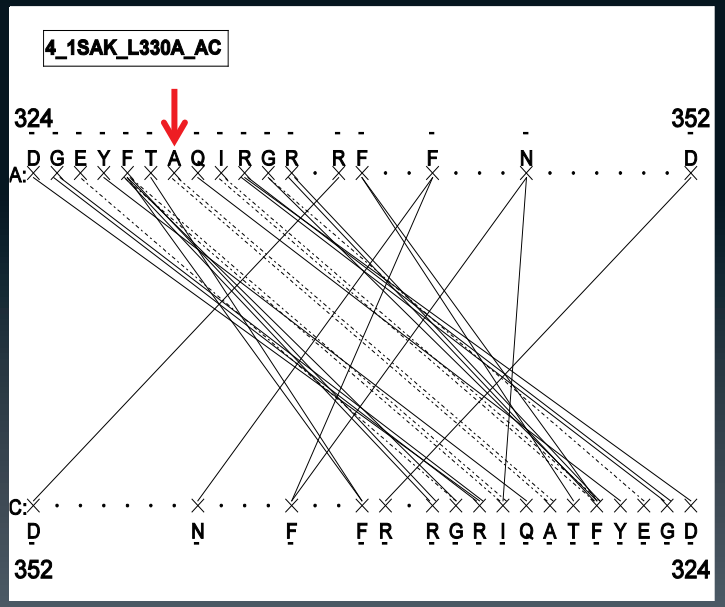
P53 WT





P53 L330A



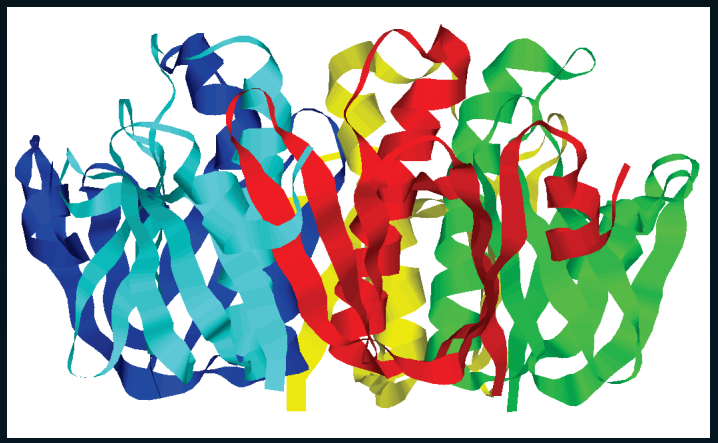


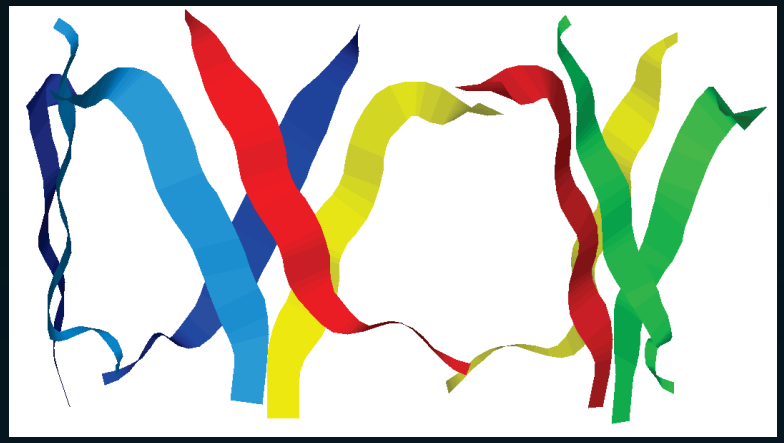
Network rewiring and node properties

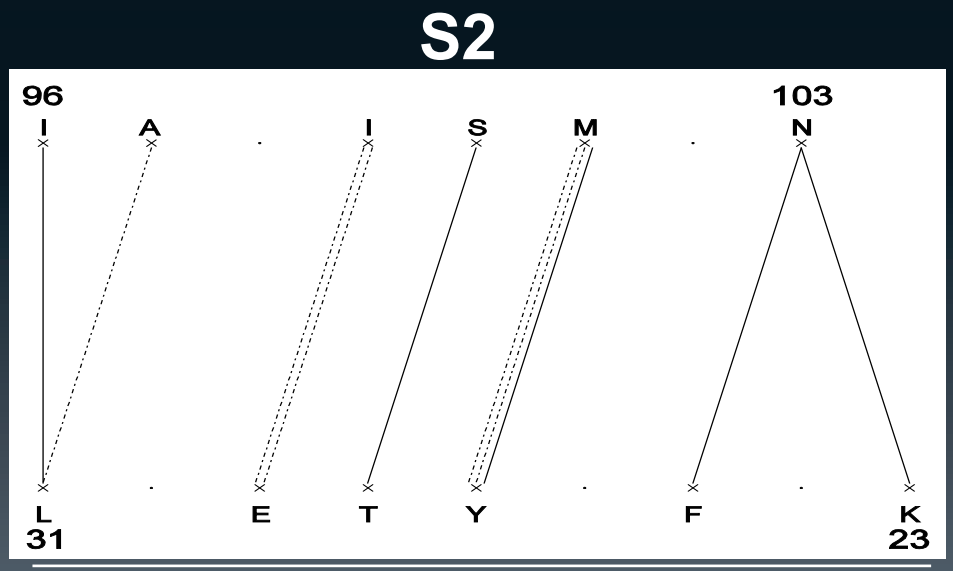
Chain dissociation and amino acid vulnerability

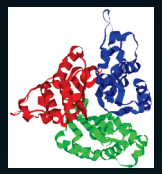
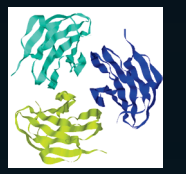
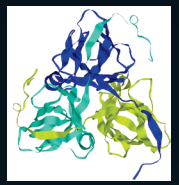
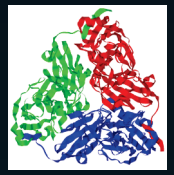
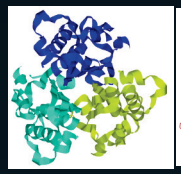
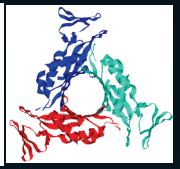
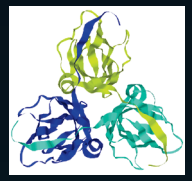
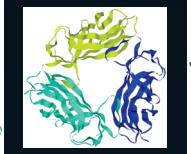
Healthy Interfaces : robust network?

