Relationship between Physico-chemical Parameters and Phylogenetics Study of Human Low Density Lipoprotein Receptor-Related Protein (LRP)

Olugbenga Samson Onile[1,2*]

[1]*Cell Biology and Genetics Unit, Department of Zoology, University of Ibadan, Nigeria,*
[2]*Biotechnology Unit, Department of Biological Sciences, Elizade University Ilara-Mokin, Ondo State, Nigeria*

Author's contribution

This work was single handedly carried out by author OSO. All protocols, analysis, literature search and draft of the manuscript regarding this study were only conducted by author OSO. Author OSO read and approved the final manuscript.

Original Research Article

ABSTRACT

In this study, 11members of human low density lipoprotein receptor-related protein (LRP) sequences was retrieved from UniProtKB/ SWISS-PROT protein database and was analyzed for information about their structural, functional and phylogenetic features. This was achieved by using many established biocomputational tools which was available at their latest version. This study shows that LRP 12 and 3 are closely related with LRP8 being their nearest neighbor. In all, it was observed that there were very low possession of certain essential amino acid like glycine, proline and a very high aliphatic in all the LRP family. Considering the evolutionary history, functional domains, high aliphatic index, overall proportion of glycine and proline and the established role of one (LRP8) of this closely related LRP in diseases, it is thus predicted that the other closely related LRP3 and 12 molecules may be important candidate in investigating the aetiopathology of Myocardial infarction1diseases or other heart related disorder.

Keywords: Low density lipoprotein receptor-related protein (LRP); biocomputational tools; domains; phylogenetics.

_____

*Corresponding author: Email: onileg@yahoo.com; olugbenga.onile@elizadeuniversity.edu.ng;

1. INTRODUCTION

The low density lipoprotein receptor-related protein (LRP) is a member of an evolutionarily ancient and highly conserved gene family, the LDL receptor (LDLR) family [1].This LDLR family includes seven family members that are closely related and include the LDL receptor, very-low-density lipoprotein (VLDL) receptor, apoE receptor 2, multiple epidermal growth factor-like domains 7 (MEGF7), glycoprotein 330 (gp330/megalin/LRP2), LRP1 and LRP1B. The family also includes members that are more distantly related, such as LRP12, LRP5, LRP6 and SorLa/ LRP11. Like other members of the LDL receptor family, the modular structures within LRP include cysteine-rich complement-type repeats, EGF repeats, β-propeller domains, a transmembrane domain and a cytoplasmic domain [2]. It is understood that there are well over 40 soluble ligands that bind to LRP [3] hence the name multiligands receptor. LRP, a 600 kDa which is one of the larger receptors in the LDLR family, is known as a receptor for chylomicron remnants, taking dietary cholesterol from the gut to the liver *via* lipoproteins. Gene knock out studies have confirmed that LRP is required for embryonic development in the mouse [4]. LRP's ability to remove proteinase and proteinase inhibitor complexes has also raised the opinion that LRP may be a multifunctional scavenger receptor [1,5].

LRP is expressed in numerous cell types including fibroblasts, hepatocytes, adipocytes, macrophages and central nervous system (CNS) cells. In the normal human brain, LRP is expressed in the pyramidal neurons. Studies have shown LRP to be expressed in microglia and astrocytes under some pathological conditions [1]. Role for LRP, and other members of the LDLR family, in synaptic transmission in the adult brain (in disorder like Alzheimer's disease) have been reported [1,6,7]. LRP1 in the liver plays an important role in facilitating the plasma removal of a number of molecules, including enzymes and cofactors involved in blood coagulation and fibrinolysis, enzyme-inhibitor complexes, and certain lipoprotein particles [2]. Its function in the liver is important for normal homeostasis of these pathways. Tissue-selective gene deletion studies in vascular smooth muscle cells, hepatocytes and macrophages have all revealed a protective role for LRP1 in the development of atherosclerosis. This study tends to characterize sequences of human LRP by using biocomputational tools, as it intend to provide insight to the various protein attributes by analyzing for the structural, functional, phylogenetic and physico-chemical properties of sequences of LRP's in Human. The outcome of this study will further help biologists in carryout investigations on this essential protein.

