ANALYSIS OF PLASMA PROTEINS ENCODED BY THE X-CHROMOSOME

[1]V.NIRANJAN, [1]R.MAHMOOD, [2]M.BALAMURUGAN, [2]NITESH KUMAR ROY, [3]SUBHO GHOSH

[1]Department of Biotechnology and Bioinformatics, Kuvempu University, Jnana Sahyadri, Shankaraghatta, Karnataka 577451
[2]Department of Biotechnology, The Oxford College of Science, J.P Nagar, 1th Phase, Bangalore-560078
[3]Department of Microbiology & Biotechnology, Al-Ameen Science College,Lalbagh Gate,Hosur Road, Bangalore-560027
E-mail: [1]niranjan.vidya@gmail.com, [2]niteshry@gmail.com

ABSTRACT

Plasma proteins are the most important proteins of human body and include enzymes, regulatory proteins, clotting factors and hormones. They perform diverse functions in the human body and play a cardinal role in its normal functioning. An aberration in structure, function or concentration of even one plasma protein may have fatal consequences. As a result, the study of plasma proteins becomes imperative to gain an insight into the reasons underlying various disorders which may be attributed to an anomaly in structure, function or variation in number of different types of plasma proteins. With the view to acquire information about the various proteins present in the blood plasma and to analyze the functions of all these proteins, a comprehensive analysis of plasma proteins was carried out. Each protein was analyzed individually on various parameters namely - the number of isoforms of that protein, the number of Single Nucleotide Polymorphisms (SNPs) present, the ectopic localization, post translational modifications(PTMs), proteolytic cleavage and expressions. Thus, a detailed and extensive analysis of each protein was carried out and this was followed by comparative analysis of all the plasma proteins. The proteins for which the locus was found to be on X-chromosome were studied for further analysis and studied for the presence of isoforms. SNPs were found in 47% of the X-chromosomal plasma proteins. PTM analysis revealed that only 6% of these proteins showed post translational modifications. No proteolytic cleavages were found. The ectopic localization of 62% of the plasma proteins is still not known, while a considerable percentage of proteins were found to be localized in the nucleus and cytoplasm. Only a small proportion of the proteins were found to be located on the plasma membrane and much lesser on the extracellular surface. The presence of a high percentage of X-chromosomal plasma proteins in nucleus and cytoplasm illustrates explicitly a deviation in the localization of plasma proteins. The data obtained indicates a leakage or malfunctioning of these proteins. Hence, we believe that the region of the X chromosome coding for these plasma proteins fall among the other important disease susceptible regions of the X chromosome.

Keywords: *Plasma proteins, SNPs, X-chromosomal plasma, plasma proteins, PTMs*

1. INTRODUCTION

An abnormality of the structure or function of the protein can lead to serious disorders and even death. For example a defect in one or more clotting factors can result in tremendous loss of blood and ultimately to the death of an individual. It is for this reason that we chose the plasma proteins coded by genes on the X chromosome for further study.

The central Dogma of molecular biology proposes the flow of genetic information from DNA to RNA to proteins. The ultimate expression of a gene is in the form of a protein. Proteins are of several types namely enzymes, hormones, transport proteins, structural proteins etc., while some proteins are confined to specific tissues or organs; some are found circulating in the body fluids e.g., blood. The active or functional form of a protein is its tertiary structure or in some cases, quaternary. The sequence of nucleotides in mRNA corresponds to the sequence of amino acids in the polypeptide chain of proteins. Any change in the sequence of nucleotides is reflected in the form of a change in the amino acid coded by that codon and hence a deviation in the protein encoded by that gene. This in turn, would lead to abnormal functioning of the tissue or organ in which that protein has expressed or functioned.

