Direct Computerized Translation of Biological data into Biological Information is now feasible: the Gains of Digital Signal Processing-based Bioinformatics Techniques

Norbert .I. Nwankwo

University of Port Harcourt, Nigeria.

Abstract

- Bio-functionality assessment e.g. Histaminic activity of drugs/extracts via clinical approaches are costly, wasteful and labor-intensive.
- Rational computerized approaches as Digital Signal Processing (DSP) Bioinformatics procedures, etc have become necessary.
- **The question is "*Direct Computerized Translation of Biological data into Biological Information now feasible*?".**
- In 1985, Serbian researchers saw proteins/peptides as signals (rather than as meat, enzymes, etc);
- They then engaged DSP procedures on them.
- Today, DSP and other procedures (e.g. geno2pheno) have made it feasible to translate data to information.
- It is demonstrated using VIPMFSALS and CAPAGFAIL.

Introduction:

- Biological functionalities e.g. disease processes, pharmacological, structural, physiochemical properties, etc such as retroviral and antiretroviral activities are encoded in the genes/proteins [1].
- Genes/proteins provide as much biological information as the therapeutic and disease causative agents. E.g. Mutations in the HIV gp120 are known to translate HIV to AIDS.
- Proteins/peptides are Amino Acids in linear formation, alphabetically codified [2].
- ✓ Proteins, now seen as signals/numerical sequences [3] can now be analyzed using DSP techniques that is the basis of Radar Technology, Speech Detector, etc.
- ✓ DSP techniques e.g. Informational Spectrum Method (ISM) [3] help uncover biological information embedded in them.

How does ISM work?

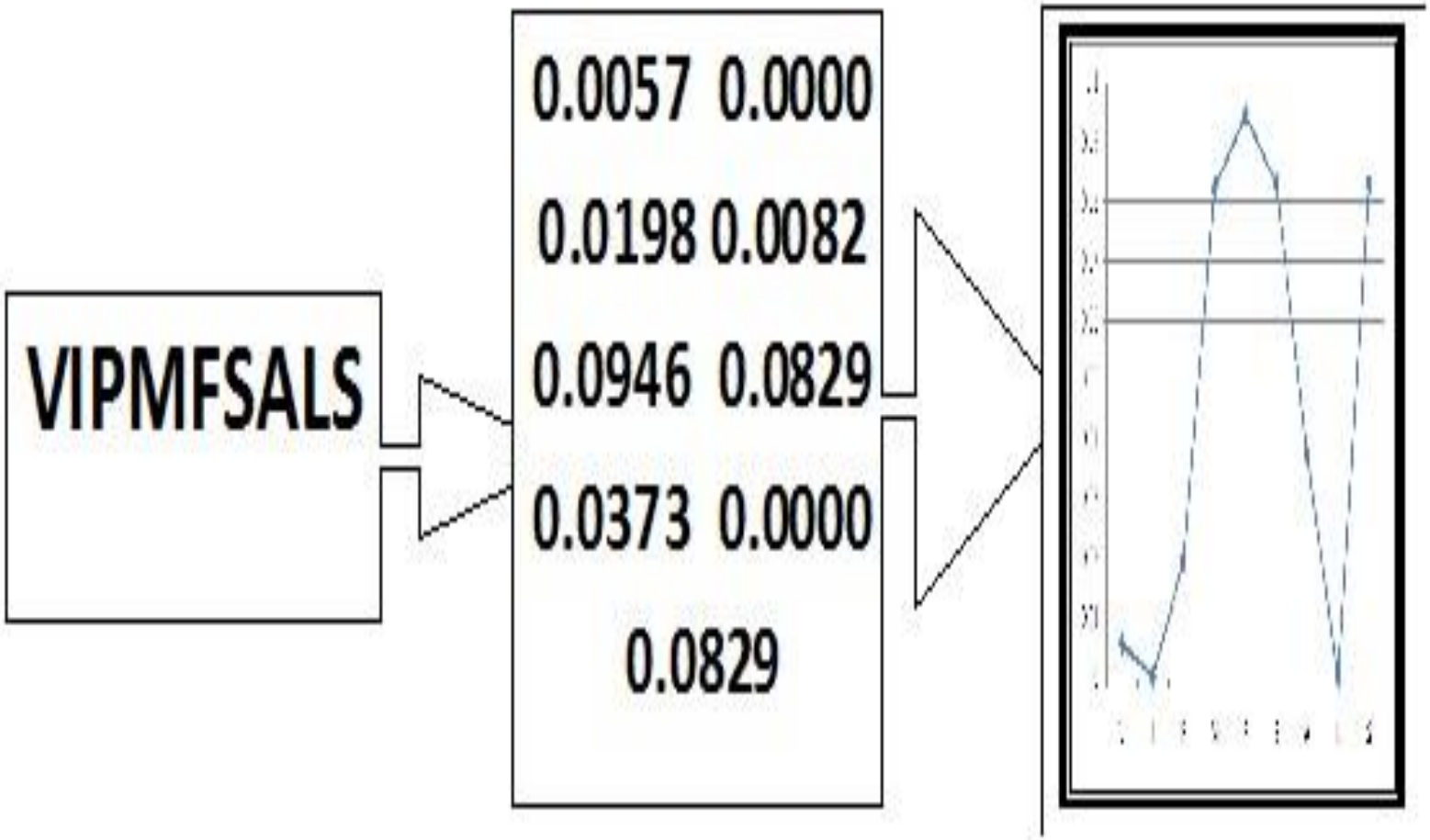
- Peptidic sequences of VIPMFSALS (Pep1) and CAPAGFAIL (Pep2) are retrieved from UNIPROT database [4].
- Their alphabetic codes are translated into numbers using an Amino Acid Scale called EIIP [5].
- Amino Acid Scales are parameters that express the level of individual participation of the 20 essential amino acids in each interaction [6], which are over 525 AASs [7].
- Translations result in two signals (numerical sequences) as shown in Slides 6 and 7.
- The signals are processed using Discrete Fourier Transform (DFT) [8].

