Recent Research Results on the Application of DSP to Genomics

Parameswaran Ramachandran

and

Andreas Antoniou

Department of Electrical Engineering,
University of Victoria, BC, Canada.

Identification of Hot Spots in Proteins

- Different regions of a protein molecule interact among themselves.
- These interactions enable the string-like protein molecules to fold into complex 3-dimensional (3-D) structures.
- Protein molecules perform their function by selectively binding to other molecules by virtue of their 3-D structures.
- The regions of a protein molecule where selective bindings occur are called hot spots.

Identification of Hot Spots (cont'd)

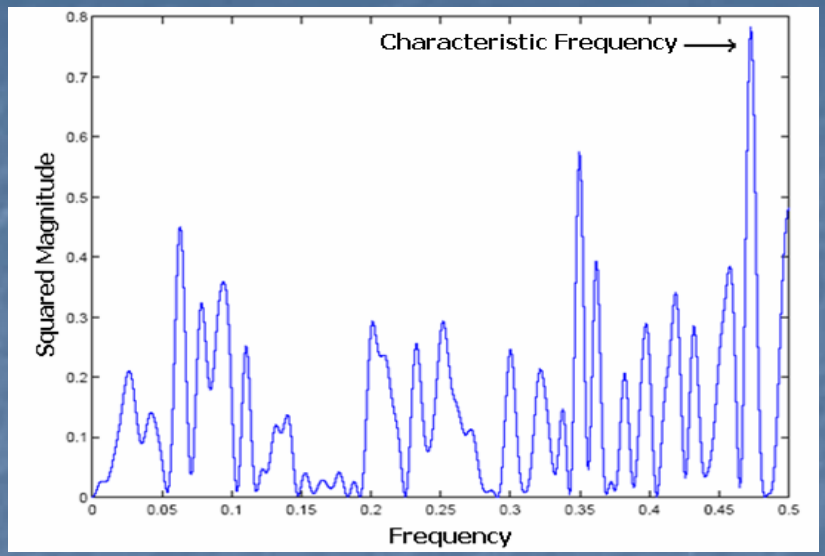
- The energies of free electrons in the amino acids can be represented by numbers called EIIP (electron-ion interaction potential) values.

  By assigning these numbers to the amino acids, a protein can be represented by a numerical sequence.

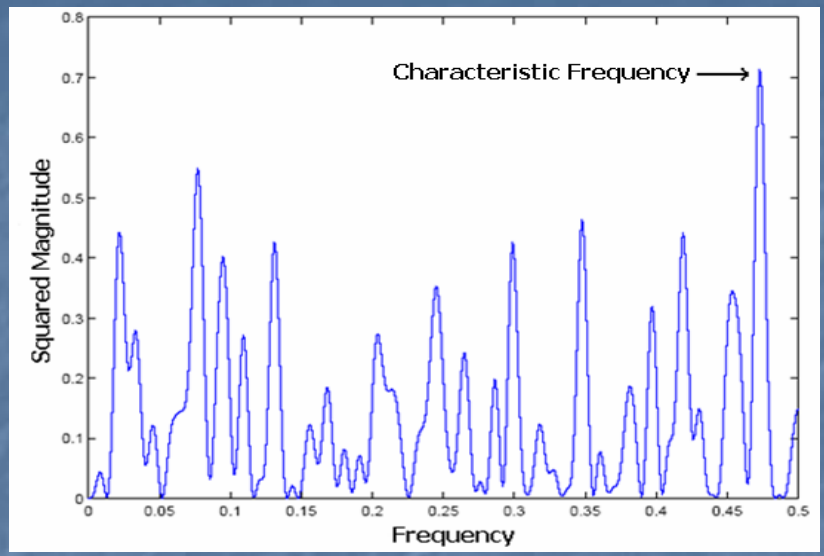
- The DFTs of the numerical sequences of a group of proteins performing similar functions have been observed to share a unique frequency component (see Cosic [1]).

  This is called the characteristic frequency of the particular group.

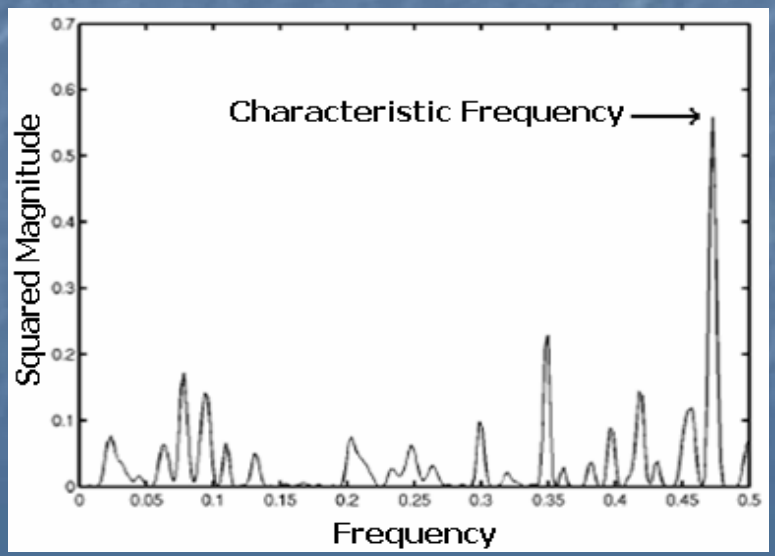
Identification of Hot Spots (cont'd)



DFT of mouse cytochrome C protein



DFT of tuna cytochrome C protein



Product of the two DFTs

- The hot spots in a protein can be located by identifying the regions where the characteristic frequency is dominant.