2. METHODOLOGY

2.1 Retrieval of Human Low Density Lipoprotein Receptor Related Protein (LRP) Sequences

11 members of human LRP family reported till date were derived from UniProtKB/ SWISS-PROT, a curated protein database (http://expasy.org/sprot/) in FASTA format with the help of the accession number provided for each LRP sequence (http://www.uniprot.org/) [8].Total information about the origin, attributes, annotation, ontologies, binary interactions and sequence of proteins was found in this knowledgebase.

2.2 Physico-chemical and Structural Characterization of LRP Family

Various features including number of amino acids, molecular weight, theoretical isoelectric point (pI), amino acid composition (%), number of positively (Arg+Lys) and negatively charged (Asp+Glu) residues, extinction co-efficient, instability index, aliphatic index and Grand Average of Hydropathicity (GRAVY) were computed using ExPASy's Prot Param tool using the protein sequence in FASTA format as the input data type (http://expasy.org/tools/protparam.html) [9]. Other physico-chemical features including hydropathy, transmembrane tendency were predicted using Kyte Doolittle Hydropathy Plot (where transmembrane regions were also identified by peaks with scores greater than 1.8 using a window size of 19.) [10]. Secondary structure of some selected member of LRP family was predicted using Phyre2 protein homology/analogy Recognition Engine V2.0 [11] and was later remodel using The PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC.

2.3 Phylogenetic Classification of Collagen Family

The phylogenetic analysis for human low density lipoprotein receptor related protein (LRP) sequences were conducted using MEGA version 5 [12]. The evolutionary history was inferred using the Neighbor-Joining (NJ) method [13] while the evolutionary distances were computed using the Poisson correction method [14].

2.4 Sequence Analysis of Functional Domains in Human LRP

Domain scan tool Pfam 27.0 (last updated March 2013, 14831 families) was used to scan and identify all functional domains and their location in all the sequences of LRP family. The Pfam database is a large collection of protein families, each represented by multiple sequence alignments and hidden Markov models (HMMs) [15]

3. RESULTS AND DISCUSSION

A multi-Align tool was employed to perform protein sequences alignment, where a decreased sequence similarity was observed while considering increase number of input sequences. It was therefore observed that the human lipoprotein receptor related protein (LRP) showed similarities and differences at several positions in all the 11members. Computation of amino acid composition of each human (LRP) sequences using Protparam tool (see Table 2) indicated very high percentage of leucine, glycine and Serine (with values ranges from 6-12.1%) in almost all the proteins as compare with other amino acids, this invariably may suggested that these amino acids could have a major role as nucleation centers in the folding and evolution of the LRP family [16], while high percentage of these amino acids could account for the stability of the LRP triple helical structure, since incorporation of large amino acids can cause steric hinderance [17]. Also, increase in the proportion of amino acids like asparagine and serine as seen in LRP4, LRP5, LRP6 and LRP10 may aid in processes like cell-cell adhesion, migration and cell signaling properties of these proteins. Their abundance will encourage further protein process/modification to help in protein function actualization, example include O and N-linked glycosylation which involve serine and asparagines respectively. The carbohydrate groups on the cell surface are important components of the recognition sites of membrane receptor, such as those involve in binding extracellular signal molecules, in antibody-antigen reactions and in intercellular adhesion to form tissues [18]. In all, the outcome of this study shows that amino acids like