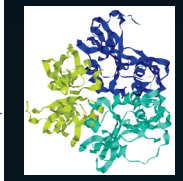
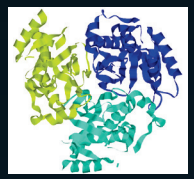
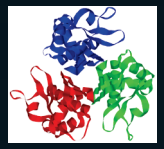
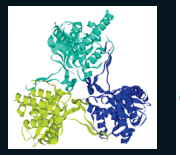
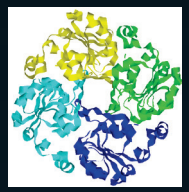
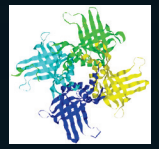
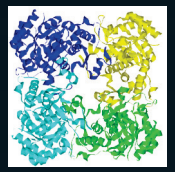
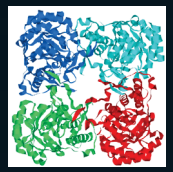
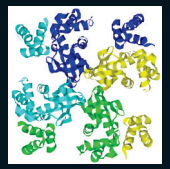
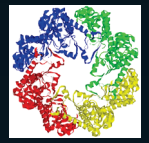
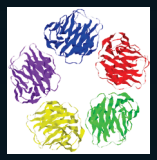
- 755 cases
- Intermolecular β-strand interactions

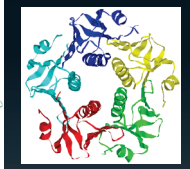
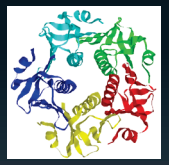
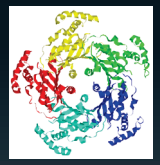
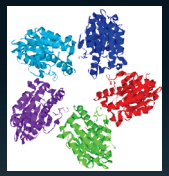
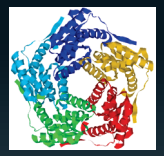
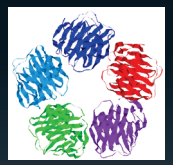
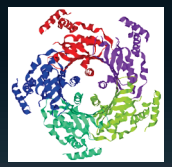
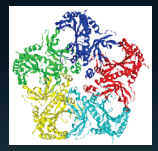


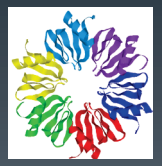
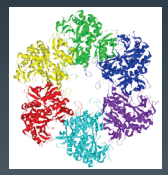
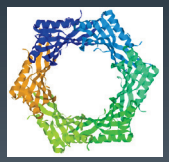
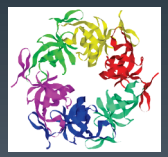
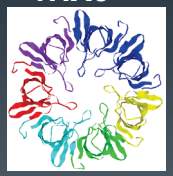
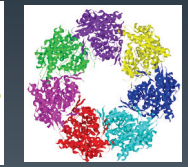


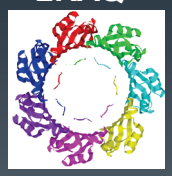
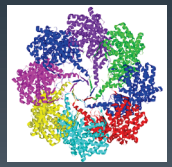