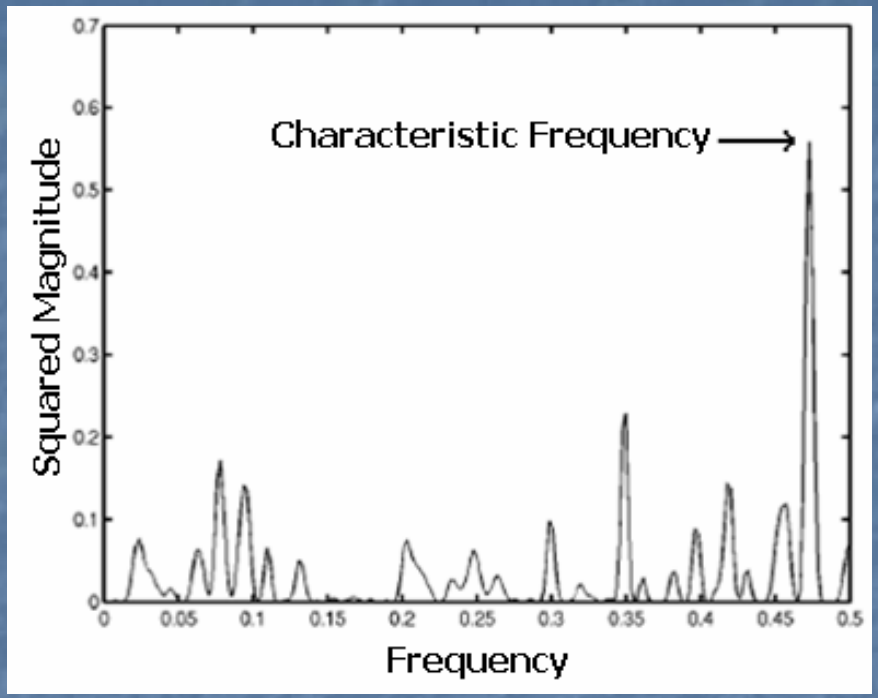
P. Ramachandran, A. Antoniou, and P. P. Vaidyanathan, *38 th Asilomar Conference on Signals, Systems, and Computers,* November 2004.

Identification of Hot Spots (cont'd)

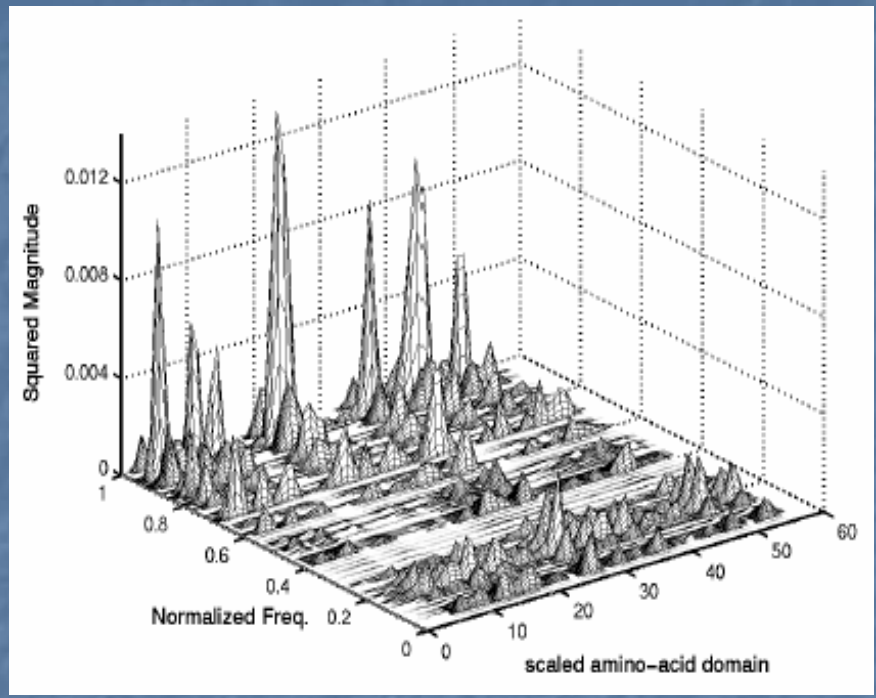
- Four steps:
  - Convert a number of proteins of a particular functional group into numerical sequences
  - Compute the DFTs of the sequences and multiply them pointwise to determine the characteristic frequency
  - Compute the STFT of the protein sequence in question using a suitable window
  - Multiply each column of the STFT by the DFT product

P. Ramachandran, A. Antoniou, and P. P. Vaidyanathan, *38 th Asilomar Conference on Signals, Systems, and Computers,* November 2004.

Identification of Hot Spots (cont'd)



Product DFTs of cytochrome C proteins



STFT of tuna heart cytochrome C protein, with peaks as the hot spots

- The 3-D plots could be displayed as color-coded contour plots, just like the spectograms!

P. Ramachandran, A. Antoniou, and P. P. Vaidyanathan, *38 th Asilomar Conference on Signals, Systems, and Computers,* November 2004.

Conclusions

- An efficient method to identify and locate hot spots in proteins was developed.
- The hot spots identified were found to match well with other available published data.
- The method is being tried on a variety of proteins from different organisms to verify its usefulness.
- Work on improving the method in terms of its accuracy and ease of use is being carried out.
- The use of digital filters for identifying hot spots is also being explored.

References

1. I. Cosic, “Macromolecular bioactivity: is it resonant interaction between macromolecules? – theory and applications,” *IEEE Trans. on Biomedical Engr.*, vol. 41, no. 12, pp. 1101-1114, Dec. 1994.
2. P. Ramachandran, A. Antoniou, and P. P. Vaidyanathan, *38th Asilomar Conference on Signals, Systems, and Computers,* November 2004.