glycine and proline are considerably low as compare to other study (17). Analysis of instability index (see Table 3) in this study, classified 7 of the proteins as unstable while LRP 1, 1B, 2 and 5 were all identified as stable (instability index<40).Instability may be suggested to aids the LRPs family to attain large amplitude vibrations, which is required for their proper function while considering their multiligand properties, and yet still maintain their native fold [19]. The pI values for all the 11 LRPs were found to lie in the acidic range. GRAVY values which signifies LRP-water interaction were observed within the range of -0.196-0.511. LRP5 is identified as the most thermo-stable of the protein family, with highest aliphatic index of 82.67 as this described the relative volume occupied by aliphatic chain the protein, with the lowest LRP1 (64.12). The hydrophobic, hydrophilic and transmembraneregion of each LRP sequences were already illustrated in Figs. 2-12 using Kyte-doolittlehydropathy plots, where region with positive score on the graph represent the hydrophobic region of each LRP sequence while region with negative scores represent the hydrophilic region. A score of 4.6 is the most hydrophobic and a score of-4.6 is the most hydrophilic. With Kyte-doolittlehydropathy plots, it was observed that most of these proteins are single spanning membrane proteins (see Figs. 2-12). The functional role of human low density lipoprotein receptor related protein (LRP) was analyzed (see Table 1). Several functional domains were identified in different LRP sequences at different regions and in differing quantities (see Table 3). It was discovered that LRP10 for example, was having two of Low-density lipoprotein receptor domain class A (Ldl_recept_a) at alignment region 139-174 and 398-433 respectively and a single CUB domain at 192-302 alignment region. All the 11 LRP family sequences were found to have Low-density lipoprotein receptor domain class A (Ldl_recept_a). Presence of this domain in all the LRPs can be correlated to their known function in removal of excess cholesterol in the blood via endocytosis [20] and their active role in cell signaling pathways between specialized cells [21]. MANEC domain was only found in LRP11 and the peculiarity of this domain in LRP11 could foster biologist understanding about this less characterized protein.

The 11 human LRP sequences were aligned based on sequences homology and phylogram was constructed with a distanced based neighbour-joining (NJ) methods to establish evolutionary regions in the LRP family (see Fig. 1). Several clusters with close relationships were identified including LRP 1 and 2, 10 and 4, 12 and 3, 5 and 11 and 6 and 1B, while 12 and 3 are relatively closely to 8. LRPs could possible then be analyze together, owing to the fact that most of the members of this family share close evolutionary history. Detailed information on LRP 1B and its possible role in diseases could better be retrieve with thorough investigation of LRP6, whose functions and role in diseases like Coronary artery diseases have been identified, also proper studying of LRP5 while considering its functions, tissues speicificity and their role in the aetiopathology of diseases like Vitreoretinopathy, exudative 4 (EVR4), Osteoporosis (OSTEOP), Osteoporosis-pseudoglioma syndrome (OPPG), High bone mass trait (HBM), Endosteal hyperostosis, Worth type (WENHY), Osteopetrosis, autosomal dominant 1 (OPTA1),will assist in providing vital information about the less known LRP 11, owing to their structural similarities and evolutionary history. The consideration of physico-chemical parameters (instability, high aliphatic index, pI at acidic range and hydropathy index) of the three closely related family of LRP (3, 12 and 8), when compared with the phylogram, further revealed certain level of similarity among the identified members of these protein family.

Table 1. Functional properties of human LRP in diseases

LRP NAME/ ACCESSION No.	PROTEIN NAME	FUNCTION	INVOLVEMENT IN DISEASES	Tissue specificity
LRP1_HUMAN/Q07954	Prolow-density lipoprotein receptor-related protein 1	Involved in endocytosis and in phagocytosis of apoptotic cells. Required for early embryonic development. Involved in cellular lipid homeostasis. Involved in the plasma clearance of chylomicron remnants. May modulate cellular events, such as APP metabolism, kinase-dependent intracellular signaling, neuronal calcium signaling as well as neurotransmission.	Deletion of hepatic LRP accelerates the development of atherosclerosis; regulate the progression of Alzheimer's Disease (AD).	Most abundant in liver, brain and lung
LRP1B_HUMAN Q9NZR2	Low-density lipoprotein receptor-related protein 1B	Potential cell surface proteins that bind and internalize ligands in the process of receptor-mediated endocytosis.	_	Thyroid gland and in salivary gland, as well as in adult and fetal brain.
LRP2_HUMAN P98164	Low-density lipoprotein receptor-related protein 2	Acts together with cubilin to mediate HDL endocytosis,. May participate in regulation of parathyroid-hormone and para-thyroid-hormone-related protein release.	Donnai-Barrow syndrome (DBS), FOAR syndrome	Absorptive epithelia, including renal proximal tubules
LRP3_HUMAN O75074	Low-density lipoprotein receptor-related protein 3	Probable receptor, which may be involved in the internalization of lipophilic molecules and/or signal transduction. Its precise role is however unclear, since it does not bind to very low density lipoprotein (VLDL) or to LRPAP1 in vitro.	_	Skeletal muscle and ovary. Expressed at intermediate level in heart, brain, liver, pancreas, prostate and small intestine. Weakly expressed in testis, colon and leukocyte
LRP4_HUMAN O75096	Low-density lipoprotein receptor-related protein 4	Mediates SOST-dependent inhibition of bone formation. Functions as a specific facilitator of SOST-mediated inhibition of Wnt signaling, function as a cell surface endocytic receptor binding and internalizing extracellular ligands for degradation by lysosomes.	Cenani-Lenz syndactyly syndrome (CLSS), Sclerosteosis 2 (SOST2)	Bone; osteoblasts and osteocytes, Expressed in several regions of the brain