1JN1

1PM4

1SJN

1SNR

1TOA

1Y13

2BAZ

2BCM

2BT9

2GVH

2I9D

2JCA

2P90

1J8D

1L3A

1PVN

2A7R

2H5X

3BFO

1BO9

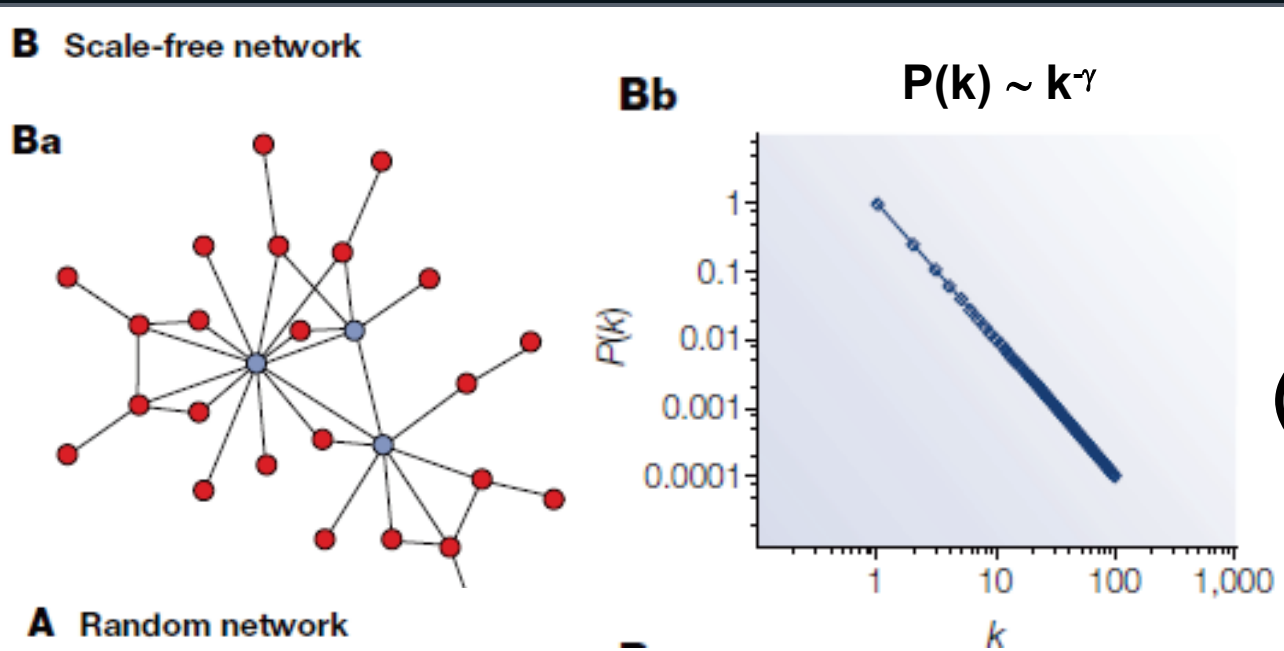
2XSC

1EEI

1EFI

1FB1

1HI9

1NQU

1SAC

1WUR

2OJW

2RCF

1U1S

2BVC

2GJV

2Z9H

1HX5

1OEL

1WNR

2RAQ

1Q3S

2V9U

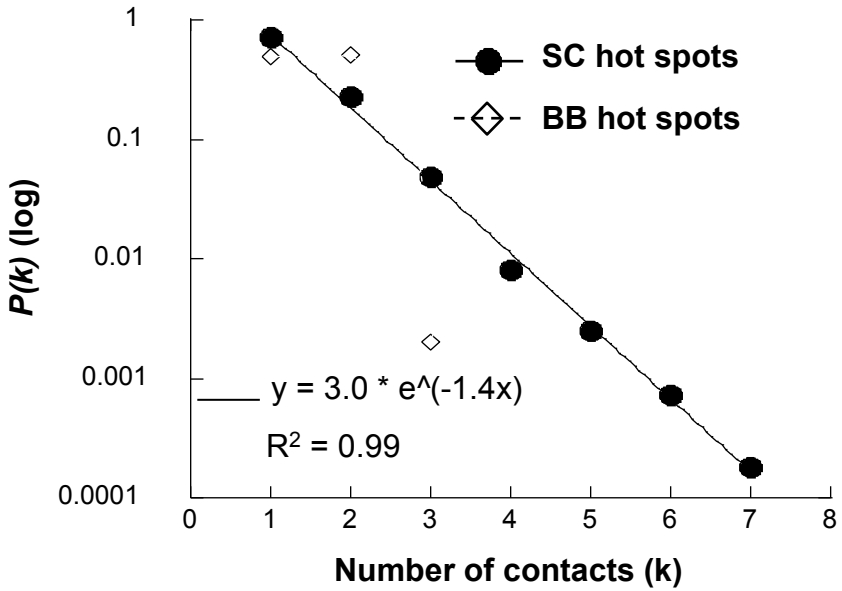
Network measures

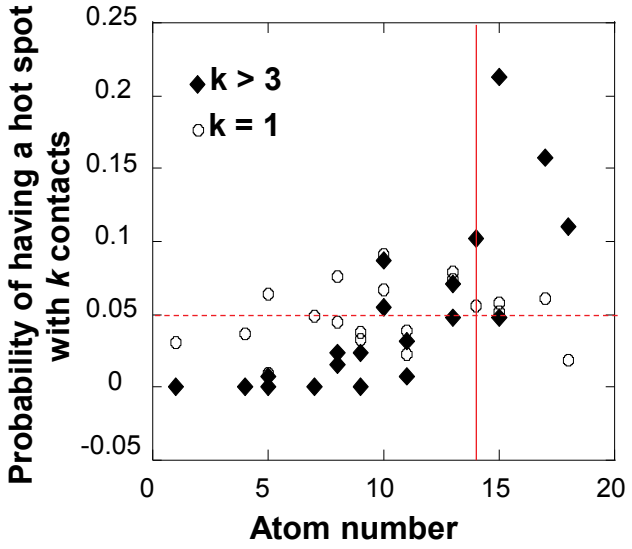


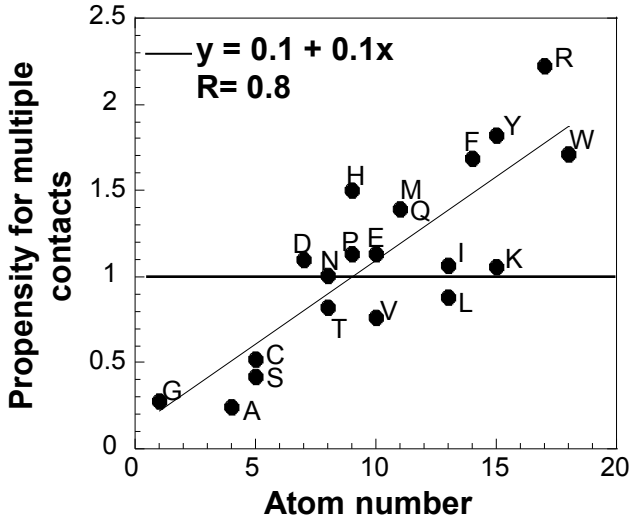
Diameter (path)

l ~ log N (small world)

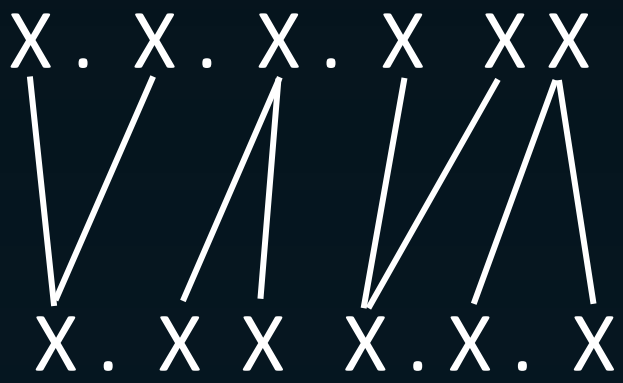
l ~ log N/log z

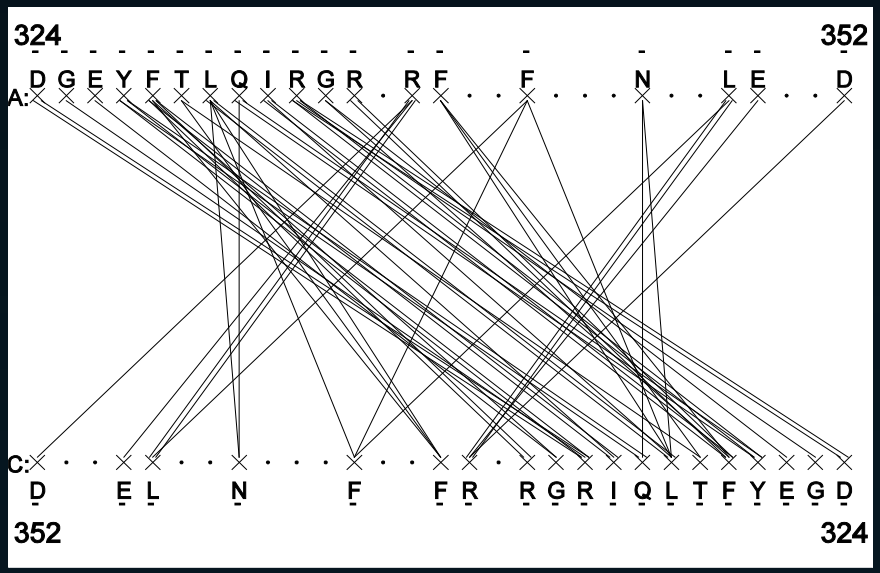

SC hot spots
BB hot spots
y = 3.0 * e^(-1.4x)
R² = 0.99
P(k) (log)
Number of contacts (k)


