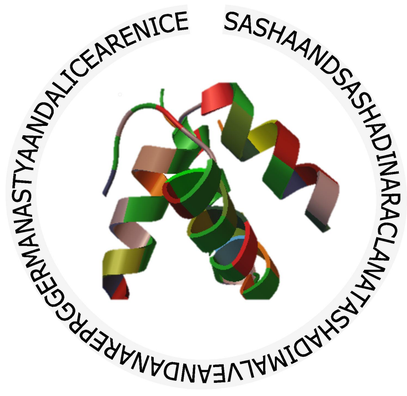
The curious case of protein mutants

Joan Ariño[1], Anastasia Gorshkova[2], Aleksandra Eremina[3], German Kanevskiy[4], Liubov Lonishin[5], Aleksandr Meister[6], Alisa Yakupova[7], Dinara Usmanova[8,9], Natalya Bogatyreva[8,9], Dmitry Ivankov[8,9].

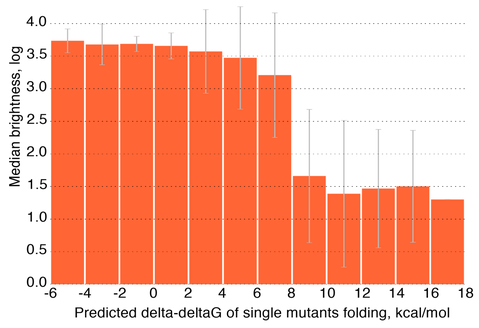
[1]Universitat Politècnica de Catalunya, Barcelona; [2]MSU, Moscow; [3]University of Edinburgh; [4]HTC MUCTR, Moscow; [5]PTSH, Saint Petersburg; [6]NSU, Novosibirsk; [7]Gymnasium №1, Ulyanovsk; [8]CRG, Barcelona; [9]UPF, Barcelona.

Protein Physics Laboratory



Abstract

It is known that globular proteins fold into their native structure, which allows them to function. Experimental and theoretical studies confirmed the relationship between protein stability (denoted as ΔG) and its function:



FoldX is one of the computer programs that estimates destabilization of mutant proteins. Our first task was to investigate a systematic shift in the prediction of destabilization by FoldX.
The main task was to discover, whether the proteins occurred during the evolutionary course are more (less) stable comparing with the alternative ones.

Protein stability ΔG is a difference between free energies of native (N) and unfolded (U) protein structures: ΔG=G(N) - G(U). For a stable protein ΔG<0. Upon mutation a protein is destabilized by ΔΔG: ΔΔG = ΔGmut - ΔG. For destabilization, ΔΔG > 0; for stabilization, ΔΔG < 0.

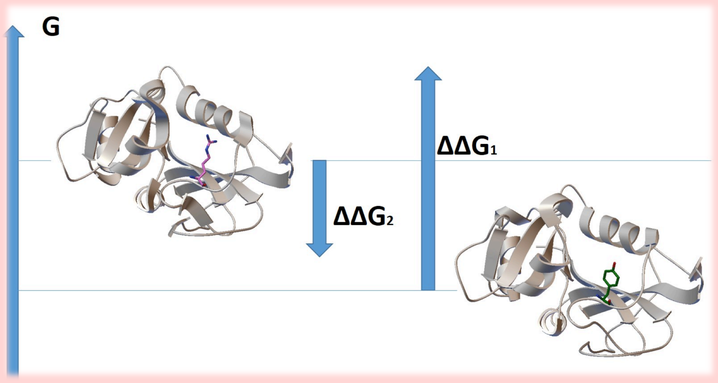


Figure 1. We have found in PDB 532 protein pairs, differing by one amino acid. The destabilization from structure 1 to structure 2 is opposite to the destabilization from structure 2 to structure 1, which causes energy shift to be zero: $\Delta\Delta G_1 + \Delta\Delta G_2 = 0$. However, computer programs have limited accuracy, therefore, the energy shift may differ from zero.