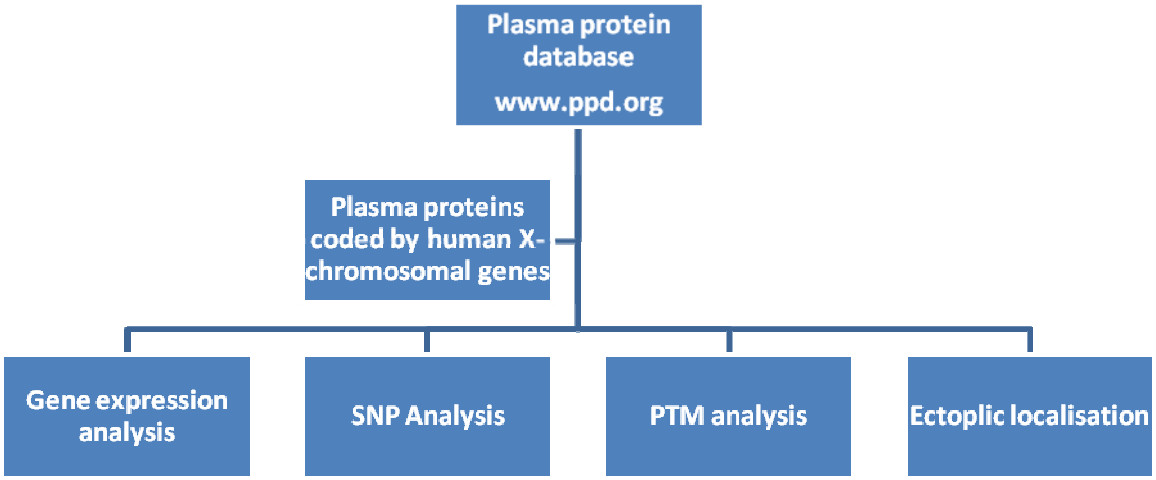
Hence a deviation from normal or original structure of a protein would result in a disorder and therefore the study of proteins become indispensable. (Anderson, N.L., and N.G. Anderson. 2002)

Plasma proteins have been early commercial targets for heterologous expression in cell systems. Much of the impetus for the design and implementation of such processes has come from the desire to supply a market with products that do not rely on the fractionation of donated blood plasma. The production of recombinant plasma proteins would avoid the hazards of blood-derived products, the most notable of which is viral contamination. Although, when processed correctly, blood-derived products are virtually free from transmitting viral infections, a perceived risk of contamination exists for the manufacturer, user and patient. These risks have been well-documented in the press over the past decade. (Goodey A R et al., 2000).

Plasma is the liquid matrix in blood. It is pale yellow colored fluid, of which 90-92% is water; a much smaller percentage of 8-10% is accounted for proteins. Plasma also contains a few inorganic ions and blood clotting factors (which help to form a blood clot on exposure to air and is instrumental in causing cessation of bleeding). Plasma comprises 20% of the extracellular fluids of body and it is very similar in composition to the interstitial fluid, the only difference being that the latter has a meager amount of proteins in it. The proteins found in plasma are called serum proteins and they play an important role in the body. Plasma proteins are mainly of 4 types namely albumins, (Mo Y *et al.,* 2007; BISCHOFF F and STAUFFER RD 1957)globulins, (Gross E *et al.,* 1992), fibrinogen and regulatory proteins. All but gamma globulins are synthesized in liver. (MILLER LL and BALE WF 1954) Besides these, plasma also contains small amount of lipoproteins. Each protein in the plasma has a different percentage of distribution and each protein has its own function. Some proteins are involved in auxiliary transport activity (Sorbara LR *et al*., 1997; Yang J and Yang FY 1992), cell adhesion activity (Tsuiji H *et al*., 2007; Brito C *et al*., 2007; Kumar S *et al*., 2007; Zhang JJ *et al*., 2007; Tsuiji H *et al*., 2007; Leik CE *et al*., 2006), DNA binding (Angelini F *et al*., 2007; Cinova J *et al*., 2007) ,DNA repair (Alrefai RH *et al*., 2007; Yang SY *et al*., 2006; Sridharan DM *et al*., 2006), Guanyl nucleotide exchange factor activity (Gross E *et al*., 1992), hydrolase activity while some other participate in the functions such as lipid kinase proteases (Bu S *et al*., 2006; Odintsova ES *et al*., 2005; Yang SA *et al*., 2004; Gratacap MP *et al*., 1998; Geltz NR and Augustine JA 1998), RNA binding (Takemoto T *et al*., 2007; Homann M *et al*., 2006; Malonga H *et al*., 2006), transcription factor activity, and ubiquitin specific protease activity. Thus the plasma proteins are necessary for maintenance of normal functioning of the body and they represent an important part of the human proteome (Powanda MC and Moyer ED 1981). Plasma proteins are represented by different percentage composition as mentioned below. Albumins constitute about 60% plasma proteins, 35% is made up by globulins, a much smaller proportion of plasma proteins are contributed by fibrinogens, while regulatory protein like enzymes and hormones account for less than one percent. Although recent proteomics research efforts focus primarily on determining the overall number of proteins present in the plasma. It is equally important to delineate protein variation among individual, for these signal the onset of disease and can be used as biological markers in diagnostics.

