Statistical Issues in Protein Folding

Ingo Ruczinski

All of this stuff is in collaboration with Kevin Plaxco, UC Santa Barbara.

Protein Folding and Structure Prediction

There is a difference between protein folding and structure prediction:

- Protein folding is concerned with the process of the protein taking its three dimensional structure.
- Protein structure prediction is solely concerned with the 3D structure of the protein.

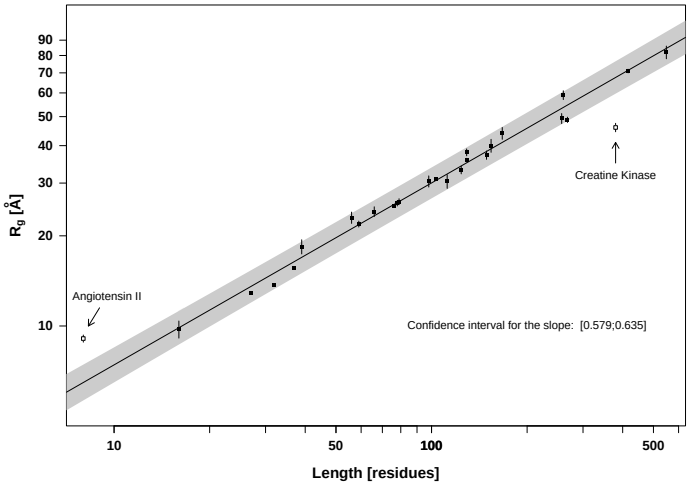
This poster explains some topics related to protein folding. The issues are:

- Do chemically denatured proteins behave as random coils?
- Are there site-specific deviations from random coil dimensions?
- What determines the folding rate of a protein?
- Are amino acids in proteins conserved because of folding kinetics?

Radius of Gyration of Denatured Proteins

Do chemically denatured proteins behave as random coils?

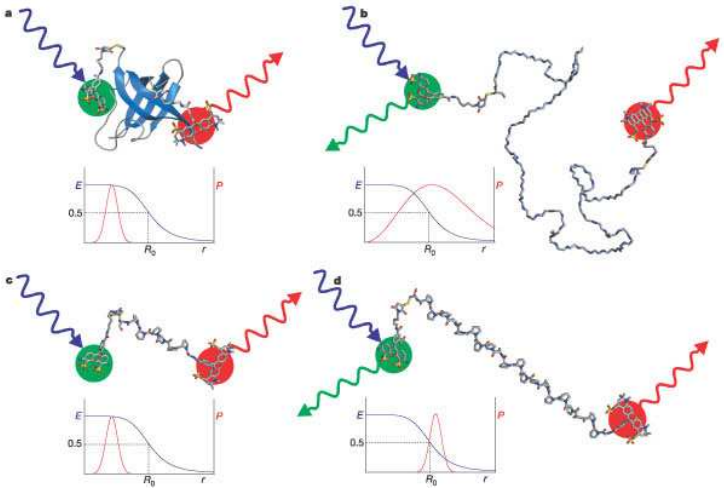
The radius of gyration R_g of a protein is defined as the root mean square distance from each atom of the protein to their centroid. For an excluded volume polymer (a polymer with non-zero thickness and non-trivial interactions between monomers) in a solvent, the average radius of gyration of a random coil is a simple function of its length N: $R_g \propto N^{0.588}$.



