



## Bioinformatics Analysis of Serologic Proteins of Prostate Cancer Patients Separated by SDS-PAGE

Roha Talat<sup>1</sup>, Zunera Tanveer<sup>2</sup>, Irshad Ahmad<sup>3</sup>, Muhammad Shehzad Farooq<sup>4</sup>, Raheela Asmat<sup>5</sup>, Hafiz Muhammad Ali<sup>6</sup>, Iqra Ayub<sup>1</sup>, Asadullah<sup>1</sup> and Mohammad Zahid Mustafa<sup>1\*</sup>

<sup>1</sup> Center for Advanced Studies in Vaccinology and Biotechnology (CASVAB), University of Balochistan, Quetta, Pakistan.

<sup>2</sup> Department of Physiology, Bolan University of Medical and Health Sciences, Quetta, Pakistan.

<sup>3</sup> Institute of Basic Medical Sciences, Khyber Medical University, Peshawar Pakistan.

<sup>4</sup> Department of Biochemistry, Faisalabad Medical University, Faisalabad, Pakistan.

<sup>5</sup> Department of Gyneacology, Sandeman Provincial Hospital, Quetta, Pakistan.

<sup>6</sup> University College of Veterinary and Animal Sciences, The Islamia University of Bahawalpur, Pakistan.

\* Corresponding author: zunzah@gmail.com

### Abstract

One of the main goals of bioinformatics is to understand and analyze the 3D structure of proteins and the relationship between amino acid sequences. With the help of amino acid sequences, the protein structure can easily be predicted as proteins are essential in natural science research and they are linked with evolution, drug development, mutation and the occurrence of different diseases directly or indirectly. Biologists used bioinformatics tools to discover different diseases by knowing protein's structure and functions rather than using different technologies/experimental tools which can't completely explain proteins, its structure and role in several diseases. Prostate Cancer is the leading cause of cancer deaths in males worldwide, it's least common in Asia and more common in western countries. The study was conducted for the bioinformatics analysis of Prostate cancer proteins.

**Keywords:** Prostate Cancer, Bioinformatics Tools, SDS-PAGE.

### ARTICLE INFORMATION

Received: 12.08.2018

Revised: 01.11.2018

Accepted: 26.12.2018

DOI: 10.31580/pjmls.v1i1.940

© Readers Insight Publication

### INTRODUCTION

Bioinformatics is the multidisciplinary field of science that combines different fields i.e., computer science, mathematics, information technology, statistics and biology to study and understand all the biological data and deals with computational management of molecular biological material. As bioinformatics provides a huge data related to molecular biology, so currently many bioinformatics projects are dealing with functional and structural features of proteins and genes (1). The two main drawbacks of experimental tools are cost and time, therefore biologists use bioinformatics tools as an alternative to practice their research studies on protein functions and their structures (2). Prostate cancer (PCa) is the common cancers in men and the second most common cause of men mortality and morbidity. Age, heredity and race are known to be risk factors for prostate cancer. Furthermore, some exogenous factors, such as a reduced intake of vitamin E and high-fat diet, selenium, isoflavones or lignans, may stimulate the progression from the dormant prostate cancer type to the clinical type (3, 4). Proteomics and bioinformatics technology development increases the study of prostate cancer, considering a commanding tool in order to study the metastasis and pathogenesis of Prostate cancer. In particular, some tumor-associated proteins have been recognized in tissues and body fluids of several cancer (5).

The software designed to identify and analyze proteins plays an important role in investigating proteins from mass

spectrometry and two-dimensional gels. To identify proteins, the obtained information about proteins matches in conflict with protein database to describe a protein which is new or identified earlier and to analyze proteins, all the information present in protein databases can be used to calculate properties about a protein, which can be useful for its experimental study. There is a software known as ExPASy (Expert Protein Analysis System) available for protein analysis and identification, which predicts protein's isoelectric point, molecular weight, physicochemical properties, transmembrane domain, structure and function of proteins (6).