2. MATERIALS & METHODS

A slip by slip approach was followed for a detailed analysis of X- chromosome plasma proteins. The composition of blood plasma proteins was obtained using the plasma proteins database (www.plasmaproteindatabase.org). The percentages of different types of plasma proteins could be figured out with the help of this database and a comparative study could be done. Each protein was individually analyzed to gain an insight into its ectopic localization, number of isoforms, no of SNPs and its post transcriptional modifications. This enabled functional annotation of each plasma protein. The plasma protein analysis results were presented in the form of graphs and excel sheets which helped in drawing correct results. The study was expedited with the use of Microsoft excel filters which segregated the proteins based on their ectopic localization and therefore made the analysis simpler. The proteins which were found to have locus on X-chromosome were opted for further analysis. These were then studied for the presence of isoforms. Graphs were plotted and the results were interpreted.



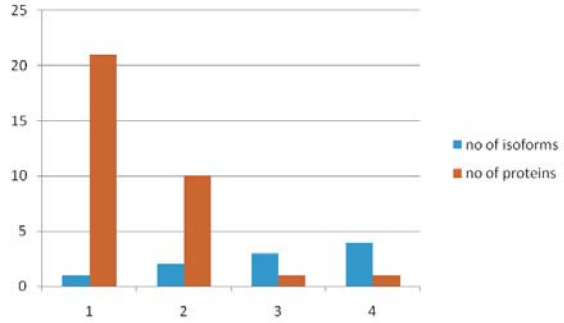
Methodology for Analysis of Plasma Proteins for Human X-chromosome

3. RESULTS & DISCUSSION

The plasma proteins as declared by the Human Protein Organization (HuPO) were shortlisted based on occurrence in the X-chromosome, further research carried on various parameters which are shown below.

Plasma protein declared by HUPO	Number of isoforms
APXL protein	1
Nucleosome assembly protein1 like 3	1
immunoglobulin binding protein1	1
melanoma associated antigen B4	1
sushi repeat containing protein SRPX	2
plexin B3	1
G-antigen family D2	4
MAGEAII	2
Praline rich gamma carboxy glutamic acid protein1	1
Adlican	1
DKFZP564B147 protein	1
DOCKII	1
Hypothetical protein FLJ11362	3
FLJ12525	2
Hypothetical protein FLJ12969	1
FLJ22965	1
FLJ33516	2
JMII protein	2
JM5 protein	1
Hypothetical protein KIAA1318	1
Leucine rich protein & calponin homology CH domain containing2	1
WD repeat domain 40B	1
melanoma antigen family C3	2
7CEAL8 protein	1
NF kappa B activating protein	1
Paraneoplastic antigen MA3	2
Paraneoplastic antigen like5	1
UTP u3 small nucleolar ribonucleoprotien homolog A	2
Hypothetical protein FLJ32867	1
plexin-A3	1
Dystrophin related protien2	1