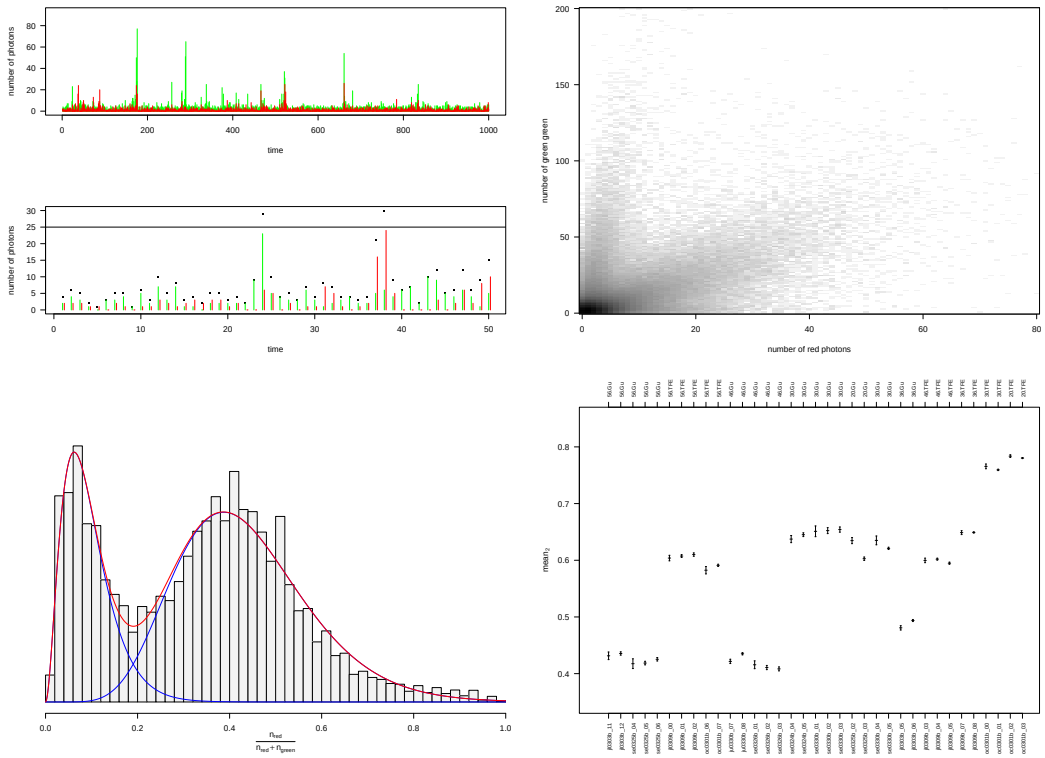
Deviations from Random Coil Behaviour

Are there site-specific deviations from random coil dimensions?

Förster Resonance Energy Transfer enables us to measure the distance between two dye molecules within a certain range. This can be used to study site-specific deviations from random coil dimensions in highly denatured peptides. Since this is quite complicated stuff, I only put up some relevant figures and will be happy to explain them to you if you are interested.



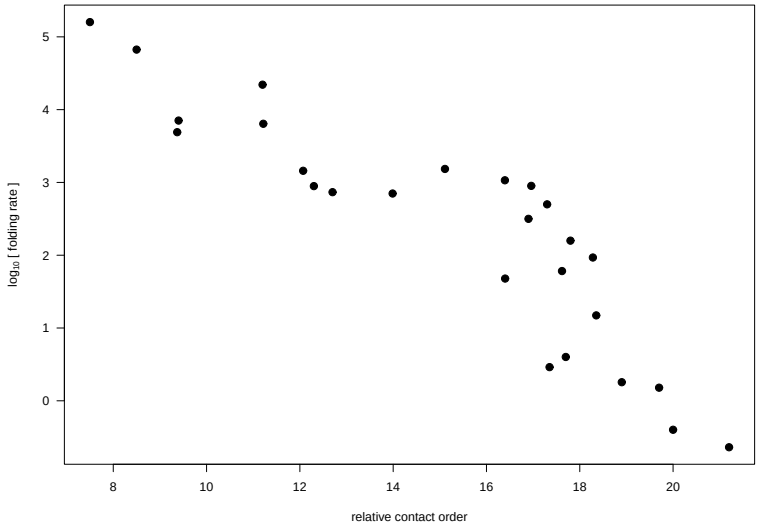
Deviations from Random Coil Behaviour



Protein Folding Rates

What determines the folding rate of a protein?