k > 3
k = 1
Probability of having a hot spot with k contacts
Atom number


y = 0.1 + 0.1x
R= 0.8
Propensity for multiple contacts
Atom number

Intermolecular β-strand networks are robust to mutation

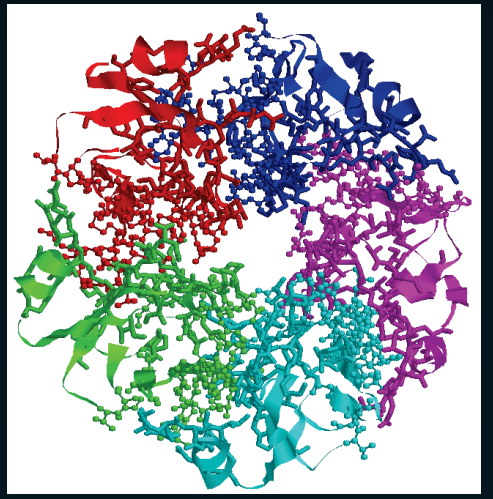


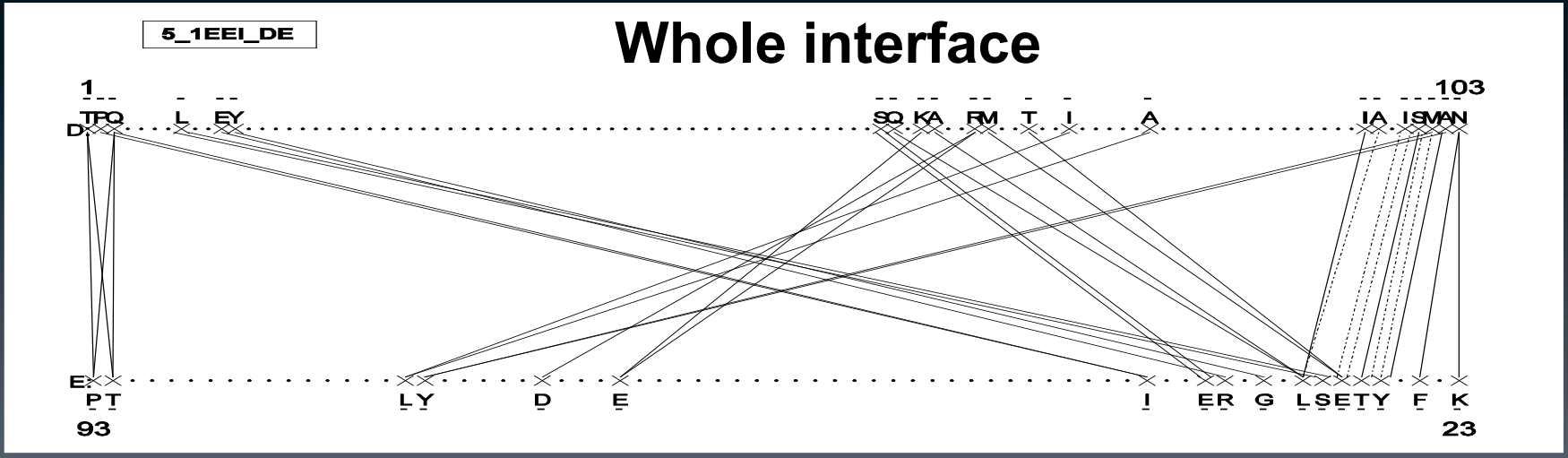


- Depleted in hubs
- No amino acid composition constraints
- Not costly in terms of links
- Disconnected networks
  - Average k is 1.4, average k for p53 is 3.0

Whole oligomeric Interface network

- SpectralPro: , whole oligomer, whole chain, 283/ 61 hot spots
  - 10 Closest neighbors, atomic distances
  - Probability of interactions (atom number)
  - Weightless: amino acid contacts
  - Weighted: number of atoms: links
  - Clustering

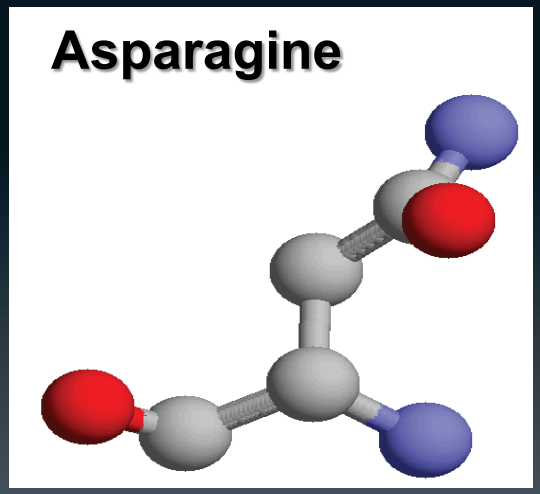




Stability/dissociation of interface

- Fold X: stability of interchain association
  - WT vs hot spot single mutation to asparagine
  - $\Delta G_{WT} - \Delta G_{mut}$

