The amino acids that are important for protein stability are more conserved in evolution: a simple lesson from the study of human hemoglobin mutational database.

Tatiana Molchanova, PhD in Biological Sciences,
Center for Advanced Research and Biotechnology,
University of Maryland, USA.
Email: tatochka@prodigy.net

Mutations are the driving force of molecular evolution. Deleterious or neutral amino acid mutations are frequently classified by the effect on the main biochemical function of the protein. However the conservation of individual amino acids is also dictated by their ability to maintain the stable protein conformation. The contribution of the individual amino acid to the stability of the protein is frequently underestimated partly because the mutant proteins are not easily available and so their stability is difficult to measure.

In this aspect the Human Hemoglobin Mutational Database offers a rare opportunity to analyze the spectrum of the globin chain substitutions in context of their effect on the protein stability. First of all some entries are abnormal hemoglobins detected in patients with a characteristic of Heinz's bodies anemia - a condition resulted from hemoglobin instability in vivo and in vitro. Secondly special protein stability tests were routinely performed for the carriers of abnormal hemoglobins. Thirdly the ratio between mutant and normal globin chain expression was measured in the majority of cases of abnormal hemoglobin identification.

Here I have tried to address the question whether the amino acid residues that are critical for maintaining stability of the hemoglobin molecule are better conserved (preserved) between mammalian species.

Human hemoglobin is a tetramer of 2 alpha (141 amino acids) and 2 beta (146 amino acids) chains. There were discovered 516 single amino acid mutations in human alpha and beta globin chains until 2001. Mutations can be conditionally called stable and unstable. The stable mutations do not change the hemoglobin stability according to the standard stability tests in a comparison with the normal hemoglobin. The unstable mutations destabilize the structure of the hemoglobin molecule under testing.

There were determined 325 mutations in the beta chain and among them there were 139 mutations that destabilize the hemoglobin structure. The respective numbers for an alpha chain are: 191 total mutations and 50 mutations that destabilize the hemoglobin. Mutations were detected in 138 positions in the beta globin chain and in 97 positions in alpha globin chain.

Destabilizing mutations are not distributed homogeneously along the globin chains. I have calculated the number of the unstable and stable mutations for each position along the human alpha and beta chain. Numbers of the mutations that were detected in each position were normalized to the number of theoretically possible mutations in the position due to the limitations by a genetic code. The assumption has been made that the mutation is a result of a single nucleotide substitution. I have found that a pattern of the distribution of destabilizing mutations along both alpha and beta globin chains is different from the pattern of the distribution of all sorts mutations.

Each alpha and beta globin sequence has been divided into conditional segments of 20 amino acids started from N-terminal amino acid (C-termini segment had less than 20 a.a.). Then the percentage of destabilizing mutations per segment have been calculated and plotted along the polypeptide chain.

Then I compared the observed profiles of instability with profiles of an evolutionary conservatism of alpha and beta globin chains. To get a profile of an evolutionary conservatism I have selected 126 mammalian alpha hemoglobin chain sequences and 106 mammalian beta hemoglobin chain sequences from SwissProt Database. Each sequence was subdivided into 20 amino acid segments according the alignment with the respective segments of human alpha and beta chains. I have calculated the percentage of the identical amino acids in each segment of the globin chains of chosen species comparatively the human one. And then I plotted calculated percentages along the sequence of alpha and beta chain respectively. As a result I obtained the simplified profile of the distribution of the level of the identity of the amino acids along alpha and beta chains in different species relatively human one.

I have observed that simplified profile of the distribution of the destabilizing mutations in both human alpha and beta chains are similar to the profile of the identity of the amino acids in the alpha and beta chains of different species. There is higher percentage of identity in the segments where is higher percentage of destabilizing mutations in human chains. The pair-wise correlation coefficients between data sets were calculated. The profile of the distribution of the destabilizing mutations and the profile of distribution of the identical amino acids in different species as compared to

human chains residues were 0.78 for the alpha chains (Fig. 1) and 0.855 for the beta chains (Fig. 3). On the contrary there is no similarity in the profile of the distribution of the stable mutations in both human alpha chain and beta chain as compared with the profile of the distribution of the level of the identity of the amino acids in different species. The correlation coefficients between these data sets were 0.056 for alpha chains (Fig. 2) and 0.084 for beta chains (Fig. 4).