The ExPASy (Expert Protein Analysis System) server used for the analysis of proteins is a proteomics server of the Swiss Institute of Bioinformatics (SIB) which evaluates protein sequences, its functions and structures in association with European Bioinformatics Institute. This server consists of a group of tools which analyze and identifies different types of proteins and characterize by using primary structure, mass spectrometric data, profile searches and pattern, secondary structure, predicts and analyze tertiary sequence and structure, visualization, molecular modeling and quaternary structure analysis (7).

One of the protein analysis tools is ProtParam, present on the ExPASy server used to calculate/computation of several physicochemical properties of proteins with the help of protein sequence. ProtParam computed various parameters of a protein which includes amino acid composition, theoretical pI, molecular weight, extinction coefficient, instability index, grand



average of hydropathicity (GRAVY), estimated half-life, aliphatic index and atomic composition (6).

Another tool is ProtScale that calculates and signifies the protein profile created by amino acid scale on a selected protein. The amino acid scale is a numerical value allotted to each amino acid. Mostly secondary structure conformational parameters and hydrophobicity/hydrophilicity scales are used, however other different scales are also exist based on physical and chemical properties of amino acids (6). A new secondary structure prediction method for proteins from its amino acid sequences is PSIPRED (8).

## MATERIALS AND METHODS

### Sample Collection

All sampling was performed in civil hospital and BMC hospital Quetta. Total 50 blood samples of prostate cancer were taken from PCa patients and 10 were taken as control from healthy subjects. Samples were collected in sterile vacutainer tubes and sera were obtained by centrifugation at 12000 rpm for 5 minutes. Samples were further processed in CASVAB (University of Balochistan) Biotechnology Lab.

### Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS -PAGE)

SDS-PAGE was performed in order to separate proteins on gel matrix. Electrophoresis is a method of separating a complex mixture of proteins. A technique by which charged molecules are displaced with electric current through a gel matrix is known as Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE). This method is used to determine the composition of the protein subunits, to control the homogeneity of the protein sample and for use in other applications to purify the proteins. The protein migration rate during SDS-PAGE is known by the charge and pore size of the gel matrix, shape and size of the protein (9, 13, 14).

### Bioinformatics tools used

Bioinformatics and proteomics tools were applied in order to check the different properties of proteins. PSIPred tool was used to predict the physicochemical properties, ProtParm tools determine the secondary structure of proteins and ProtScale tool was used to find the transmembrane domains of the proteins.

Simple random sampling method will be used to collect data from teachers working in schools of Islamabad and Rawalpindi. A closed ended questionnaire used will be and measures will be adopted from the previous studies existing in literature.

## RESULTS AND DISCUSSION

From the serum samples of prostate cancer patients the protein profiling was performed using SDS-PAGE. Out of 50 serum samples five proteins (which may correspond to PSA, TRAP1, ANO7, CDC37, and Clusterin) were differently expressed comparing a pool of ten control sera. These proteins were further subjected to the bioinformatics analysis using different tools in order to check their transmembrane domains, secondary structure and physicochemical properties.

### Prediction of transmembrane domains in detected proteins using ProtScale Bioinformatics tool.

The knowledge of whether the detected proteins have trans-membrane domains tells us a lot about possible biological functions of the proteins, the physicochemical properties and the behavior during purification. According to observations of Wang et al., our detected proteins with molecular weights 32kDa., 75kDa., 99.8kDa, 44kDa, and 75kDa might correspond to Triosephosphate isomerase 1, Peroxiredoxin 4, Prohibitin, 60S acidic ribosomal protein P0, Cytoplasmic actin. We performed ProtScale analysis to predict their transmembrane domains. The accession numbers (accession numbers of the proteins were confirmed by UniProt data base) of the proteins and 'range' to see the hydrophobicity profile is given in table no II.