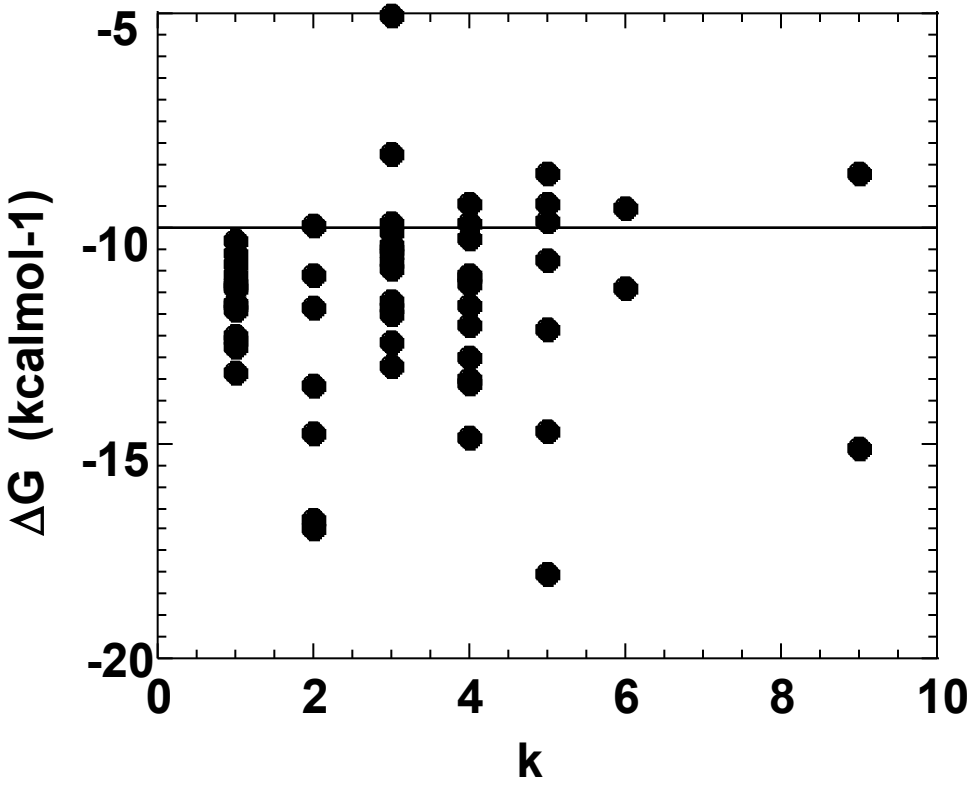
Network measures

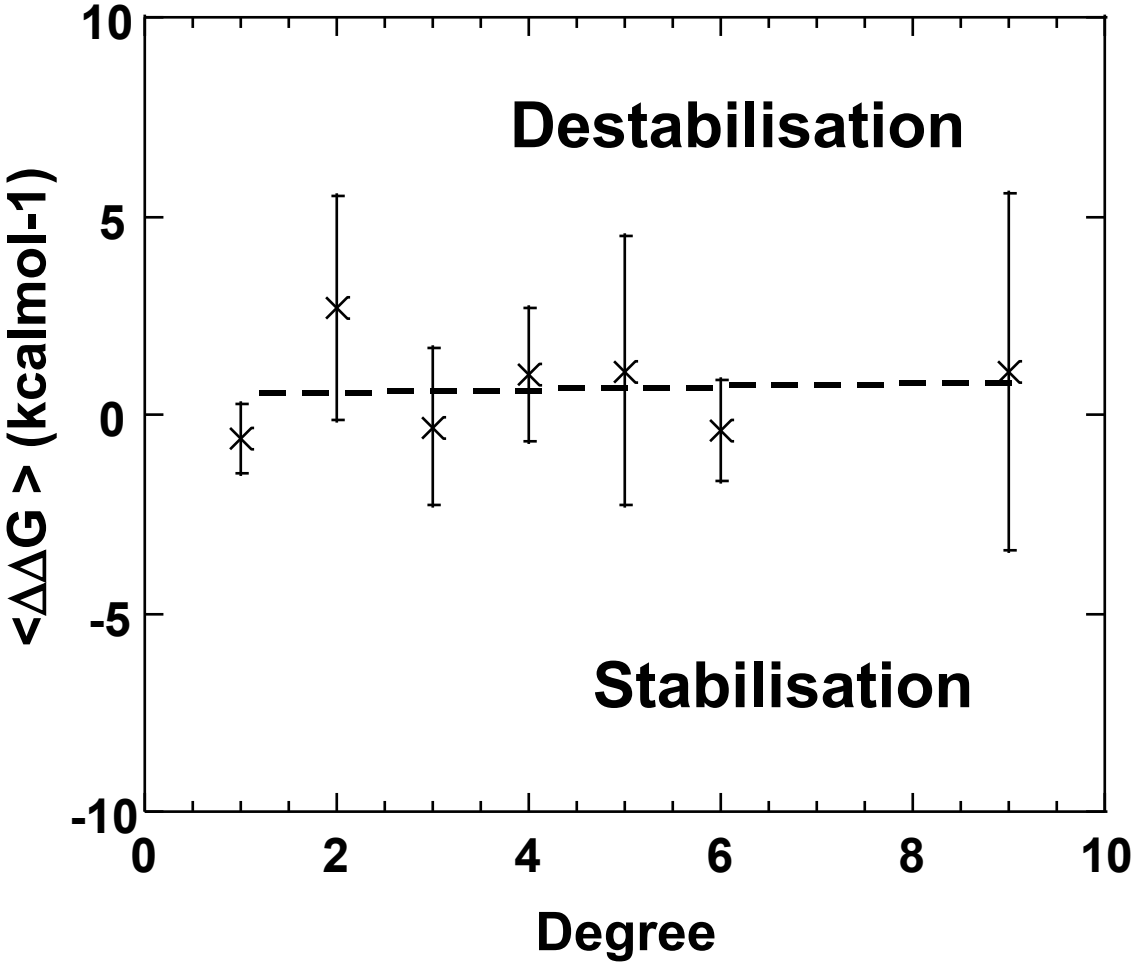


- Hot spot mutation affects free energy of interface
  - Hot spots/nodes plastic to mutation