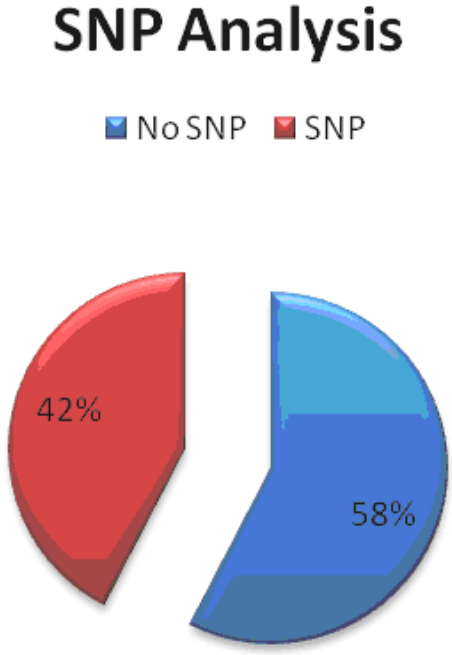
4. NUMBER OF ISOFORMS IN PLASMA PROTEINS OF HUMAN X-CHROMOSOME



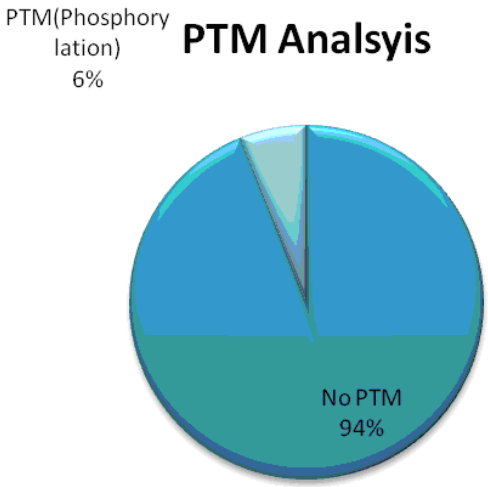
Isoform Analysis of X-chromosomal Plasma Proteins

It illustrates the number of isoforms for various plasma proteins. These isoforms were found out and their combined result was obtained in the form of bar graph. The study revealed that majority of plasma proteins did not have a large number of isoforms. 20-22% plasma proteins were found to have only 1-2 isoforms while only a much smaller percentage of proteins i.e., 4-5% had 4-5 isoforms. Thus it could be inferred from the analysis that a high proportion of plasma proteins did not have many isoforms and only a few had 4-5 isoforms.

The plasma proteins that were found to have their locus on X- chromosome we extensively analyzed using few more parameters. Single nucleotide polymorphism analysis for all the X-chromosome plasma proteins revealed that 42% of these proteins consisted of SNPs. SNPs were not found in the remaining 58% proteins. This analysis explicitly indicates the presence of SNPs in the X-chromosomal proteins which highlights the possibility that mutations might occur in this region of X-chromosome. These mutations in turn may lead to disorders thus showing this region of the X chromosome to be a highly disease susceptible region.



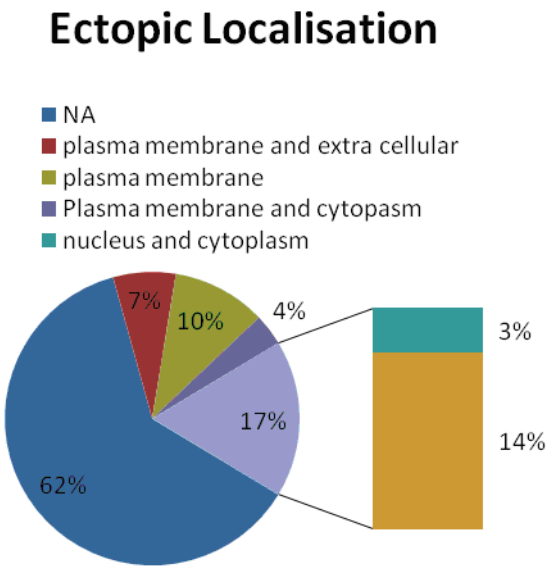
Percentage of SNPs in Plasma Proteins



Percentage Of Ptms In Plasma Proteins

The information coded in a gene (a stretch of DNA) is transferred to mRNA which is translated into proteins. After translation, the post translational modification of amino acids extends the range of functions of the protein by attaching to it other biochemical functional groups such as acetate, phosphate, various lipids and carbohydrates, by changing the chemical nature of an amino acid (e.g. citrullination) or by making structural changes, like the formation