In conclusion the profile of the distribution of the naturally observed destabilizing mutations along the human alpha and beta chains strongly correlates with the profile of distribution of the relatively conserved regions in the mammalian hemoglobins (Fig. 5). These data demonstrate that those hemoglobin destabilizing mutations have the selective pressure during the evolution of the hemoglobin.

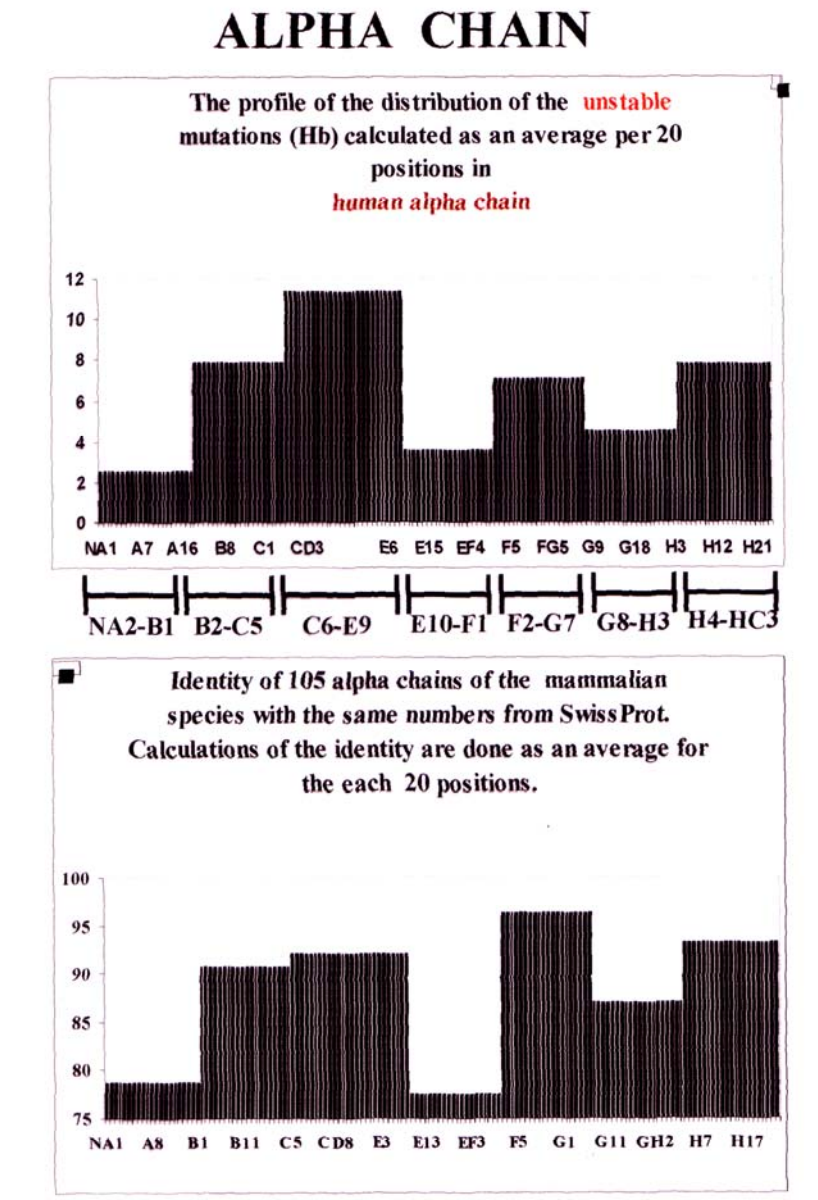


Fig. 1. The profile of the distribution of the destabilizing mutations in the human alpha chain and the profile of distribution of the identical amino acids in different species.