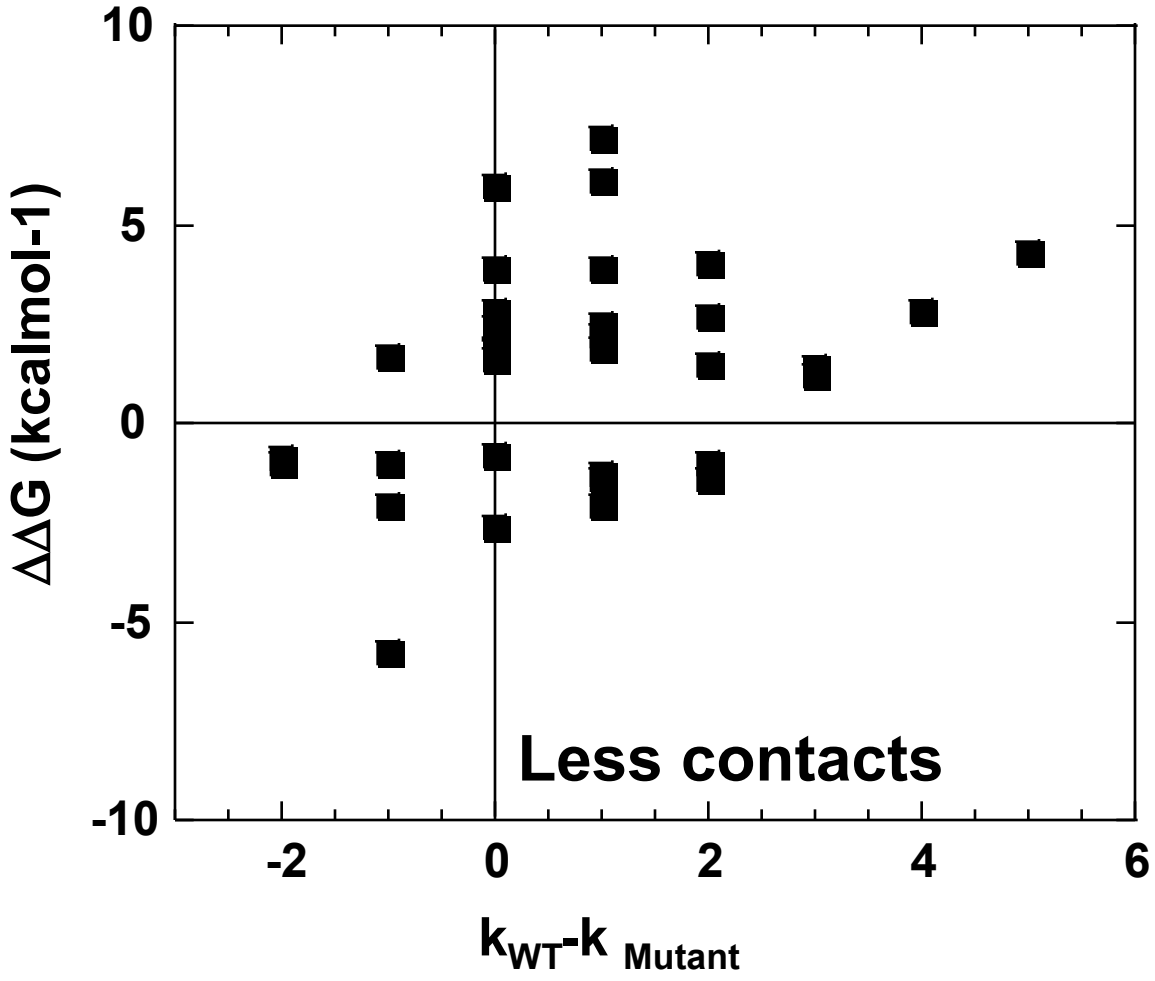
- Hot spot mutation does not affect the free energy of interface
  - Hot spots/nodes robust to mutation

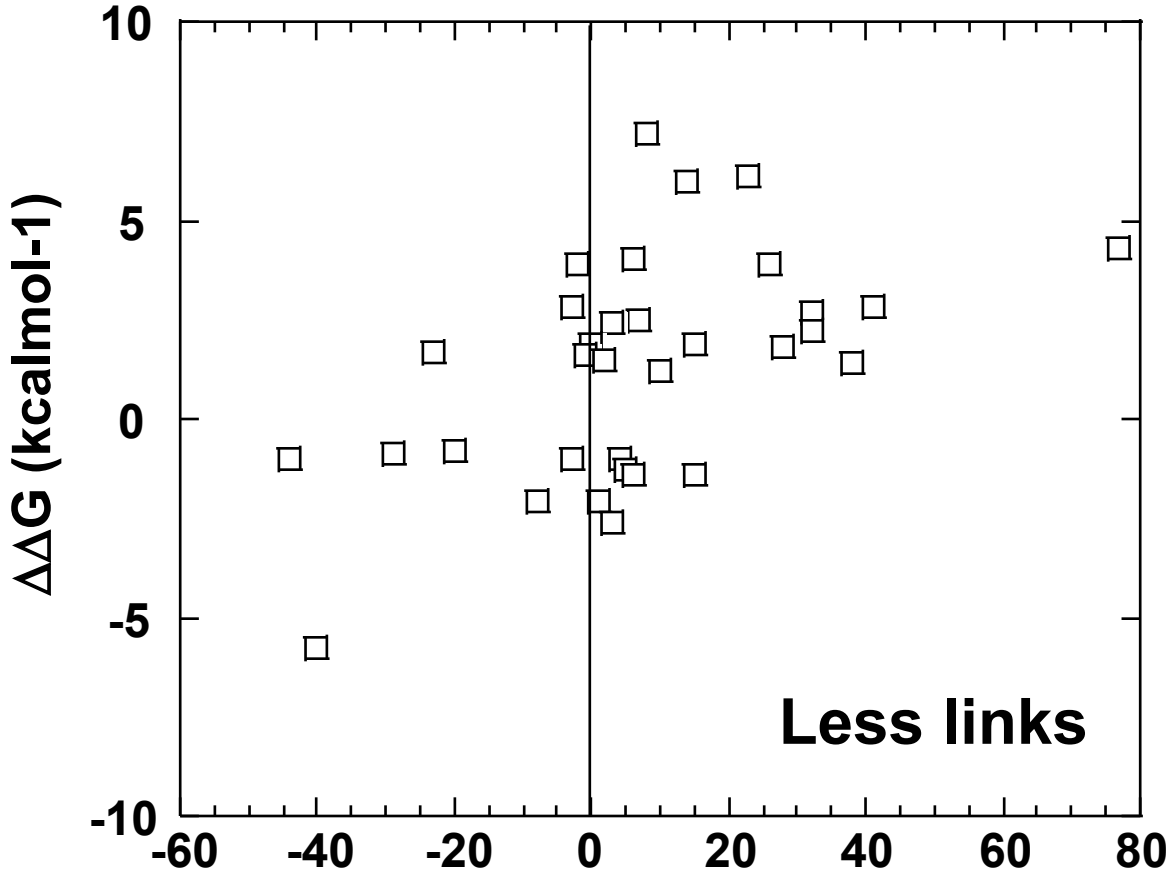
Network measures: the degree of the nodes