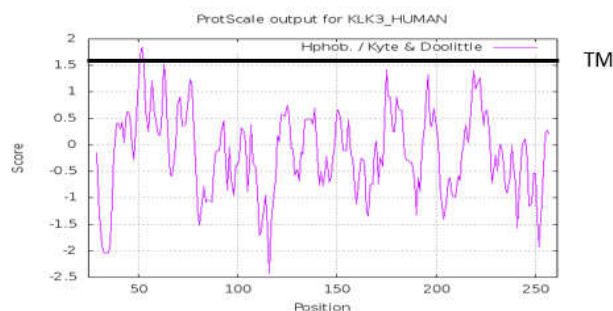
The computation was performed on the complete sequence of amino acids of detected proteins. Putative transmembrane (TM) domains were obtained using the scale Hphob. / Kyte & Doolittle (1982), hydrophobicity profile of proteins shown at y-axis and position of the amino acids shown at x-axis on the algorithms. Weights for window positions were 1... 19, using linear weight variation model (19 is optimal for transmembrane domains).

The results revealed that the transmembrane domains were found more in ANO7, few or less in PSA and CDC37, moderate in TRAP1 while null in clusterin. For finding transmembrane domains the values calculated should be above 1.5 (10). According to the Physicochemical properties the theoretical molecular weight of Prostate Specific Antigen, Tumor necrosis factor type 1 receptor-associated protein, Anoctamin-7, CDC37 and clusterin is 26.09, 73.5, 105.5, 16.4 and 24.1 kDa respectively, whereas the observed experimental molecular weight is (32, 75, 99.8, 44 and 75kDa). However the Theoretical PI of Prostate Specific Antigen, Tumor necrosis factor type 1 receptor-associated protein, Anoctamin-7, CDC37 and clusterin was (7.26, 6.13, 8.11, 5.24 and 6.27), whereas Physicochemical properties are shown in Table II.

Transmembrane domains compromise of nonpolar amino acid residue which go across the bilayer many times it also contain  $\alpha$  helices. The membrane-associated proteins and Transmembrane domains reside locations in the cell membrane where the protein concentration in two dimensions is excessive which employ a large impact on functions and clustering (11).

PSIPRED predicts protein secondary structure and gives to prediction server, GenTHREADER recognizes the structure and folds of a protein and MEMSAT 2 is a latest method hypothesizes topology of transmembrane proteins and protein structure (12). Whereas in the secondary structure the properties observed were helix, coil, disorders, membrane interaction, and membrane helix, polar, non-polar and hydrophobic regions.