Table 1 Continued ……..

LRP5_HUMAN O75197	Low-density lipoprotein receptor-related protein 5	Cell-surface coreceptor of Wnt/beta-catenin signaling, which plays a pivotal role in bone formation.	Vitreoretinopathy, exudative 4 (EVR4), Osteoporosis (OSTEOP), Endosteal hyperostosis, Worth type (WENHY)	Liver and in aorta
LRP6_HUMAN O75581	Low-density lipoprotein receptor-related protein 6	Component of the Wnt-Fzd-LRP5-LRP6 complex that triggers beta-catenin signaling	Coronary artery disease, autosomal dominant, 2 (ADCAD2)	Adult tissues.
LRP8_HUMAN Q14114	Low-density lipoprotein receptor-related protein 8/ Apolipoprotein E receptor 2	Cell surface receptor for Reelin (RELN) and apolipoprotein E (apoE)-containing ligands.	Myocardial infarction 1 (MCI1)	Brain, placenta, platelets and megakaryocytic cells
LRP10_HUMAN Q7Z4F1	Low-density lipoprotein receptor-related protein 10	May be involved in the uptake of lipoprotein APOE in liver, Probable receptor, which is involved in the internalization of lipophilic molecules and/or signal transduction	_	blood leukocyte, lung, placenta, small intestine, liver, kidney, spleen, thymus, colon, skeletal muscle and heart
LRP11_HUMAN Q86VZ4	Low-density lipoprotein receptor-related protein 11	_	_	_
LRP12_HUMAN Q9Y561	Low-density lipoprotein receptor-related protein 12/ Suppressor of tumorigenicity 7 protein	May act as a tumor suppressor, Probable receptor, which may be involved in the internalization of lipophilic molecules		heart, skeletal muscle, brain, lung, placenta and pancreas, but not in tissues consisting of a large number of epithelial cells

Table 2a. Amino acid composition of the human low density lipoprotein receptor related protein (LRP) in %

LRP	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
LRP5	5.5	6.1	5.5	8.6	7.3	3.6	5.2	8.3	2.8	4.3	6.7	3.8	1.6	2.9	5.1	6.8	5.7	1.9	2.9	5.4
LRP1B	4.3	5.1	6.3	8.8	7.5	3.3	5.4	7.7	2.8	6	6.8	4.8	1.5	3	4	6.9	5.4	1.9	3.4	4.7
LRP2	4.6	5.3	6.3	8.7	7.1	3.5	5.3	7.2	2.9	5.4	6.2	3.5	1.5	3.7	5	7.3	5.5	1.9	3.5	5.6
LRP3	9.5	8.4	2.1	6.4	5.8	4.9	4.9	10.1	1.9	2.1	9	1.4	0.9	3.1	9.7	8.2	3.5	1.4	2.3	4.2
LRP4	5.9	6.7	5	9	4.9	3.2	5.2	8.1	2.8	4.9	8.4	3.3	1.7	2	4.6	7.7	6.1	2.7	2.5	5.1
LRP5	7.1	6.1	3.8	8.5	3.5	3.5	4.3	7.3	2.4	5.4	9.5	3.2	1.8	3	6.4	6.8	6.1	2.2	3.1	6
LRP6	6	6	4.3	7.6	3.3	3.3	5.5	6.9	2.3	6	8.7	3.5	1.9	2.9	5.6	8.5	5.8	2.2	3.8	6.1
LRP8	7.6	5.2	3.5	8.5	6.5	3.2	6.6	7.3	2.7	4	8.9	4.6	1.1	2.5	6.5	7.4	5.3	1.8	2.1	4.6
LRP10	8.7	7	2.1	5.5	4.6	4.2	5.2	10.1	2.1	2.2	12.1	1	0.8	2.2	10.1	7	5.3	1.4	2.4	5.9
LRP11	11	6.4	1.8	5.4	3.6	4.8	5.8	8.8	2.2	2.2	11.4	3.6	1.4	2	7.6	8.8	4.8	0.8	2.2	5.4
LRP12	6.1	6.6	5.5	6.9	5.7	4	5.8	7.5	2	4.5	7.2	2.6	0.9	4	6.5	9.3	5.4	1.7	2.7	5.2