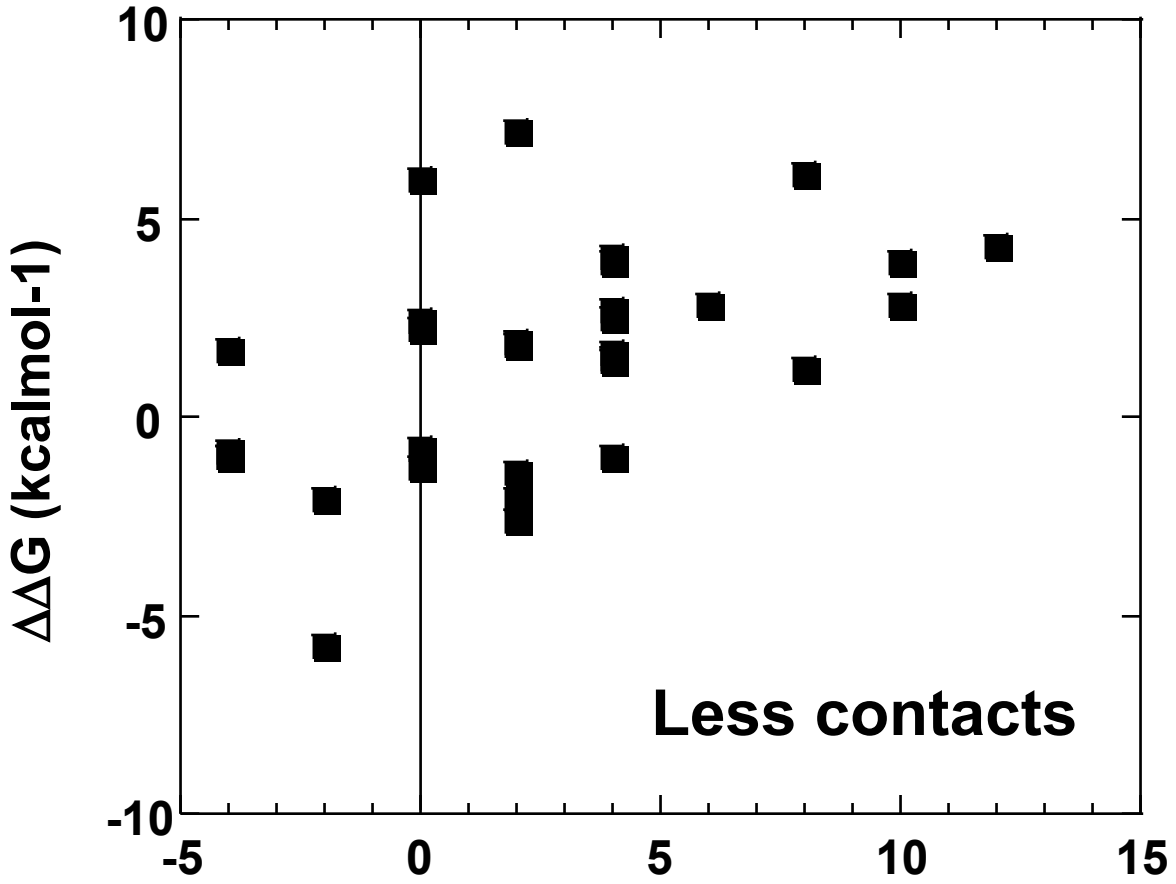

Destabilisation
Stabilisation
<ΔΔG > (kcalmol-1)
Degree
10
5
0
-5
-10
0
2
4
6
8
10