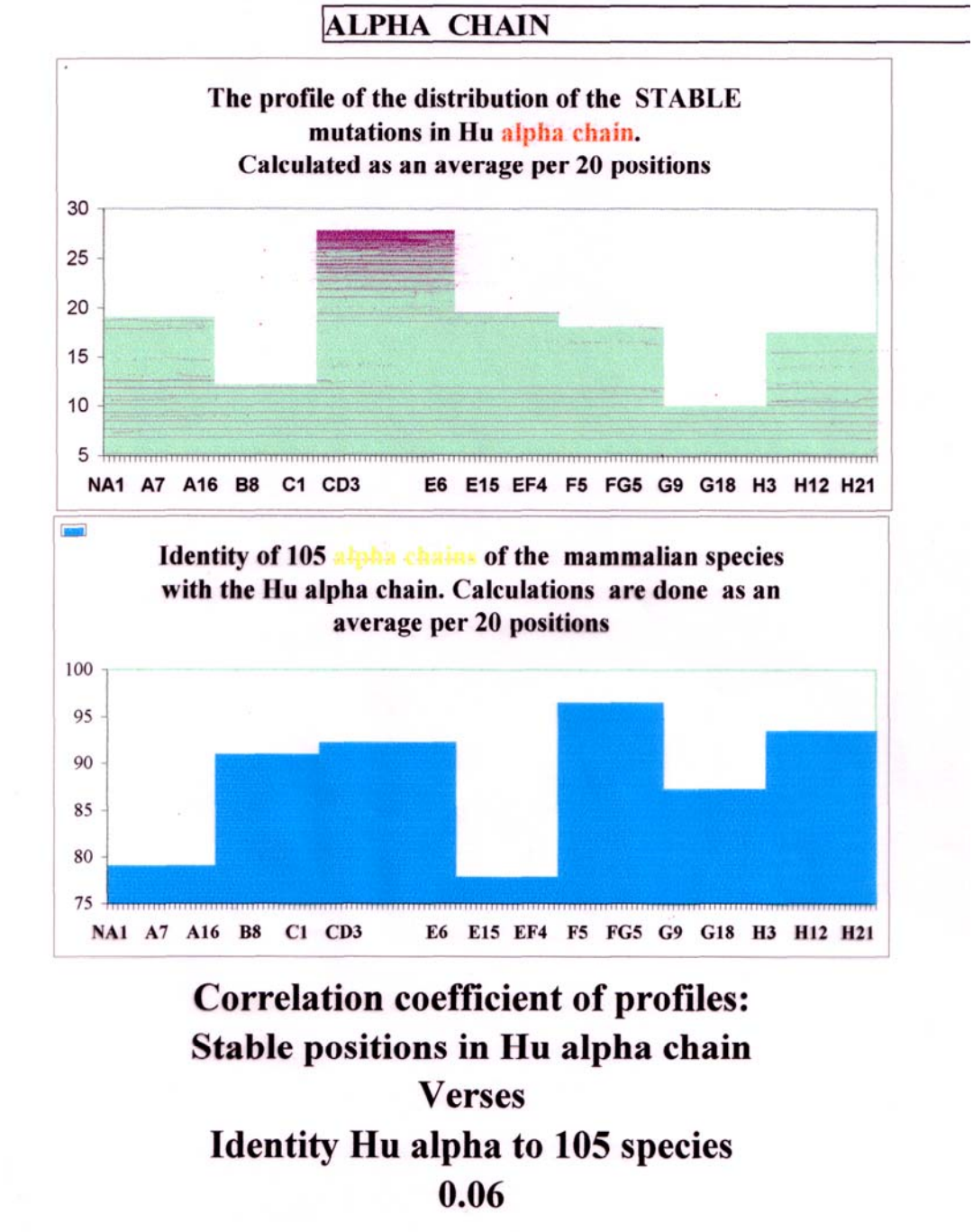
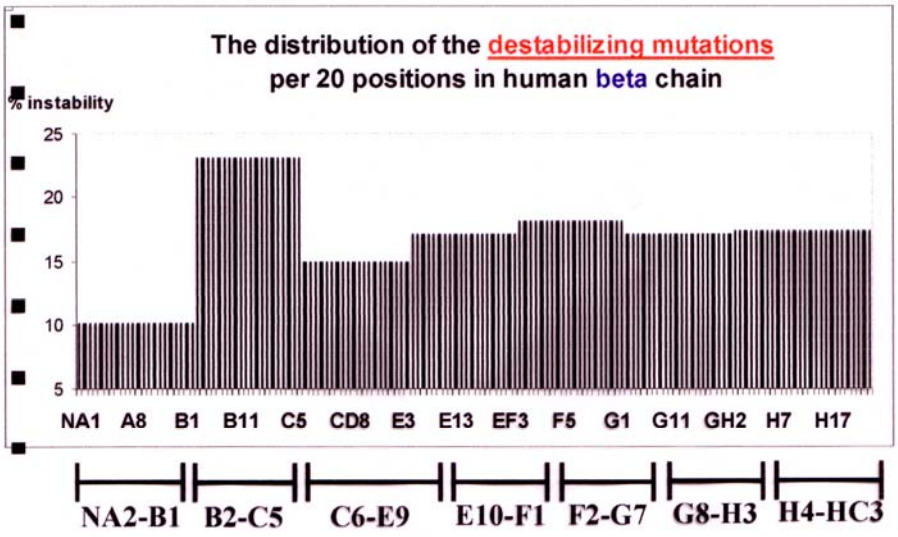