of disulfide bridges. Post translational modification might lead to a change in the type of protein, which is different from the one actually coded by the gene. PTM analysis of X-chromosome plasma proteins showed that 6% of these proteins undergo PTMs. The only post transcriptional modification found in these proteins was phosphorylation. There were no indications of any proteolytic cleavage.



Ectopic Localization

Each protein was studied individually for its location and an integrated result was obtained taking all the proteins into consideration. The result revealed that 17% proteins were found to be located in nucleus and cytoplasm, while 10% were found to have plasma membrane localization). Plasma membrane and extracellular space accounted only for 7% of plasma proteins. The localization of 62% protein is yet unknown. The most significant finding of this analysis is that, localization of proteins in nucleus and in cytoplasm is quite unusual for plasma proteins and hence, it images a deviation from their normal localization.

It could be clearly inferred from this analysis that there could be leakage of the plasma proteins or possible malfunctioning. Hence, these proteins would be of greater interest which stands for the possibility that this region of X-chromosome might be responsible for causing diseases.

CONCLUSION

The plasma proteins for which the locus was found to be on X-chromosome were considered for further analysis and studied for the presence of isoforms. It was discovered that 47% of X-chromosome plasma proteins revealed the presence of SNPs. PTM analysis revealed that only 6% of these proteins showed post transcriptional modification. No proteolytic cleavages were found. The localization of 62% plasma proteins is still not known, while a considerable percentage of proteins were found to be in the nucleus and cytoplasm. A relatively small proportion of protein was found to be located on plasma membrane and still lesser on extracellular surface. A high percentage of X-chromosome plasma protein in nucleus and cytoplasm illustrates explicitly a deviation in the localization of plasma proteins. The data obtained indicates a leakage or malfunctioning of these proteins. Since plasma proteins were found to be involved in many diseases, these proteins would be of great interest which lends support in the consideration of this region of X-chromosome as a disease causing region.

REFERENCES

[1]. Mo Y et.al (2007). Human serum albumin nanoparticles for efficient delivery of Cu, Zn superoxide dismutase gene Mol Vis. May 23;13:746-57.

[2]. BISCHOFF F and STAUFFER RD (1957). Orientation of circulating human estrogens by albumin Am J Physiol. Nov;191(2):313-8.

[3]. Gross E et.al (1992). Anti-Cdc25 antibodies inhibit guanyl nucleotide-dependent adenylyl cyclase of Saccharomyces cerevisiae and cross-react with a 150-kilodalton mammalian protein. Mol Cell Biol. Jun;12(6):2653-61.

[4]. MILLER LL, and BALE WF(1954). Synthesis of all plasma protein fractions except gamma globulins by the liver; the use of zone electrophoresis and lysine-epsilon-C14 to define the plasma proteins synthesized by the isolated perfused liver. J Exp Med.Feb;99(2):125-32.

[5]. Sorbara LR et.al(1997). thrombin-induced translocation of GLUT3 glucose transporters in human plateletsBiochem J. Dec 1;328 (Pt 2):511-6.

[6]. Yang J and Yang FY (1992). he effect of selenium on the function of the anion transporter (Band 3) of erythrocyte membranes,Biofactors. Dec;4(1):29-32.

[7]. Tsuiji H et.al(2007). Cadherin conformations associated with dimerization and adhesion J Biol Chem. Apr 27;282(17):12871-82. Epub 2007 Mar 8.