Local degree perturbation upon mutation

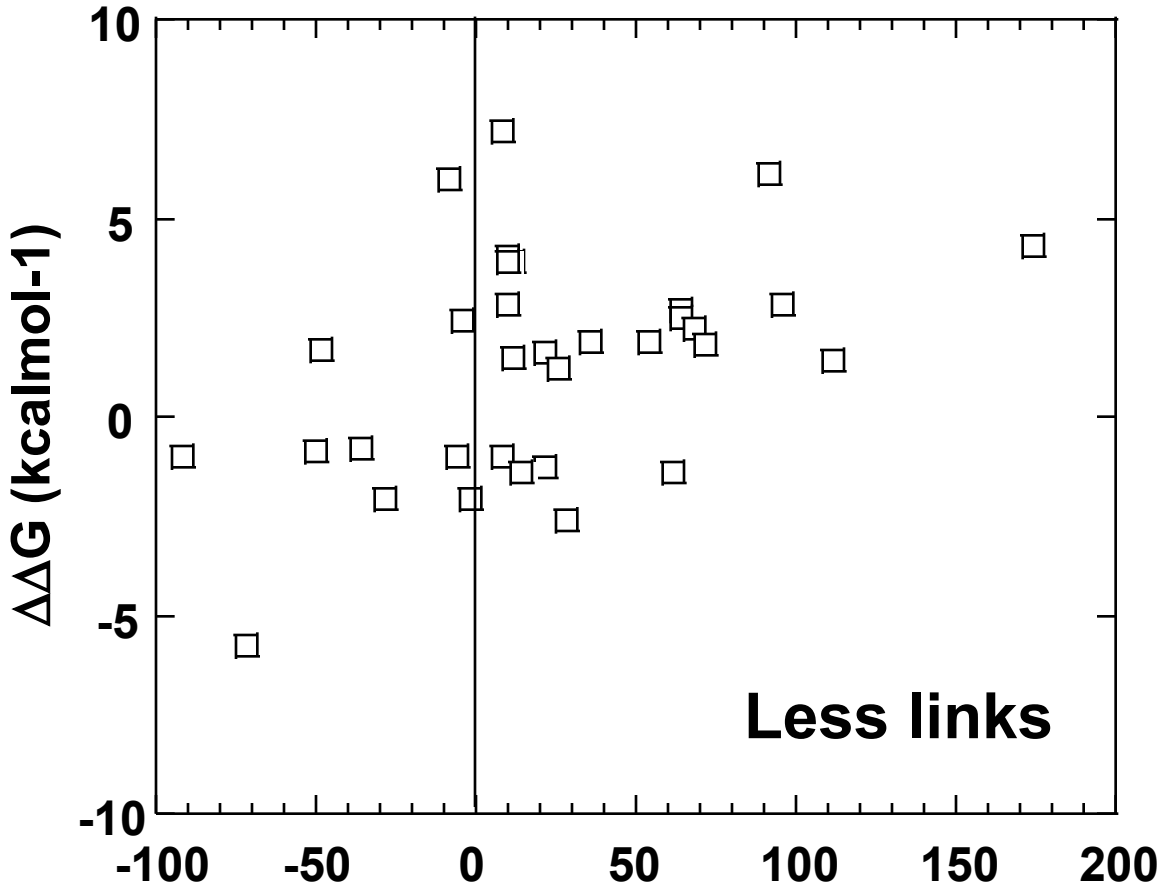



10
5
0
-5
-10
-60
-40
-20
0
20
40
60
80
ΔΔG (kcalmol-1)
Less links

Global degree perturbation upon mutation

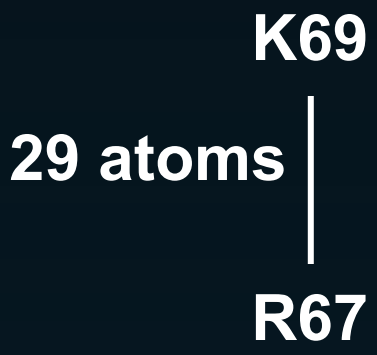


Global degree perturbation upon mutation

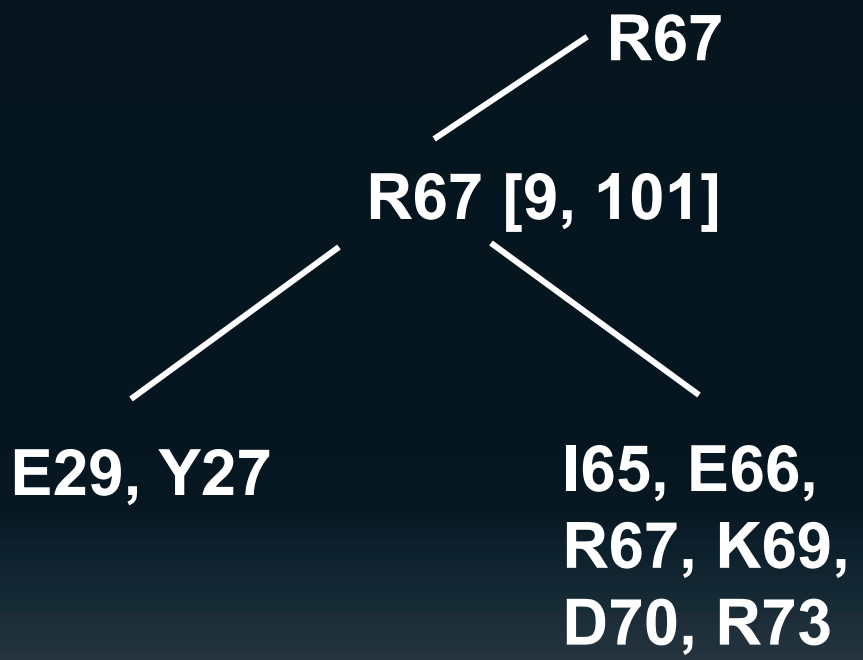


Zoom in on the nodes R67 and K69

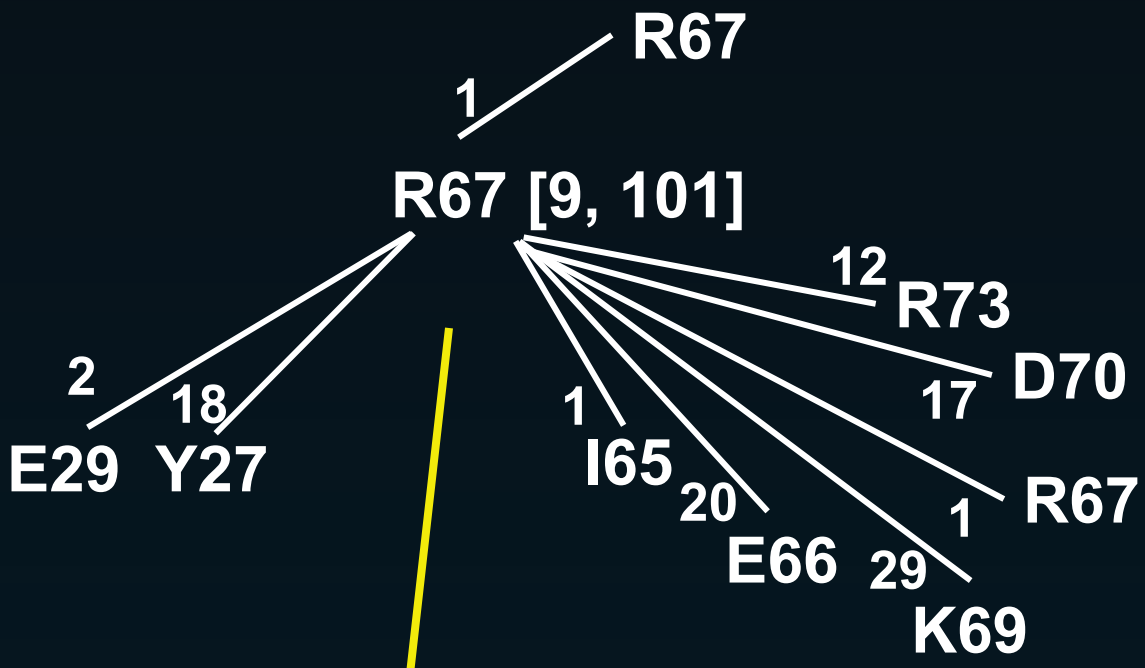
K69N: max stabilisation

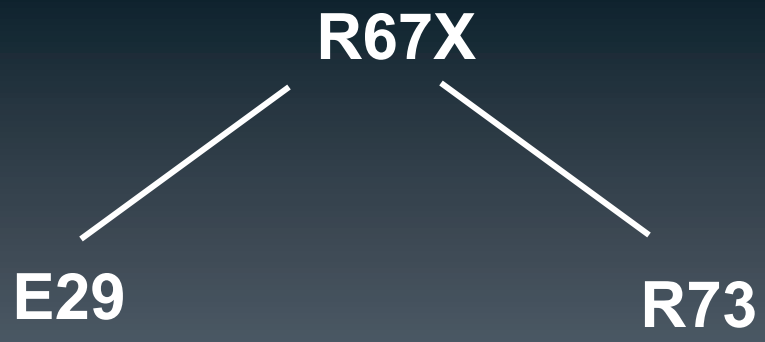


R67N: 40 % stabilisation



Paradox: node with sequence constraint more robust to mutation than node flexible in sequence


R67
1
R67 [9, 101]
12
R73
D70
17
R67
1
29
K69
E66
20
I65
1
2
E29
18
Y27


R67X
E29
R73

Conclusion

- Too many contacts/links : risk
- Robustness by redundancy (many members of the same communities
- Robustness by duplicate (not shown)
- Risk: single contacts all the asset (bank)

- Correlation degree $P_k.P_{k'} = P_{kk}$

- High degree-low degree pairs: propagation

The degree of the nodes is not enough information to infer the robustness or the plasticity towards mutation