### (a) PSA



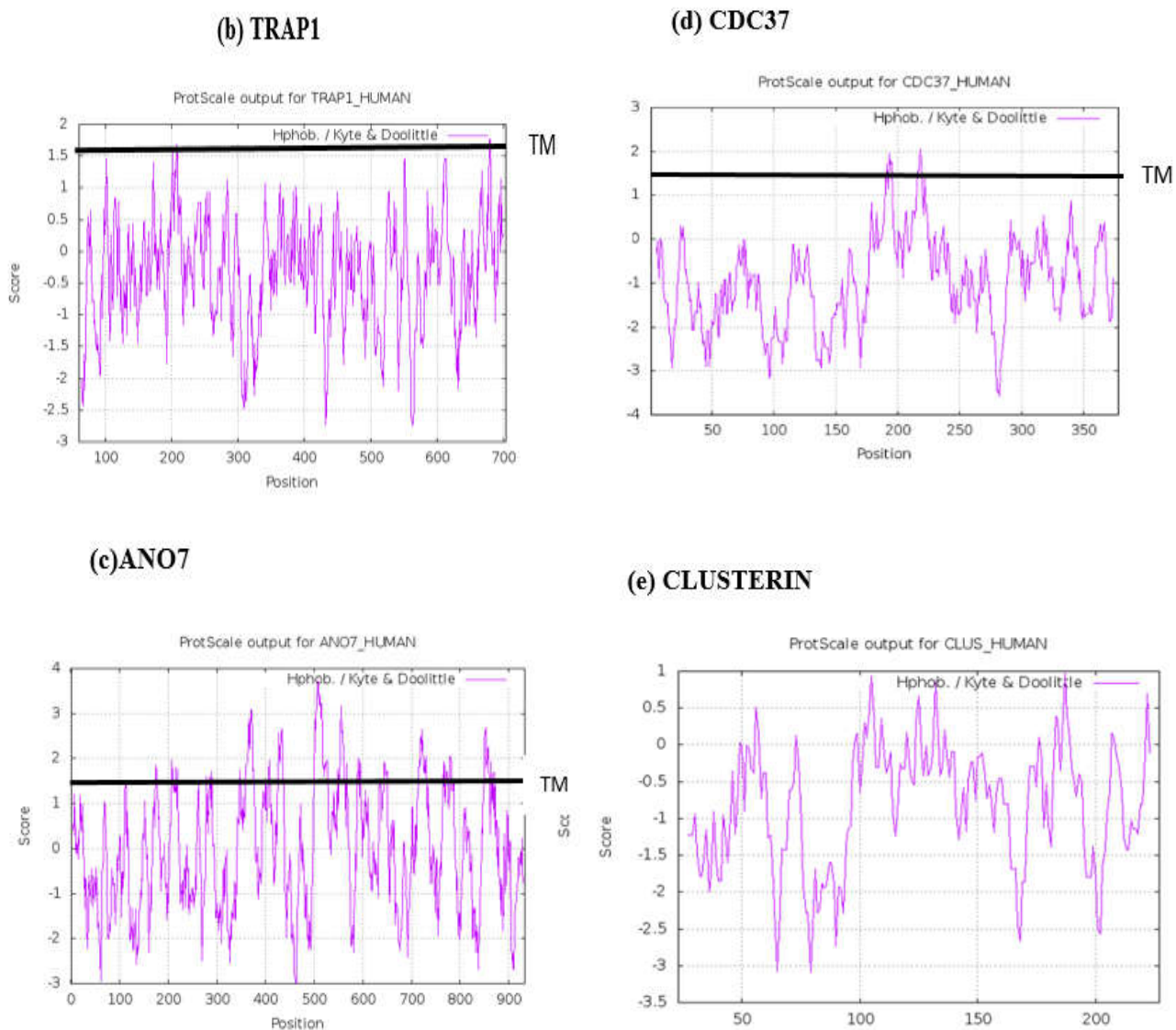
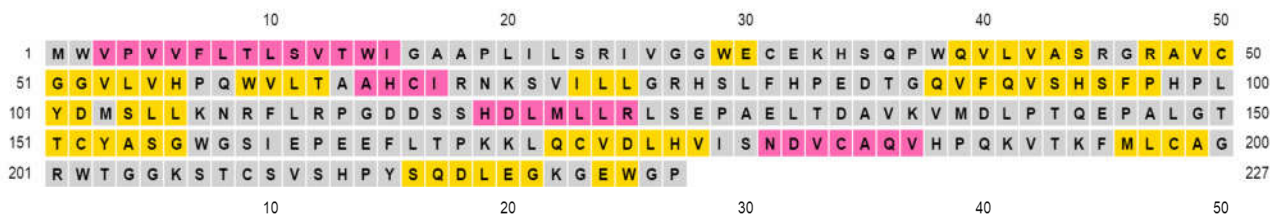


Fig. 1. Putative Transmembrane Domains of A. PSA, B. TRAP1, C. ANO7, D. CDC3 and E. Clusterin were obtained using the scale Hphob. / Kyte & Doolittle.

