Dissecting alpha helices in HIV proteins

Hossein Fallahi
Dep. Of Biology
Razi University
Kermanshah, I.R of Iran
E-mail: fallahi.hossein@gmail.com

Ali Bidmeshkipour
Dep. Of Biology
Razi University
Kermanshah, I.R of Iran
E-mail: abidmeshki@razi.ac.ir

Abstract—Many studies have focused on HIV (Human immune deficiency virus) because of its importance to human beings. To find appropriate drug for treating this virus studies have been conducted to find a suitable target for drug design. Here a comparison was made between alpha helices in HIV and some other proteins to investigate if there are any differences between the content of amino acids in different positions of HIV proteins and those of other proteins. Secondary structures in all nine proteins of HIV were predicted and dissected. The results indicate that alpha helix number and size varies among the HIV proteins and between HIV proteins and other proteins in our dataset. Long alpha helices were found in HIV proteins, which could not be found in other proteins. On the other hand, some amino acids showed differential preferences toward HIV and other proteins in the dataset. These finding could be useful for choosing appropriate target for drug design.

Keywords- alignments; alpha helices; substitution matrices; secondary structure; protein evolution.

I. INTRODUCTION

Many studies indicated that different amino acids have different preferences for different positions in protein secondary structures. As a result of these studies, substitution matrices have been developed to compare or align the secondary and tertiary structures of proteins [1,2,3]. In all of these matrices the score of replacement of an amino acid in alpha helices is the same as replacement of this amino acid in a beta sheet. To understand if amino acids have similar affinity toward these secondary structures in HIV proteins we have studied the secondary structures of proteins based on their amino acid compositions. In this work we have presented the results of our analysis of alpha helices in HIV proteins and some other proteins.

II. METHODS

The amino acid sequences of HIV and one hundred proteins containing alpha helices were obtained from PDB [4]. Alpha helices in these proteins were then categorized based on the number of amino acids into classes with 3 to 25 amino acids. In addition alpha helices in HIV proteins were predicted using SABLE protein prediction server and all the alpha helices were identified. The amino acid contents of each class and each position within the classes were then determined. Amino acid contents of the same position in different classes were also compared. All the scripts were written using PERL language programming and tested locally on a Mac operating system (Mac OSX ver.10.5.8).

Figure 1. Distribution of alpha helices in different size classes. Alpha helices from HIV proteins and all other proteins were classified according to the number of amino acids.

Figure 2. Distribution of amino acids in the alpha helices of HIV and all other proteins. Percentage of amino acid contents of alpha helices were presented in the graph.
III. RESULTS AND DISCUSSIONS

We observed significant differences in the size distribution of alpha helices in the HIV and other proteins in our data set. It appears that alpha helices with 9-14 amino acids were more abundant in HIV while 3-amino acids helices were the predominant form of helices in the rest of proteins (Fig.1).

Besides, distribution of amino acids in the protein data set was different from those of HIV proteins. While Ala, Pro, Asp, Tyr were mostly present in our protein data set, HIV proteins contained higher percentages of Glu, Ile, Leu, Gln and Arg (Fig.2). In addition, in HIV proteins Leu was the most abundant amino acids, while Ala was found in rest of proteins at higher level compared to other amino acids (Fig. 2).

Analyzing the positional effect of amino acids on the structure of HIV proteins and positional preferences of the amino acids toward different positions is underway and the results will be presented at the conference.

ACKNOWLEDGMENT

This work was partially supported by Razi University, through a research fund provided to Dr. Fallahi.

REFERENCES