Amino Acids Scales: e.g. Electron-Ion Interaction Pseudo-potential (EIIP) i.e. binding interaction

Some of the Amino Acid Scales are available at www.genome.jp/aaindex [5]

Amino Acid	EIIP	Amino Acid	EIIP	Amino Acid	EIIP	Amino Acid	EIIP
A	0.0373	Q	0.0761	L	0.0000	S	0.0829
R	0.0959	E	0.0058	K	0.0371	T	0.0941
N	0.1263	G	0.0050	M	0.0823	W	0.0548
D	0.0036	H	0.0242	F	0.0946	Y	0.0516
C	0.0829	1	0.0000	P	0.0198	V	0.0057

The making of the signals: VIPMFSALS



The making of the signals: CAPAGFAIL

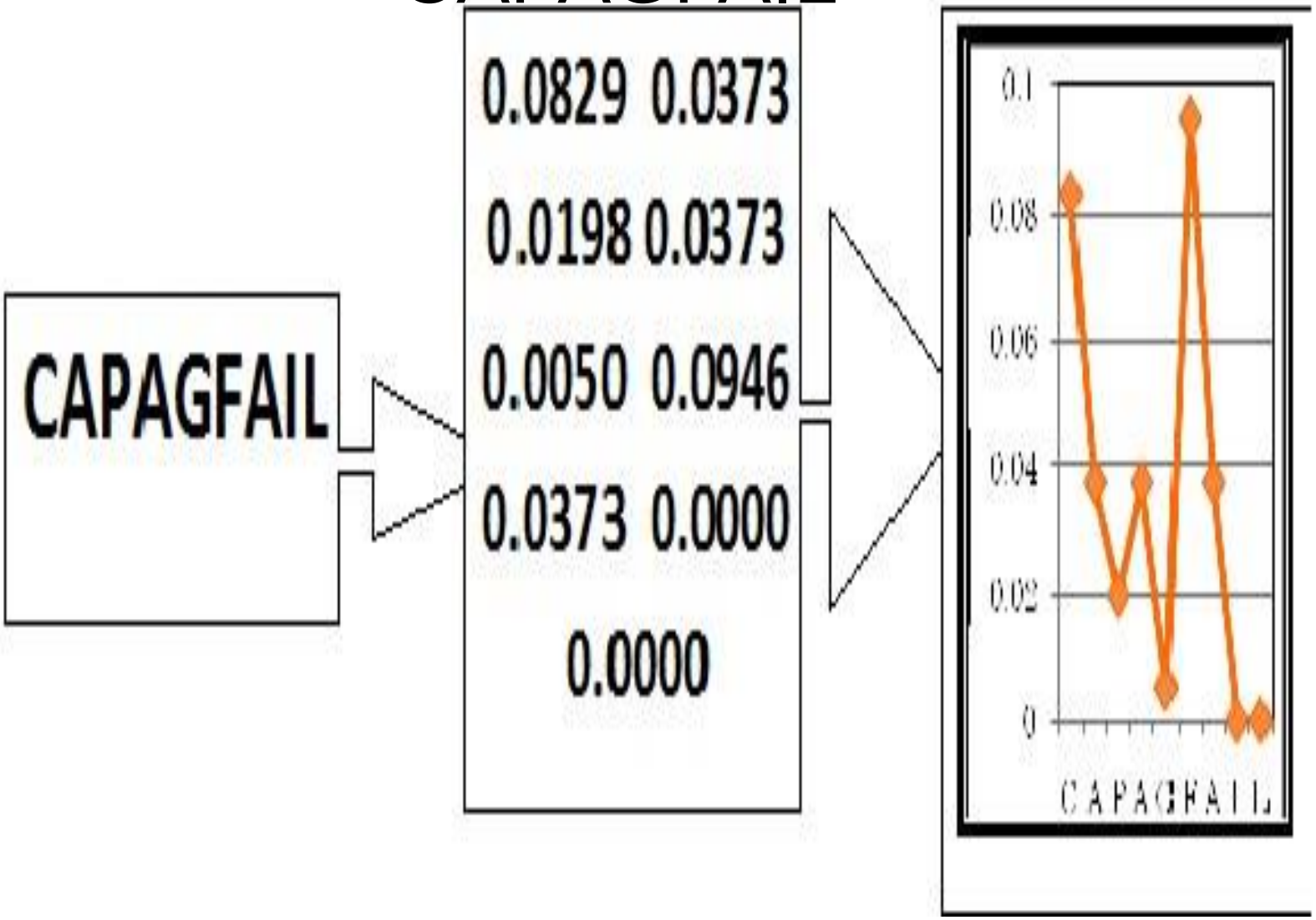
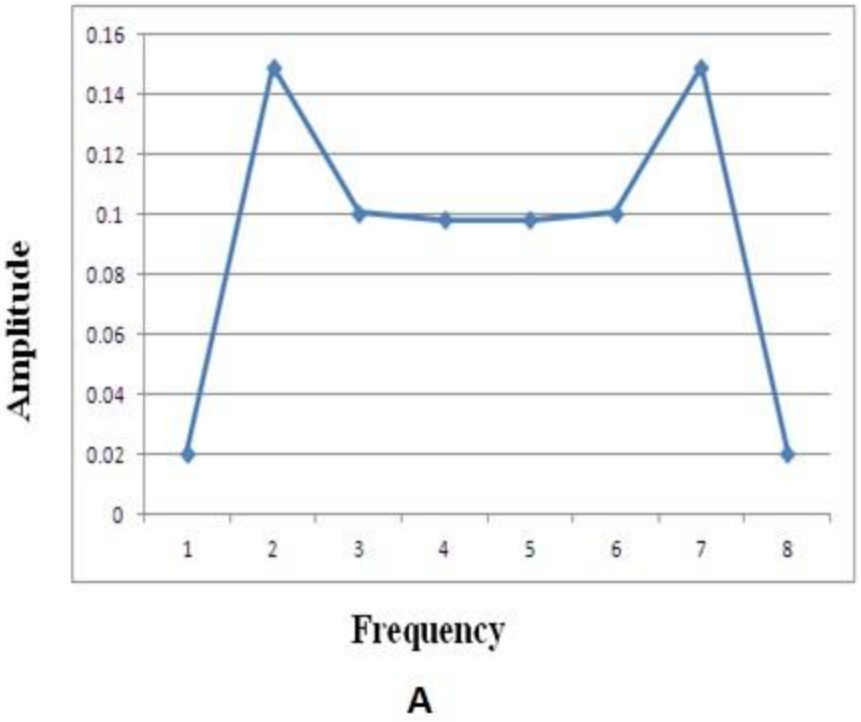