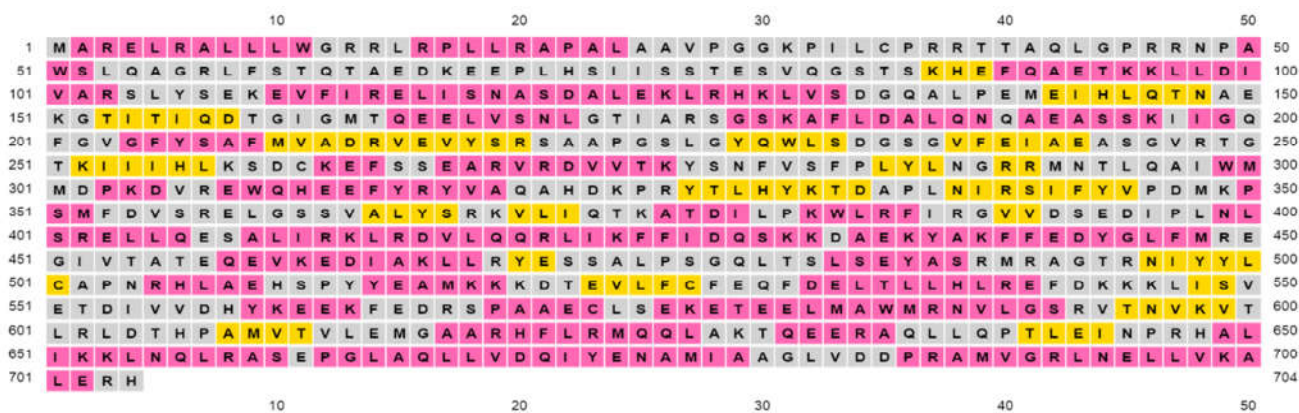
Table I. Physicochemical properties of detected proteins on the basis of their degree of hydrophobicity.

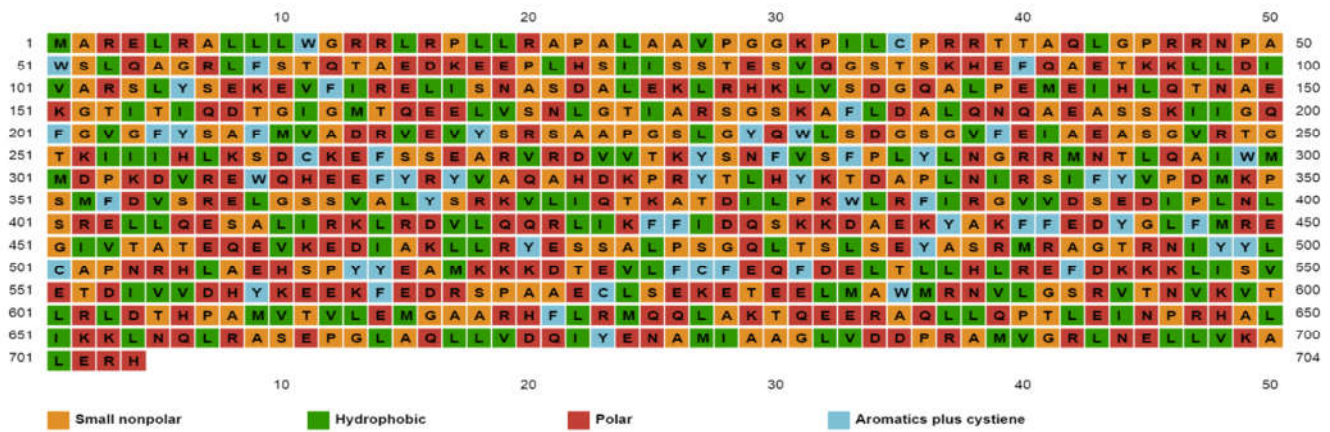
Name of Proteins	No. of Amino Acids	Th Molecular weight (kDa)	Experimental Molecular weight (kDa)	Theoretical PI	Ext. coefficient	Estimated half life	Instability index	Aliphatic index	Grand average of hydrophaticity
PSA	237	26.09	32	7.26	45085	20 hours	40.55	86.29	-0.140
TRAP1	645	73.5	75	6.13	60530	1.9 hours	40.69	90.14	-0.364
ANO7	933	105.5	99.8	8.11	191930	30 hours	45.78	90.36	-0.025
CDC37	137	40	44	5.42	23615	30 hours	49.47	53.36	-1.361
Clusterin	205	69	75	6.27	18700	1.1 hours	63.38	65.12	-0.843

(a) PSA



(b) TRAP1





(c) ANO7



(d) CDC37



(e) CLUSTERIN



Fig. 2. Secondary structure of protein a. PSA, b. TRAP1, c. ANO7, d. CDC3 and e. Clusterin are predicted using PSIPRED.