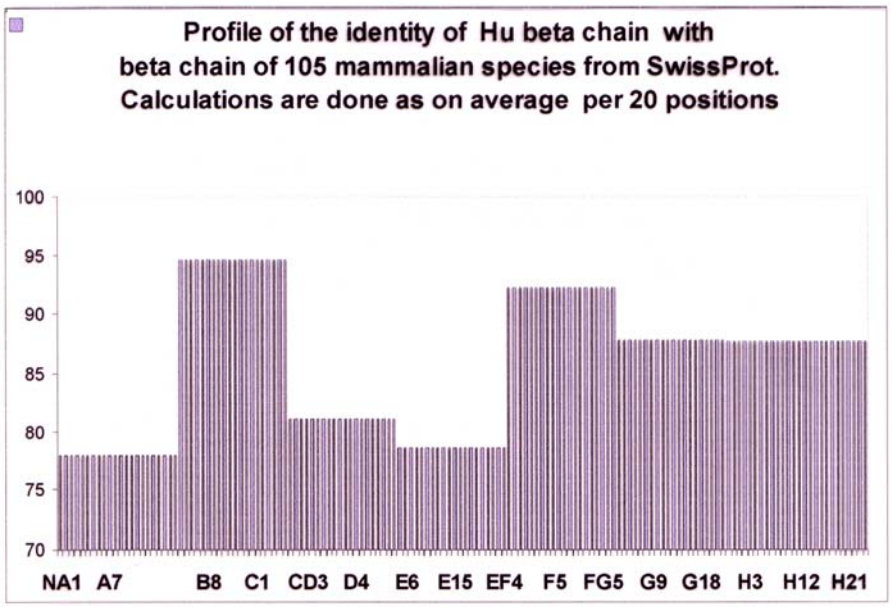