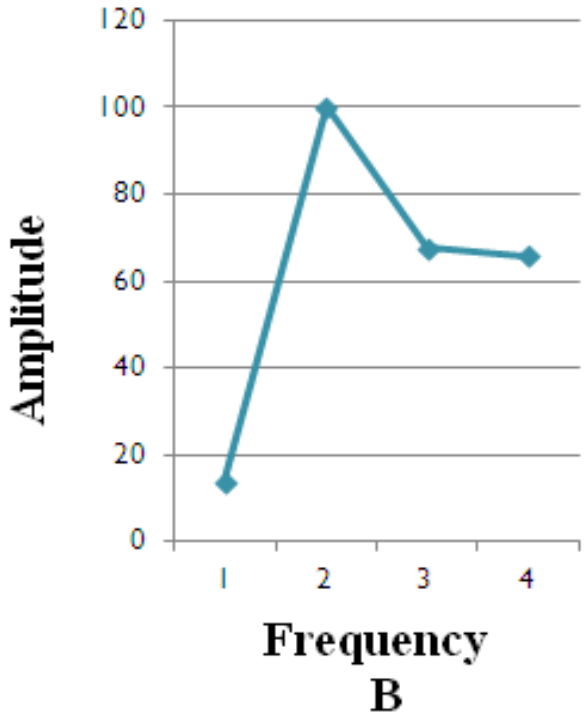
Table 2: DFT results for Peptides 1 and 2 each (Informational Spectrum) and combined (Common Informational Spectrum)

	Peptide1	Peptide2	Peptides1*2
n	Amplitude	Amplitude	Point-wise Multiplication
0	DC	DC	-
1	0.1292	0.0205	0.0026
2	**0.1282**	**0.1492**	**0.0191**
3	0.1188	0.1008	0.0120
4	0.0919	0.0982	0.0090
5	0.0919	0.0982	0.0090
6	0.1188	0.1008	0.0120
7	**0.1282**	**0.1492**	**0.0191**
8	0.1292	0.0205	0.0026

Plots of the DFT of the numerical sequence showing mirror image (A), and its half (B).