Table 2b. Physico-Chemical Parameters of the human low density lipoprotein receptor related protein (LRP)

LRP's	No. of amino acids	Molecular weight	pI	'-' charged residue	'+' charged residue	Formular	Extinction Coefficient	Instability index	Aliphatic index	GRAVY
LRP1	4544	504606.3	5.16	628	450	$C_{21574}H_{33338}N_{6308}O_{6913}S_{403}$	698325	35.99	64.12	-0.511
LRP1B	4599	515497.9	5.09	655	459	$C_{22150}H_{34180}N_{6322}O_{7076}S_{413}$	739430	36.33	68.09	-0.490
LRP2	4655	521957.5	4.89	652	409	$C_{22512}H_{34333}N_{6375}O_{7180}S_{399}$	761350	37.41	65.76	-0.471
LRP3	770	82884.0	5.83	87	76	$C_{3572}H_{5557}N_{1071}O_{1107}S_{52}$	90070	51.77	64.58	-0.439
LRP4	1905	212044.8	5.06	272	189	$C_{9132}H_{14192}N_{2664}O_{2917}S_{127}$	363395	41.60	72.80	-0.494
LRP5	1615	179144.5	5.11	208	150	$C_{7893}H_{12258}N_{2192}O_{2408}S_{86}$	270500	39.54	82.67	-0.280
LRP6	1613	180429.3	5.12	211	153	$C_{7944}H_{12306}N_{2192}O_{2451}S_{84}$	288130	46.69	80.63	-0.341
LRP8	963	105633.6	4.88	146	94	$C_{4547}H_{7101}N_{1291}O_{1463}S_{74}$	127175	50.88	71.45	-0.433
LRP10	713	76193.3	5.30	76	57	$C_{3333}H_{5237}N_{955}O_{1016}S_{39}$	82330	54.91	81.57	-0.196
LRP11	500	53311.3	6.04	56	50	$C_{2311}H_{3697}N_{673}O_{725}S_{25}$	39515	58.40	79.70	-0.289
LRP12	859	94983.9	5.07	109	79	$C_{4104}H_{6330}N_{1182}O_{1308}S_{57}$	119770	52.71	67.10	-0.454

Table 3. Functional Domains/Motifs in human low density lipoprotein receptor related protein (LRP)

LRPs	DOMAINS/ DOMAIN ID							
	Ldl_recept_a/F00057	CUB/PF00431	FXa_inhibition/ PF14670	Ldl_recept_b/PF00058	EGF_CA/PF07645	cEGF/PF12662	MANEC/PF00057	EGF/ PF00008
LRP1	P	-	P	P	P	P	-	-
LRP1B	P	-	P	P	P	P	-	P
LRP2	P	-	P	P	P	P	-	-
LRP3	P	P	-	-	-	-	-	-
LRP4	P	-	P	P	-	P	-	-
LRP5	P	-	P	P	-	-	-	-
LRP6	P	-	P	P	-	-	-	-
LRP8	P	-	P	P	-	-	-	-
LRP10	P	P	-	-	-	-	-	-
LRP11	P	-	-	-	-	-	P	-
LRP12	P	P	-	-	-	-	-	-