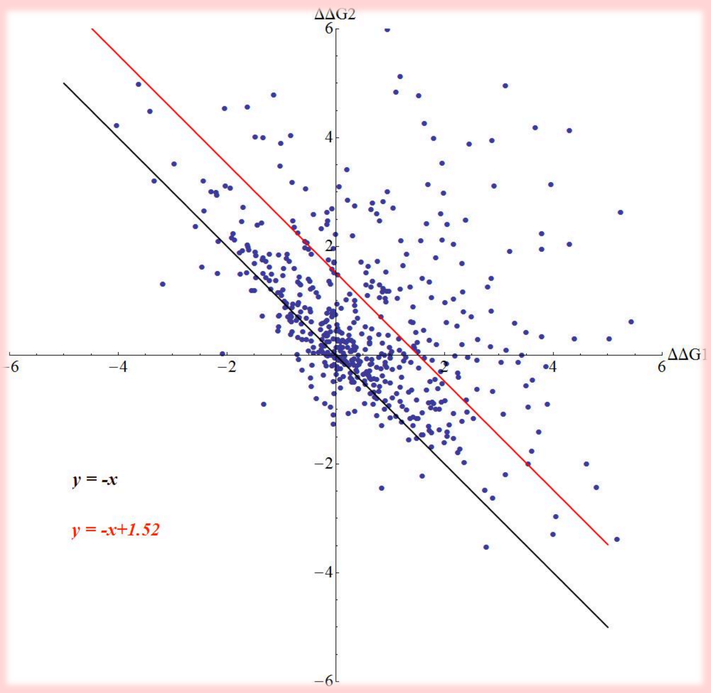


Figure 3. Relationship between $\Delta\Delta G_1$ и $\Delta\Delta G_2$, estimated by FoldX. Line y = --x corresponds to the case when $\Delta\Delta G_1 + \Delta\Delta G_2 = 0$ (expected in the absence of the energy shift). As you can see, using FoldX causes the energy shift equal to 1.52 kcal/mol for two mutations, which must be taken into account.

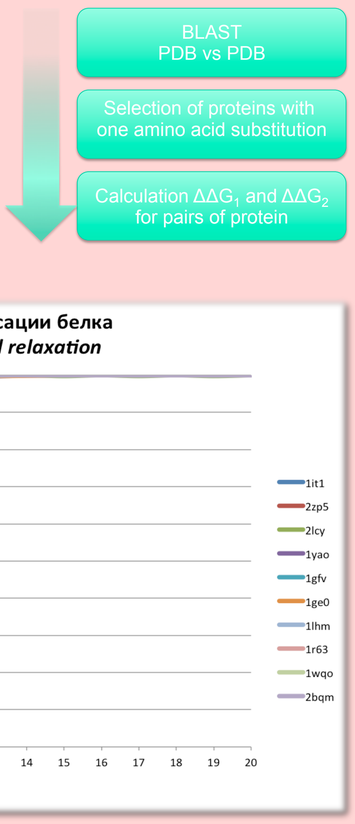


Figure 2. Before conducting any FoldX calculations we need to relax a protein structure. We found that one iteration is not enough for a full relaxation, and it is better to do 10.

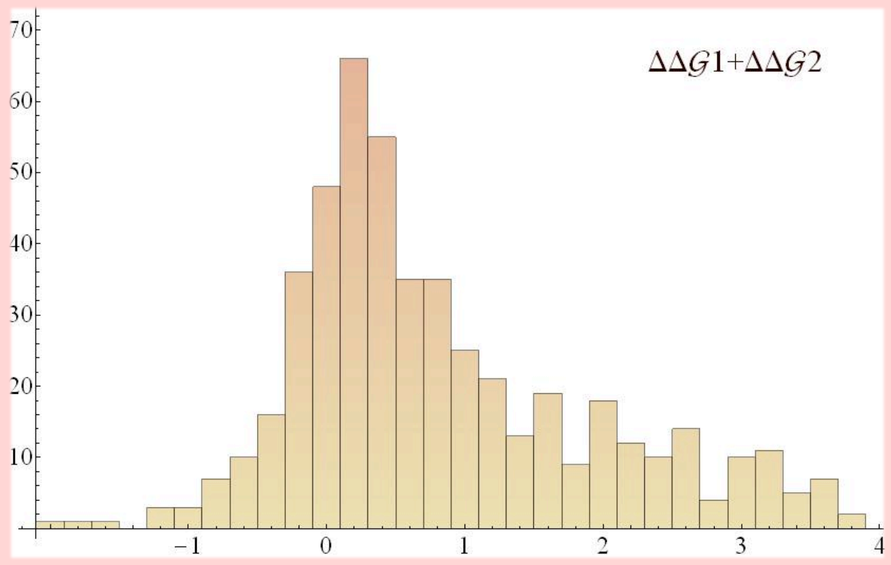


Figure 4. Histogram of energy shift $\Delta\Delta G_1 + \Delta\Delta G_2$

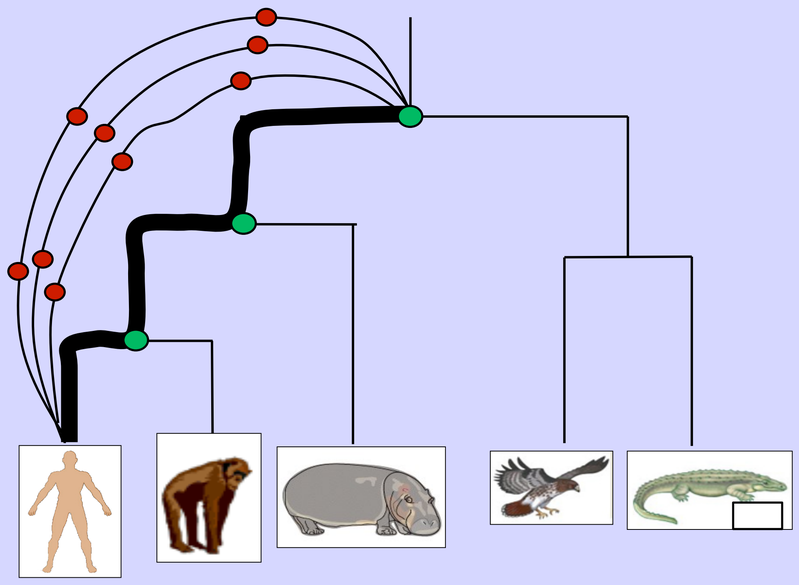


Figure 5. From comparison of vertebrate protein sequences we reconstructed ancestral proteins and the order in which substitutions have occurred. We generated artificial evolutionary trajectories, with the order of substitutions on these trajectories being differing from the really happened one. For all proteins, both from real and artificial trajectories, we modeled their 3D structure and estimated their stabilities.