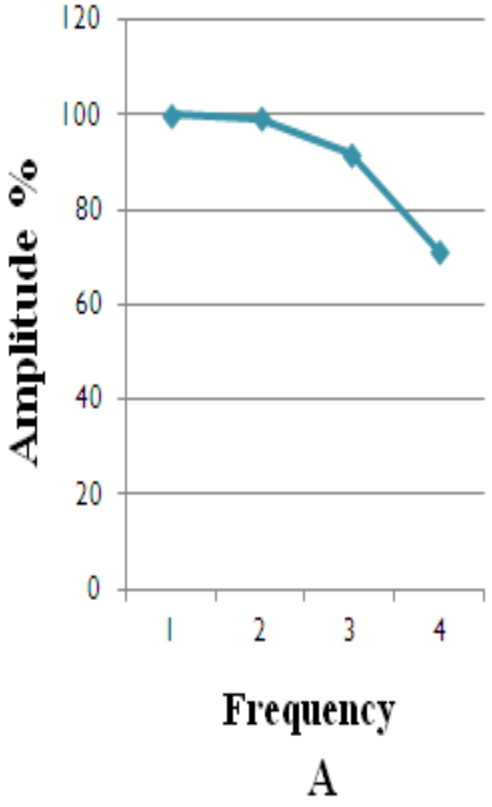


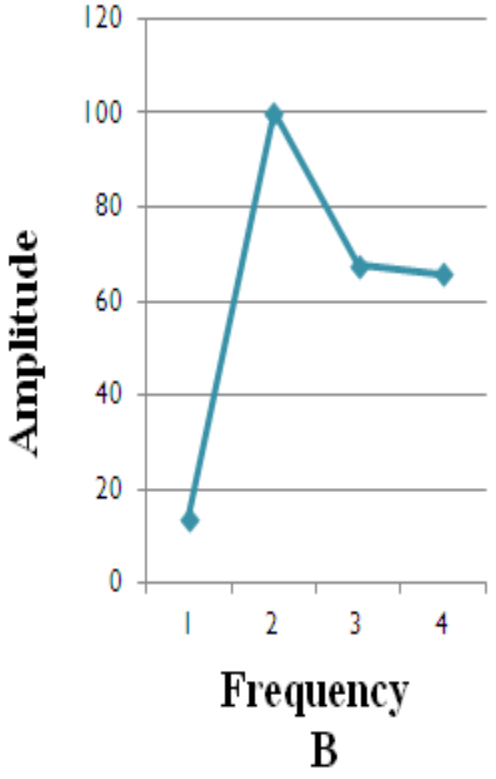
A

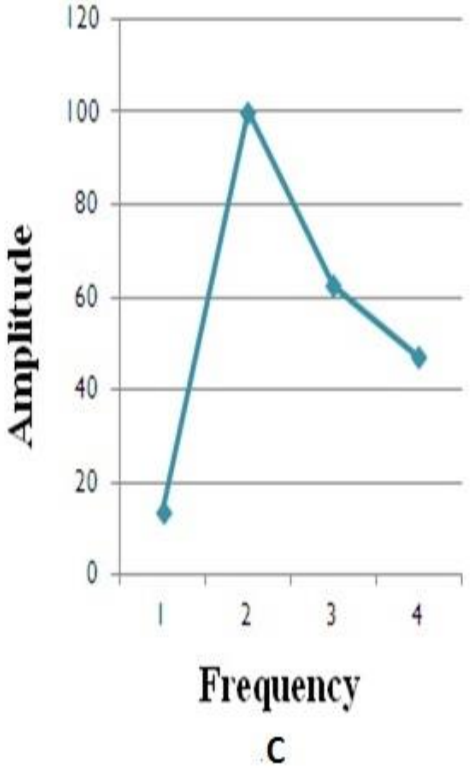


B

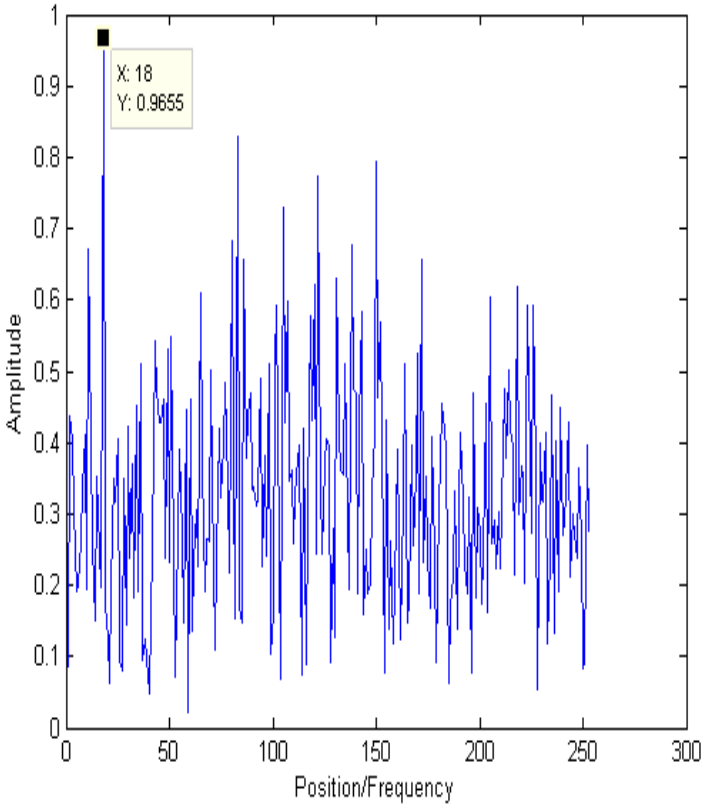
Informational Spectrum of Pep1 (A) and Pep2 (B). and the Common Informational Spectrum of both (C)



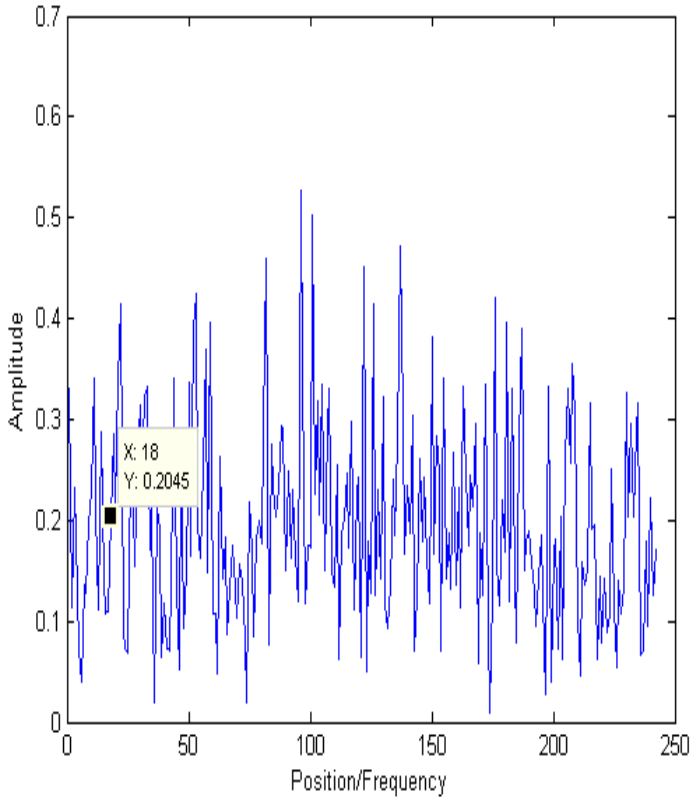




Typical Informational Spectra of HIV-1 T-tropic HXB3 (A) and M-Tropic YBF30 (B), showing affinities 96.55% and 20.45% respectively for the host CD4 at position 18,



A

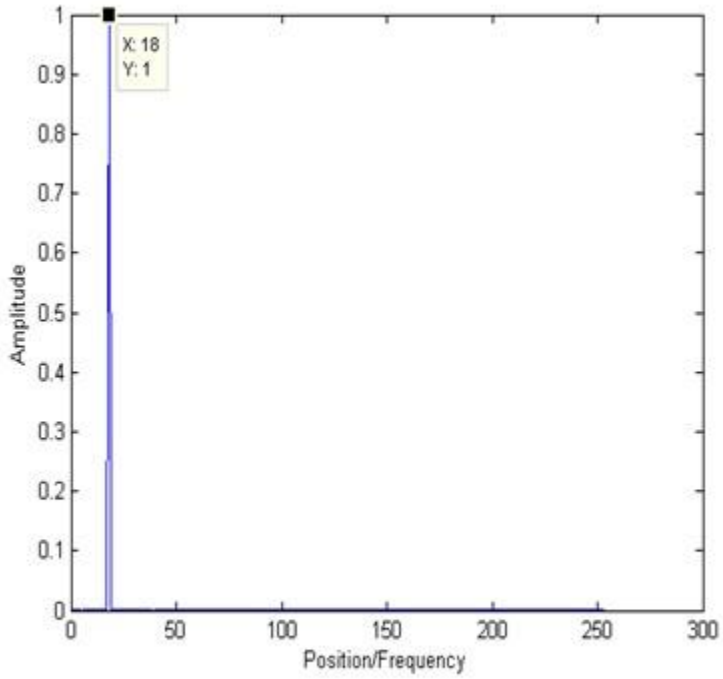


B

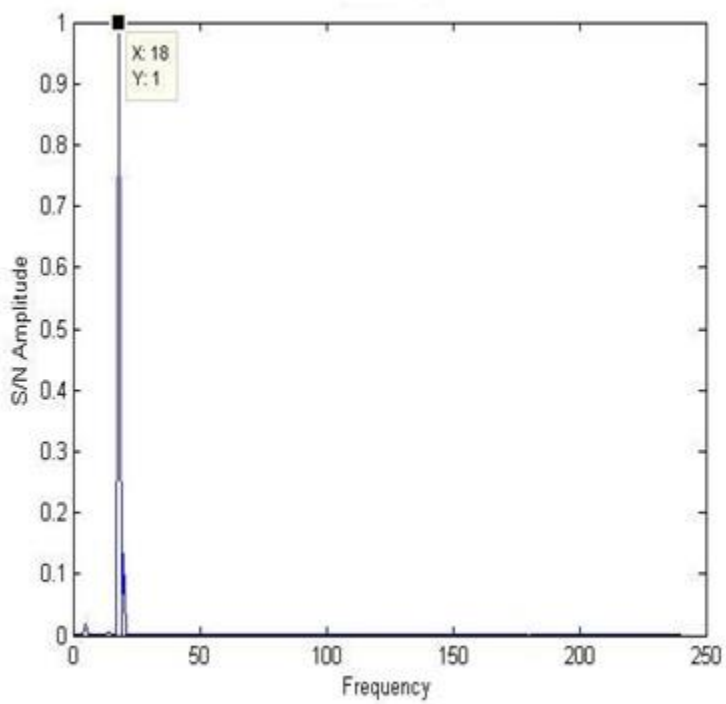
The ISM principles:

Proteins with common biological characteristics share same Consensus Frequency (CF) or point of interaction.