**Table II:** Hydrophobicity profile of detected Proteins for prediction of transmembrane domains

No.	Proteins	Accession Number	Sequence Length (Amino acids)	Range
1	Triosephosphate isomerase 1 (TPIS_HUMAN)	P60174	286	1-286
2	Peroxiredoxin 4 (PRDX4_HUMAN)	Q13162	234	38-271
3	Prohibitin	P35232	271	2-272
4	(PHB_HUMAN)	P05388	317	3-223
5	60S acidic ribosomal protein P0 (RLA0_HUMAN)	P60709	375	1-317

## CONCLUSION

The aim of the present study was to predict the physicochemical properties, transmembrane domains and secondary structure of proteins extracted by PCA patients using different bioinformatics tools. Biotechnologists are currently using bioinformatics tools for predicting 3D structure of proteins and finding relationship between amino acid sequences. In future advanced bioinformatics tools should be introduced for their better analysis.

## ACKNOWLEDGEMENT

I am very thankful to my supervisors Dr. Mohammad Zahid Mustafa and Dr. Asadullah, Department of Biotechnology and Molecular Biology, CASVAB (University of Balochistan) Quetta, for their insightful supervision and strong mentorship, and guidance throughout the research that helped me in every step.

## REFERENCES

- Vaisla DK, Rana J. Bioinformatics Tools. 2012. p. 158-75.
- Ghorbani M, Ghorbani F, Karimi H. Bioinformatics Tools for Protein Analysis 2016. 107-11 p.
- Antonelli JA, Jones LW, Banez LL, Thomas JA, Anderson K, Taylor LA, et al. Exercise and prostate cancer risk in a cohort of veterans undergoing prostate needle biopsy. The Journal of urology. 2009;182(5):2226-31.
- Ma RW, Chapman K. A systematic review of the effect of diet in prostate cancer prevention and treatment. Journal of human nutrition and dietetics : the official journal of the British Dietetic Association. 2009;22(3):187-99; quiz 200-2.
- Davalieva K, Kiprijanovska S, Komina S, Petrussevska G, Zografaska NC, Polenakovic M. Proteomics analysis of urine reveals acute phase response proteins as candidate diagnostic biomarkers for prostate cancer. Proteome science. 2015;13(1):2.
- Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 2003;31(13):3784-90.
- Haas J, Roth S, Arnold K, Kiefer F, Schmidt T, Bordoli L, et al. The Protein Model Portal--a comprehensive resource for protein structure and model information. Database (Oxford). 2013;2013:bat031.
- McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics. 2000;16(4):404-5.
- Manns JM. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) of Proteins. Current Protocols in Microbiology. 2011;22(1).
- Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. Journal of molecular biology. 1982;157(1):105-32.
- Grasberger B, Minton AP, DeLisi C, Metzger H. Interaction between proteins localized in membranes. Proceedings of the National Academy of Sciences of the United States of America. 1986;83(17):6258-62.
- Chen H, Gu F, Huang Z. Improved Chou-Fasman method for protein secondary structure prediction. BMC Bioinformatics. 2006;7 Suppl 4:S14.
- Bibi Zainab , AsadUllah , Zunera Tanveer, Muhammad Umar, Farah Sabeen Bugti , Muhammad Naeem, Taseer Ahmed Khan, Azam Jan, M. M. Tariq Kiani, Saad Ullah Jan, Ferhat Abbas and Mohammad Zahid Mustafa. Serologic detection of proteins associated with Rheumatoid Arthritis patients by Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE). International Journal of Biology and Biotechnology. 2017;14(1):11-15.
- Bugti FS, Mustafa MZ, Asmat TM, Asadullah, Zainab B, Achakzai SK, Rizwan M, Samad A, Taj MK, Rashid N. Detection of differentially expressed proteins in sera of patients with hepatocellular carcinoma by sodium dodecyl sulphate poly-acrylamide gel electrophoresis (SDS-PAGE). Pure and Applied Biology. 2017;6(1):140-145.