Fig. 2. The profile of the distribution of the stable mutations in the human alpha chain and the profile of distribution of the identical amino acids in different species.

BETA CHAIN





Correlation coefficient
Hu beta-globin chain Instability
verses
Identity of beta chain of 105 species
0.81558613

Fig. 3. The profile of the distribution of the destabilizing mutations in the human beta chain and the profile of distribution of the identical amino acids in different species.

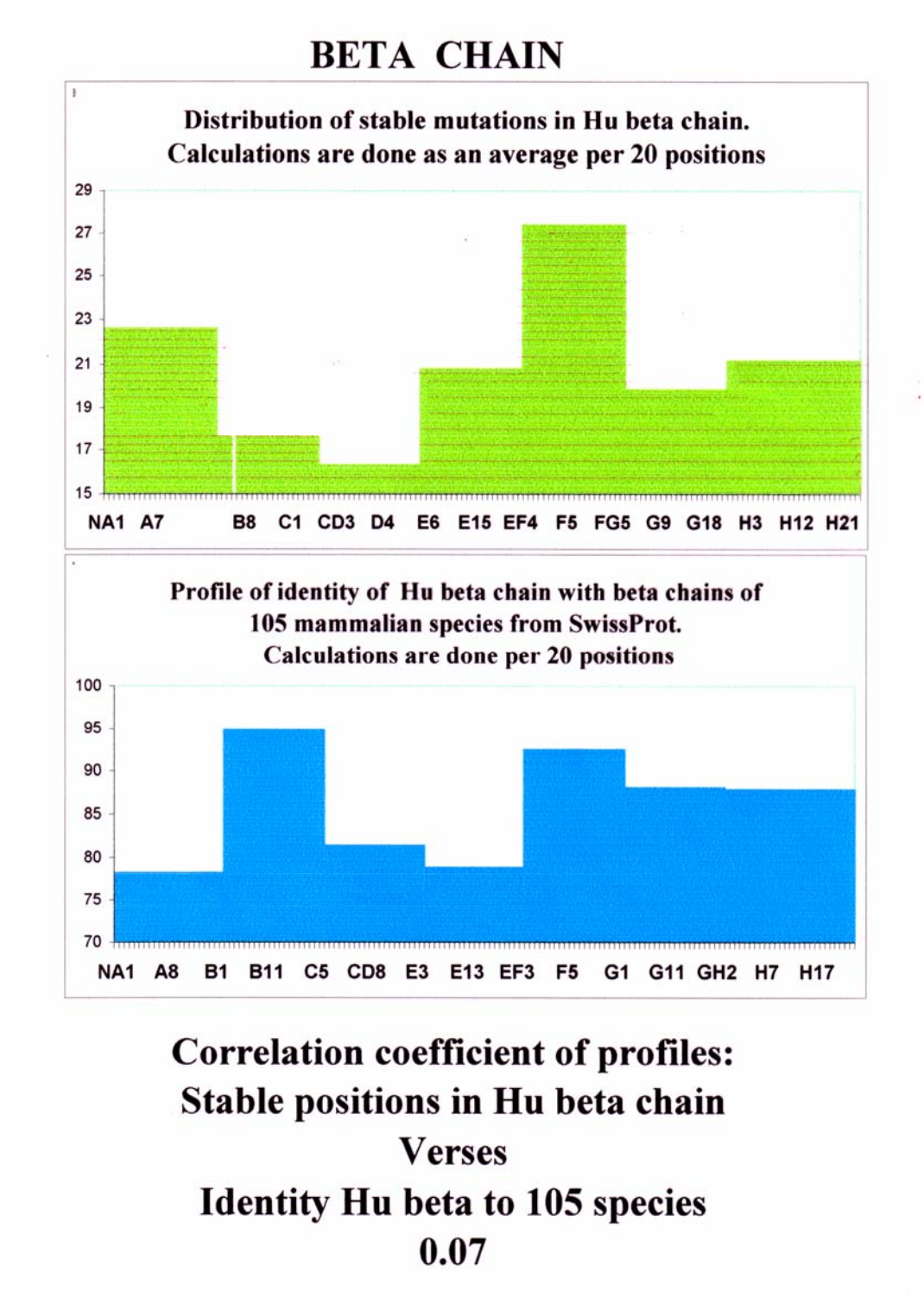


Fig. 4. The profile of the distribution of the stable mutations in the human beta chain and the profile of distribution of the identical amino acids in different species.

Correlation Coefficients of the Profiles of the Destabilizing Mutations, 'Stable' mutations in Human Alpha and Beta Hemoglobin Chains With the Profiles of Identity of Human Alpha and Human Beta Hemoglobin Chains To Alpha and Beta chains of Mammalian Species

	ALPHA	BETA
Destabilizing / Identity	0.779	0.815
'Stable' / Identity	0.056	0.065

Fig. 5.
The correlation coefficient of the profiles: the distribution of the destabilizing mutations versus the profile of distribution of the identical amino acids in the hemoglobin of different species.
The correlation coefficient of the profiles: the distribution of the stable mutations versus the profile of distribution of the identical amino acids in the hemoglobin of different species.

Presentation in the National Center for Biotechnology Information,
National Library of Medicine, National Institutes of Health,
Bethesda MD, USA. July 26, 2001.