Example : gp120 of 53 HIV isolates(A) and CD4 18 host (B), both have 100% affinity for each other at position 18 [6, 9]

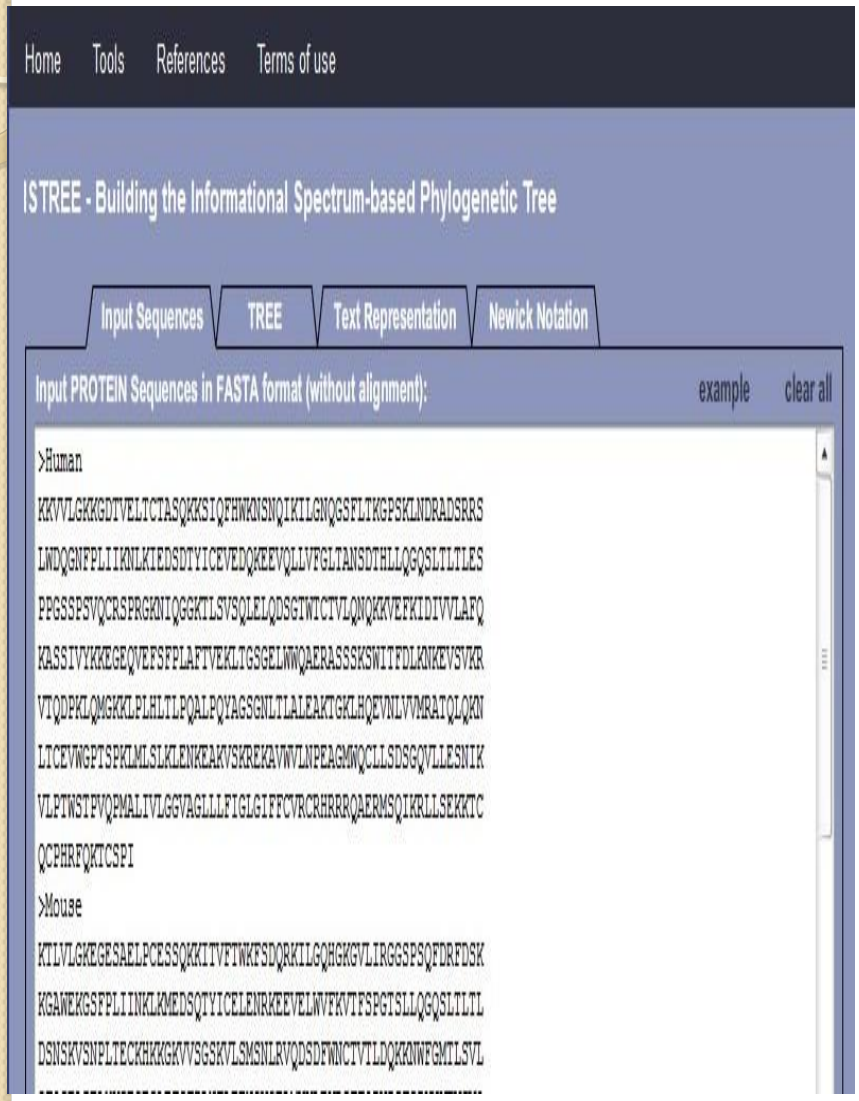


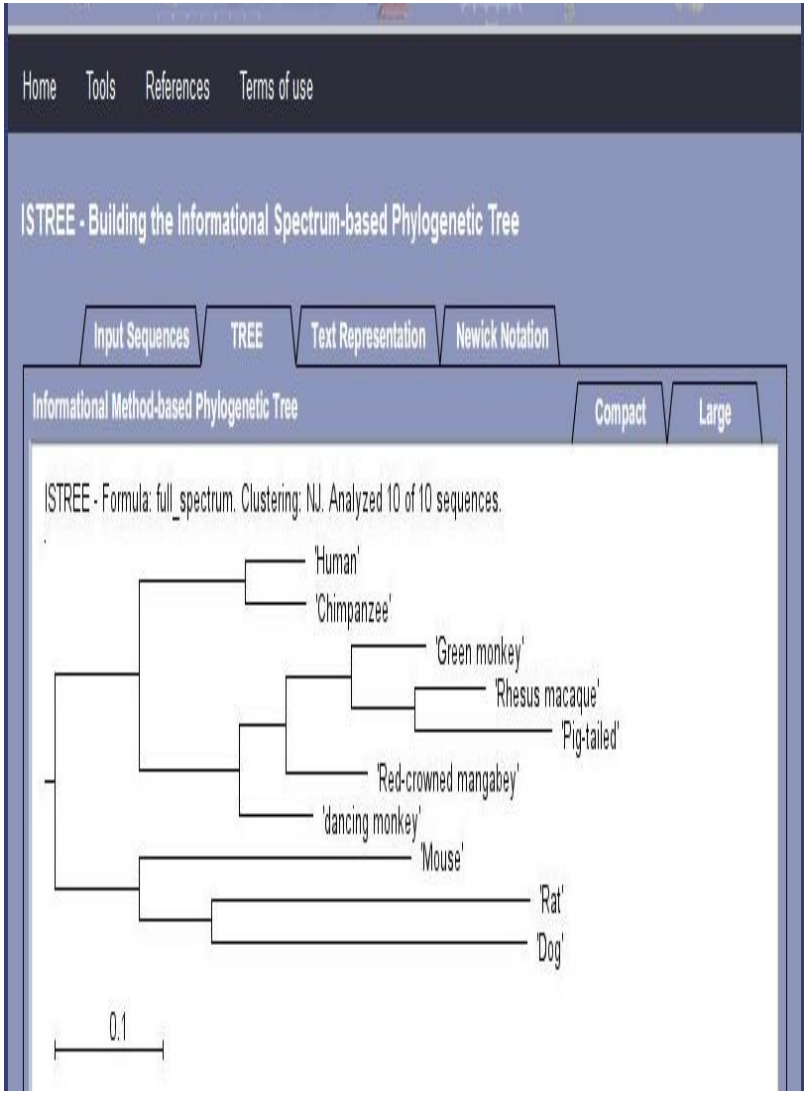
A



B

ISM-based, Sequence Information-oriented Biological data Transformation to Bio-functionalities: Example No 1