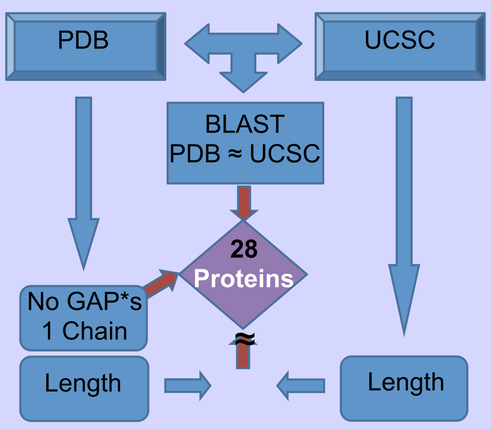


Figure 6. Pipeline for protein selection for evolutionary analysis.

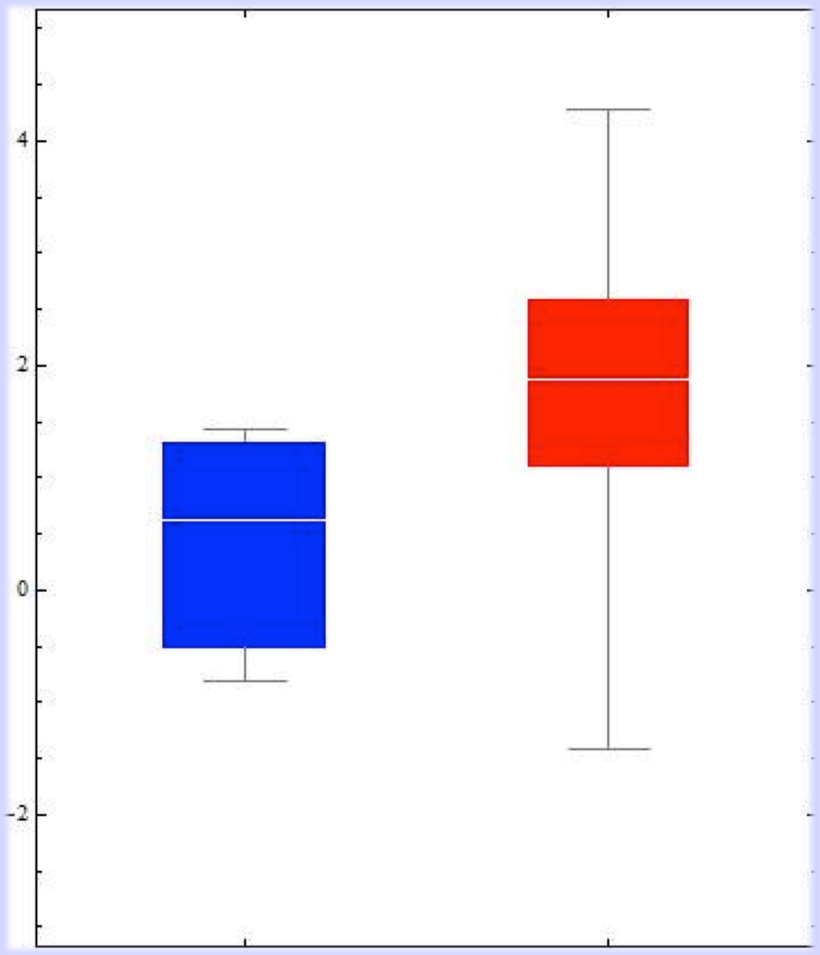


Figure 7. Barplot for destabilization (Y-axis) of different evolutionary states: 1) states observed during the evolution (blue); 2) evolutionary non-observed states (red). Comparison of observed and non-observed protein states in Kolmogorov-Smirnov test gives p-value = 0.004, which means that the destabilizations differ statistically significantly.

Plans:

To study the relationship between the FoldX energy shift and solvent accessibility of residue or mutation type.
To carry out the same computer experiment applying alternative programs estimating mutant protein stability (iMutant, Modeller + FoldX) and compare received data with the result from FoldX alone.
To take into account FoldX energy shift when estimating destabilization of observed and non-observed states in protein evolution.
To check the conclusions about evolutionary analysis on a bigger dataset (additional CPU time is needed).

Conclusions:

1. We discovered FoldX energy shift when estimating destabilization, it equals 0.77 kcal/mol for one mutation, it should be taken into account when using FoldX.

2. We found that evolutionary non-observed states are on average more stable comparing with non-observed protein states.