[8]. Brito C et.al(2007). Increased levels of fucosyltransferase IX and carbohydrate Lewis(x) adhesion determinant in human NT2N neurons J Neurosci Res. May 1;85(6):1260-70.

[9]. Kumar S et.al(2007). Isoliquiritigenin inhibits IkappaB kinase activity and ROS generation to block TNF-alpha induced expression of cell adhesion molecules on human endothelial cells Biochem Pharmacol.May 15;73(10):1602-12. Epub 2007 Jan 13.

[10]. Zhang JJ et.al(2007). PECAM-1 modulates thrombin-induced tissue factor expression on endothelial cells J Cell Physiol. Feb;210(2):527-37.

[11]. Leik CE et.al(2006). Effect of pharmacologic plasminogen activator inhibitor-1 inhibition on cell motility and tumor angiogenesis. J Thromb Haemost. Dec;4(12):2710-5. Epub Oct 2.

[12]. Angelini F et.al(2007). Nuclear factor kappaB activity is increased in peripheral blood mononuclear cells of children affected by atopic and non-atopic eczema Int J Immunopathol Pharmacol. Jan-Mar;20(1):59-67.

[13]. Cinova J et.al(2007). Gliadin peptides activate blood monocytes from patients with celiac disease. J Clin Immunol. Mar;27(2):201-9. Epub Jan 27.

[14]. Alrefai RH et.al(2007). Nucleotide excision repair in an immunoglobulin variable gene is less efficient than in a housekeeping gene. Mol Immunol. Apr;44(11):2800-5. Epub Mar 1.

[15]. Yang SY et.al(2006). Control of gene conversion and somatic hypermutation by immunoglobulin promoter and enhancer sequences. J Exp Med. Dec 25;203(13):2919-28. Epub Dec 18.

[16]. Sridharan DM et.al (2006). alphaII-Spectrin interacts with five groups of functionally important proteins in the nucleus. Cell Biol Int. Nov;30(11):866-78. Epub Jun 30.

[17]. Bu S et.al (2006). Dihydrosphingosine 1-phosphate stimulates MMP1 gene expression via activation of ERK1/2-Ets1 pathway in human fibroblasts FASEB J. 2 Jan;20(1):184-6. Epub Nov 8.

[18]. Odintsova ES et.al (2005). Casein-hydrolyzing activity of sIgA antibodies from human milk J Mol Recognit. Sep-Oct;18(5):413-21.

[19]. Yang SA et.al (2004). Rho and Rho-kinase mediate thrombin-induced phosphatidylinositol 4-phosphate 5-kinase trafficking in platelets. J Biol Chem. Oct 1;279(40):42331-6. Epub Jul 23.

[20]. Gratacap MP et.al (1998). Phosphatidylinositol 3,4,5-trisphosphate-dependent stimulation of phospholipase C-gamma2 is an early key event in FcgammaRIIA-mediated activation of human platelets. J Biol Chem. Sep 18;273(38):24314-21.

[21]. Geltz NR and Augustine JA (1998). The p85 and p110 subunits of phosphatidylinositol 3-kinase-alpha are substrates, in vitro, for a constitutively associated protein tyrosine kinase in platelets. Blood. Feb 1;91(3):930-9.

[22]. Takemoto T et.al (2007). RBMX is a novel hepatic transcriptional regulator of SREBP-1c gene response to high-fructose diet ,FEBS Lett. Jan 23;581(2):218-22. Epub Dec 14.

[23]. Homann M et.al (2006). serum-stable RNA aptamers to an invariant surface domain of live African trypanosomes. Comb Chem High Throughput Screen. Aug;9(7):491-9.

[24]. Malonga H et.al (2006). transfer RNA binding to human serum albumin: a model for protein-RNA interaction. DNA Cell Biol. Jul;25(7):393-8.

[25]. Powanda MC and Moyer ED (1981). Plasma proteins and wound healing. Surg Gynecol Obstet. Nov;153(5):749-55.