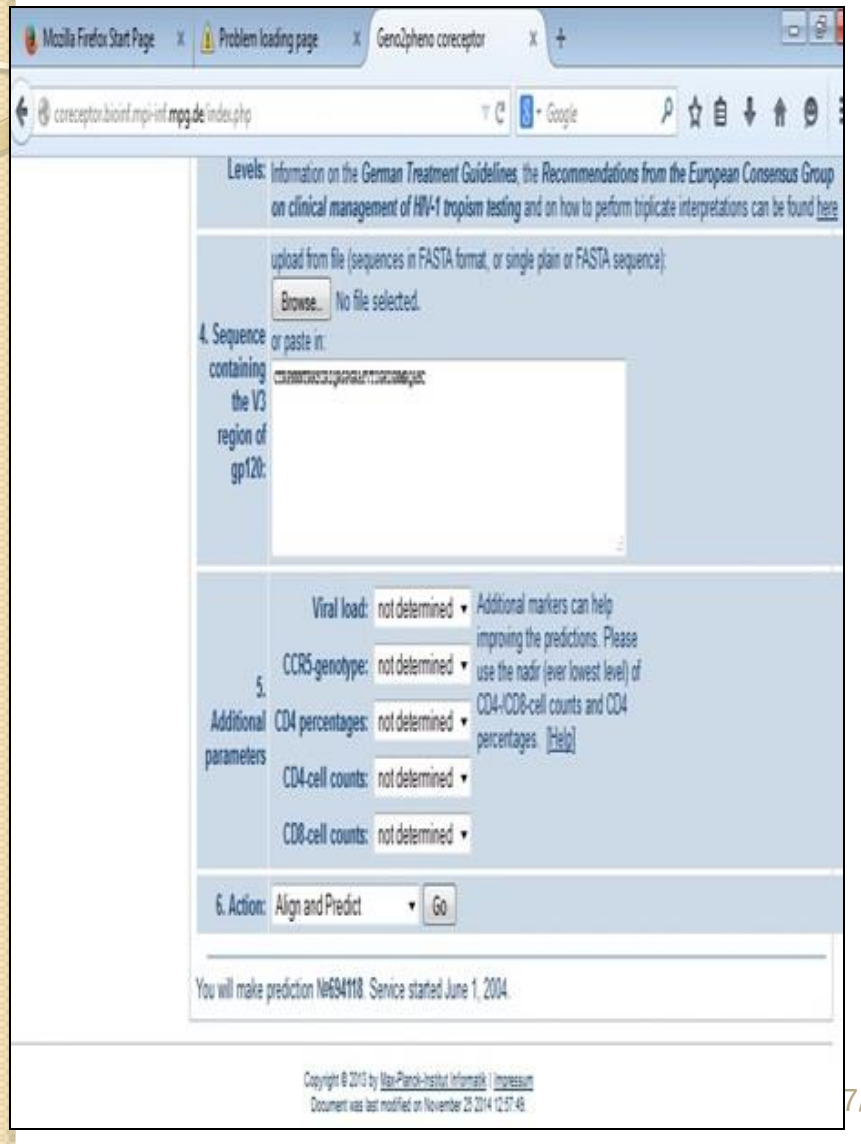
Other Biological functionalities obtained by Analyzing sequence information 1

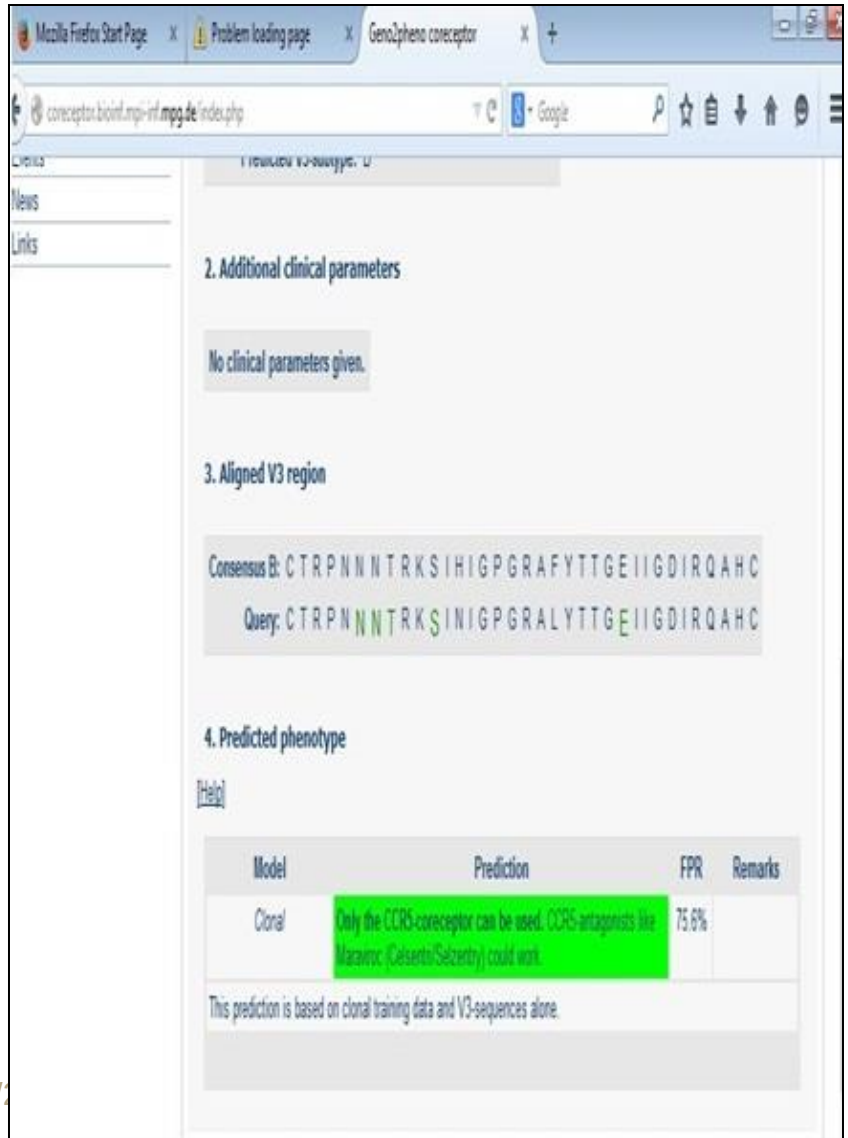
1. Explanation disease processes e.g. HIV progression to AIDS [6. 9].
2. Assessment of vaccine potency (Innocentive Challenge Award-ID: 9933477).
3. Identification of the origins of HIV-1 non B subtypes harboured by American soldiers [6. 10].
4. Calculation of biological functionalities [6. 11].
5. Comparison of potencies of Pharmacological activities of two anti-retroviral agents [6].
6. Development of Biomedical device: Computer-Aided Drug Resistance Calculator [6. 12].