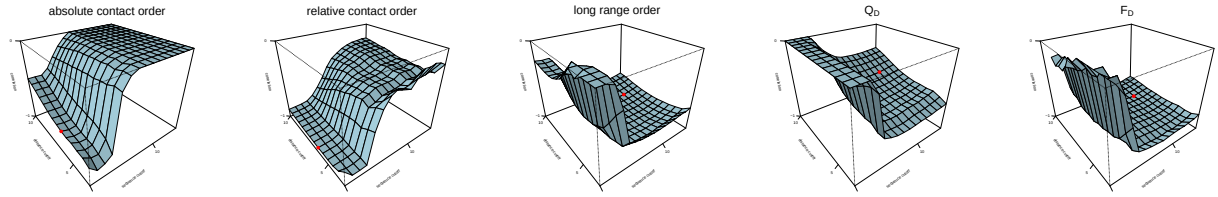
The folding rate of a protein heavily depends on the number of non-local interactions between amino acids. The relative contact order of a protein is defined as the average sequence separation between amino acids in contact, divided by the length of the protein. Contact is defined as two heavy atoms within 6Å.



Protein Folding Rates

What about all those other measures for non-local interactions?

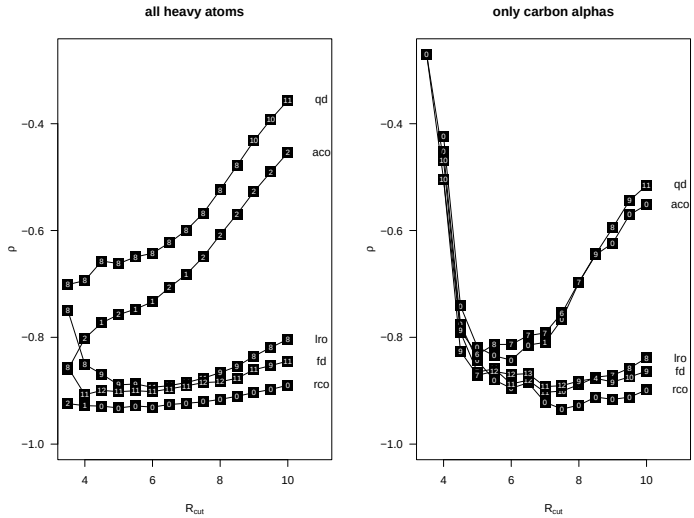
Since the finding that the folding rate of a protein depends on the number of non-local interactions between amino acids, many other measure similar to relative contact order have been suggested (and most of them are actually only slight “perturbations” of the formula for relative contact order). All of those measures depend on two thresholds: the distance cutoff to define contact and the sequence cutoff to exclude sequence neighbours from the calculations.



Protein Folding Rates

Which measure to use?

Are some of those measures better in a certain sense than others? Which ones perform equally well? All of those measures are pretty much ad-hoc. Maybe we can gain some insight in the actual folding process by comparing those measures - which aspects of the folding process do they capture? But mainly this comparison is to settle arguments along the line *my measure is better than yours!”*



Evolution and Folding Kinetics

Are amino acids in proteins conserved because of folding kinetics?

To what extent does natural selection act to optimize the details of protein folding kinetics? Is there a relationship between an amino acid's evolutionary conservation and its role in protein folding kinetics?

Some comments:

- Our studies of sequence conservation among residues known to participate in the folding nuclei of all of the appropriately characterized proteins reported to date have not provided any evidence that highly conserved residues are more likely to participate in the protein folding nucleus than poorly conserved residues.
- This is in contrasts to some of the beliefs stemming from theoretical considerations (good science, good people).
- This is also in contrast to some of the reported findings in the literature (no good science here, and really crappy statistics).
- The latter people do not like us.