Other Biological functionalities obtained by Analyzing sequence information 2

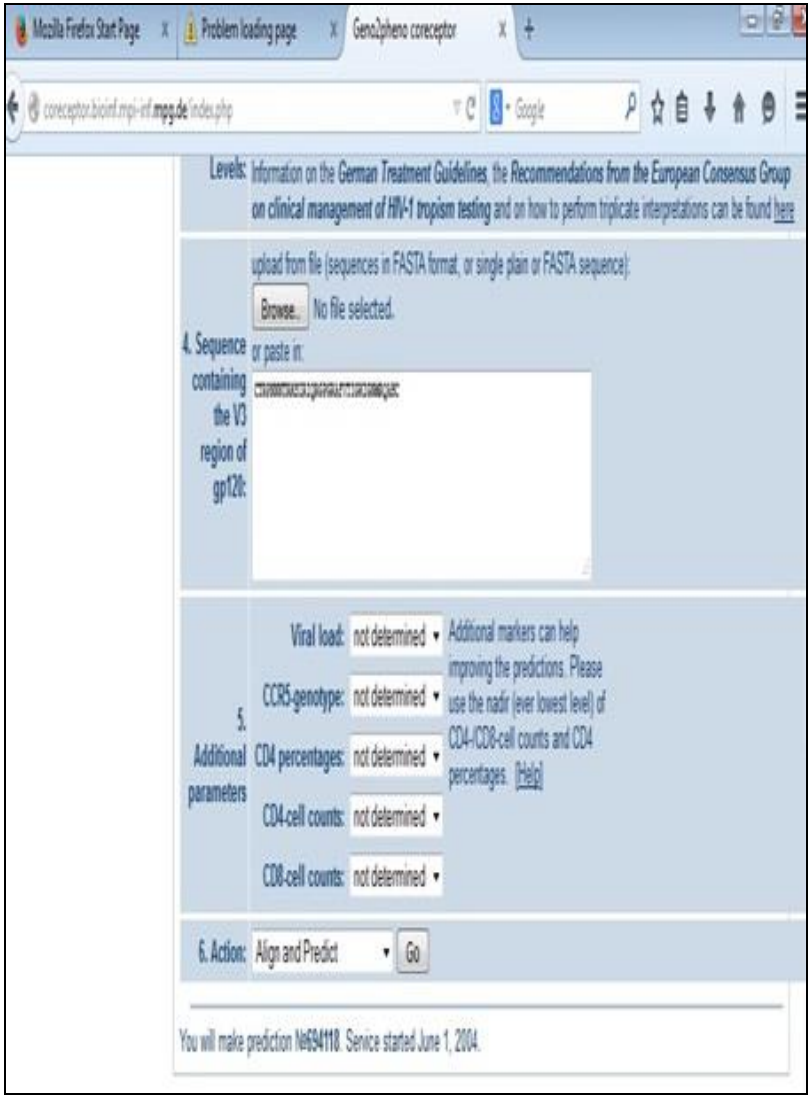
7. Cosic *et al* have been studied over 1000 proteins from 25 functional groups e.g. Oncogenes, Heat Shock Proteins, Protease Inhibitors, SV40 Enhancer, Tumor Necrosis Factors, etc using RRM [13].
8. Evolutionary roadmap for HIV [6] and Influenza [14].
9. Determine HCV protein sequences responsible for Interferon/Rabavirin therapy [15].
10. Ebola Virus/Endothelium interaction and its role in the Ebola Virus Disease process, prevention and therapy [16].
11. Designing of targerted bioactive peptide analogue with cytotoxic effects on tumor cells [17].
12. Determine possible mechanism by which Influenza vaccine prevents cardiovascular diseases

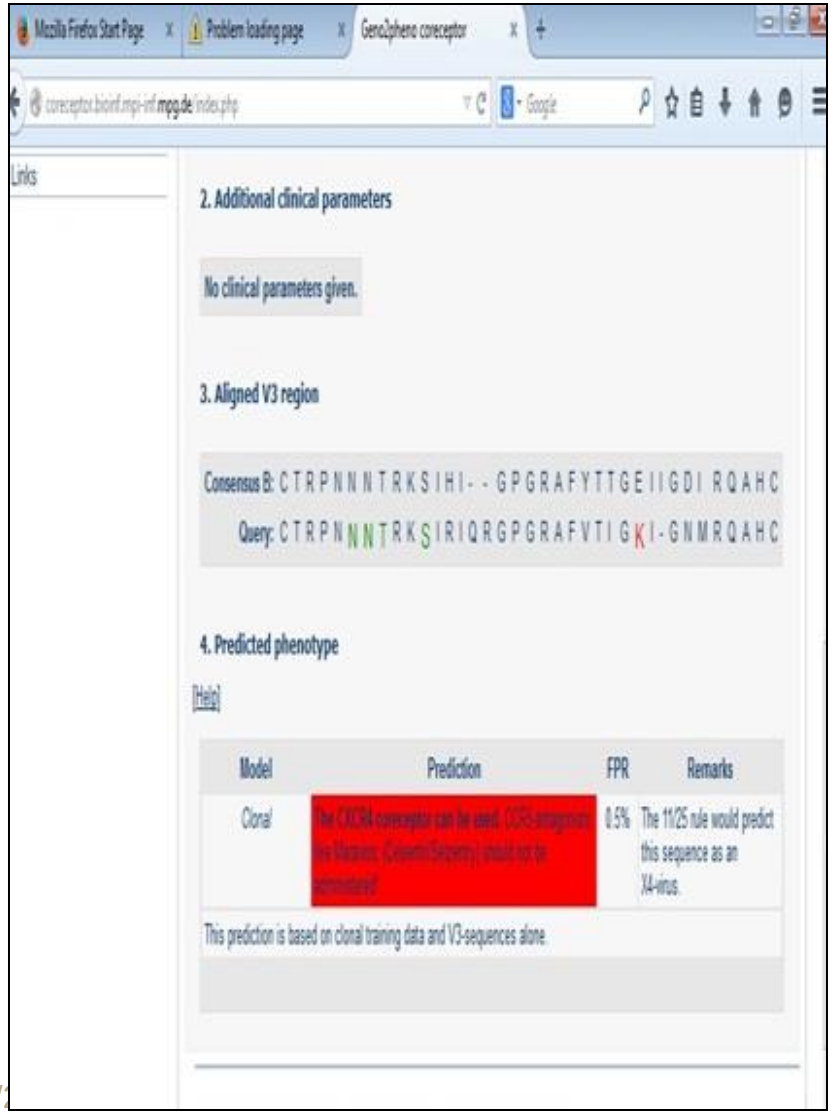
Non DSP-based, Sequence Information-oriented Biological data Transformation to Bio-functionalities: HIV-MN Isolate's Response to CCR5 Antagonist therapy [19]





Non DSP-based, Sequence Information-oriented Biological data Transformation to Bio-functionalities: HIV-HXB2 Isolate's Response to CCR5 Antagonist therapy





Discussion

- Biological data obtained from clinical experimentation (e.g. amino sequence alterations) can now be transformed into biological information being studied.
- As shown, several such vital DSP-based transformations have been made leading to:

i. a biomedical device: Computer-Aided Drug Resistance Calculator
ii. Innocentive winning solution for “Assessing Vaccine Potency” ID:9933477.
iii. Assessment of over 1000 proteins and the designing of several pharmacologically active drug and vaccines as well as candidates.
iv. Several other biological data transformation correlated with biological information obtained DSP and other procedures

Conclusion

1. The FEASIBLITY that translation of biological data e.g. sequence information into bio-functionalities is A GOOD REALIZATION.
2. These procedures are applicable to ALL biologically active agents by engaging the sequences of:

i. Peptidic and protein-based agents;

ii. genes/proteins encoding them;

iii. Their target proteins all employable.

3. Therefore, effective engagement of these simple and rational approaches will revolutionize our bench to clinic accomplishments.
4. The techniques will gainfully be employed in seeking therapeutic interventions.
5. Rational, computerized, informatics- and robotics-based procedures are the in-thing today. This

References:

1. Kelada SN, Shelton E, Kaufmann RB, et al. 2001. "delta-Aminolevulinic Acid Dehydratase Genotype and Lead Toxicity: A HuGE Review". Am J Epidemiol. 154:1–13.

2. Veljkovic V, Cosic I, Dimitrijevic B, Lalovic D .1985. "Is it possible to analyze dna and protein sequence by the method of digital signal processing". IEEE Trans Biomed Eng. 32(5):337-341.

3. Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell J, Molecular Cell Biology. W. H. Freeman and Company., 2000.

4. Jain E, Bairoch A, Duvaud S, et al. 2009. "Infrastructure for the life sciences: design and implementation of the uniprot website." BMC Bioinformatics. 10:136-154.

5. Tomii K, Kanehisa M. 1996. "Analysis of amino acid indices and mutation matrices for sequence comparison and structure prediction of proteins". Protein Engineering. 9(1):27-36.

6. Nwankwo N. 2012. Signal processing-based Bioinformatics methods for characterization and identification of Bio-functionalities of proteins. PhD Thesis (submitted). De Montfort University, Leicester, United Kingdom; available at www.openthesis.org

7. Hoj L, Kjaer J, Winther O, Cozzi-Lepri A, Lundgreen DJ, "In silico identification of physicochemical properties at mutating positions relevant to reducing susceptibility to Amprenavir," XVII International HIV Drug Resistance Workshop, vol. Poster No.113, 2008.

8. Smith SW, The Scientist and Engineer's Guide to Digital Signal Processing. California Technical Publishing, 2002.

9. Nwankwo N, Seker H. 2013. "HIV Progression to AIDS: Bioinformatics Approach to Determining the Mechanism of Action". Curr HIV Res. 11(1):30-42.

10. Nwankwo N, Seker H. 2011. "Digital Signal Processing Techniques: Calculating the Biological Functionalities of Proteins" J. Proteomics Bioinform 4: 260-268. doi:10.4172/jpb.1000199.
11. Nwankwo N, Seker H. 2011. "Digital Signal Processing Techniques: Calculating the Biological Functionalities of Proteins" J. Proteomics Bioinform 4: 260-268. doi:10.4172/jpb.1000199.
12. Nwankwo N, Seker H. 2010. "A signal processing-based bioinformatics approach to assessing drug resistance: human immunodeficiency virus as a case study". Conf. Proc. IEEE Eng Med Biol Soc. 2010:1836-1839.
13. Cosic I. 1994. “Macromolecular Bioactivity: Is It Resonant Interaction between Macromolecules?-Theory and Applications," IEEE Transactions on Biomedical Engineering” . 41(I):1101-1114.
14. Veljkovic V, Niman HL, Glisic S, Veljkovic N, Perovic V, Muller PC. 2009. “Identification of hemagglutinin structural domain and polymorphisms which may modulate swine h1n1 interactions with human receptor," BMC Structural Biology”. 9(62):1-11.
15 Veljkovic V, Glisic S, Veljkovic N, et al, et al. 2014.”Assessment of Hepatitis C Virus protein sequences with regard to interferon/ribavirin combination therapy response in patients with HCV genotype 1b”. Vaccine l 32(48): 6569-75.
16. Veljkovic V, Glisic S, Muller P, et al 2015. “Insilico analysis suggests interaction between Ebolavirus and the extracellular matrix”. Virology 6(135):1-11
17. Istivan TS, Pirogova E, Gan E et al. (2011) Biological Effects of a De Novo Designed Myxoma Virus Peptide Analogue: Evaluation of Cytotoxicity on Tumor Cells. PLoS ONE 6(9): e24809. doi:10.1371/journal.pone.0024809
18. Obermeier M, Ehret R, Berg T, *et al*, “Genotypic HIV-coreceptor tropism prediction with geno2pheno [CORECEPTOR]: differences depending on HIV-1 subtype“, Journal of the International AIDS Society, vol. 15(Suppl 4), pp. 1821-1824, 2012
19. Veljkovic V, et al. Influenza vaccine as prevention for cardiovascular diseases: Possible molecular mechanism. Vaccine (2014), http://dx.doi.org/10.1016/j.vaccine.2014.07.007

Thank you.

Comments/Questions