*Ldl_recept_a- Low-density lipoprotein receptor domain class A; CUB- CUB Domain; FXa_inhibition- Coagulation Factor Xa inhibitory site; Ldl_recept_b - Low-density lipoprotein receptor repeat class B; EGF_CA-Calcium-binding EGF domain; cEGF-Complement Clr-like EGF-like; MANEC- MANEC domain; EGF- EGF like domain. Where P is Present and (–) is absent
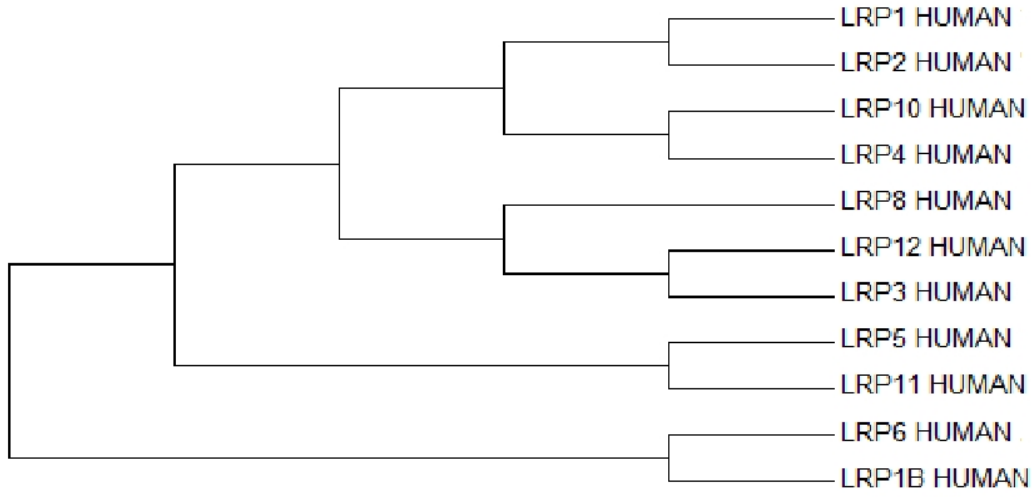
Fig. 1. Phylogram of the human low density lipoprotein receptor related protein (LRP) constructed using NJ methods, showing cluster with similar relations like LRP 12 and3, with LRP8 as their nearest neigbour
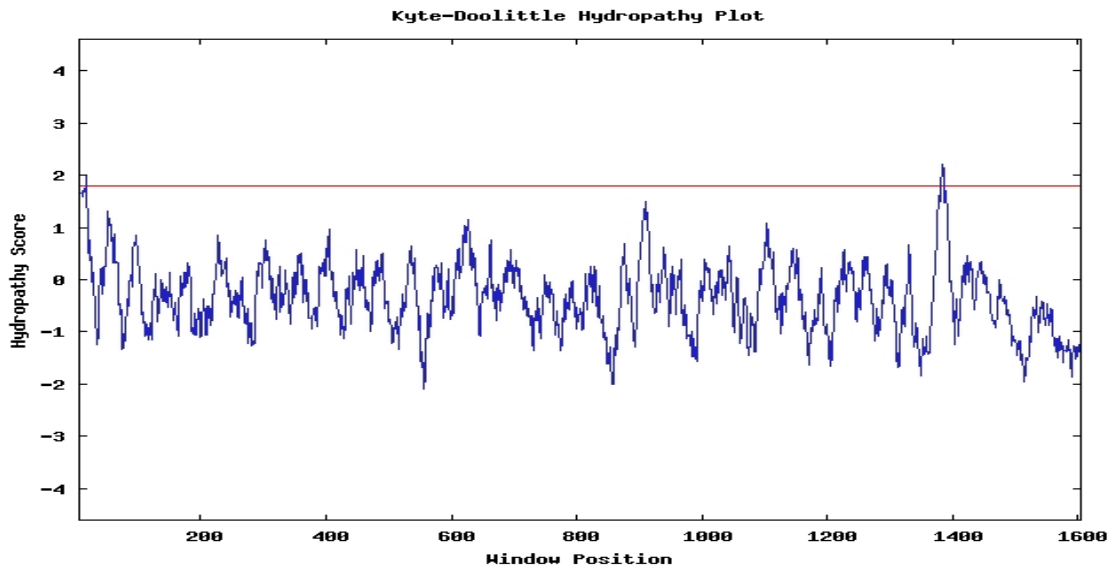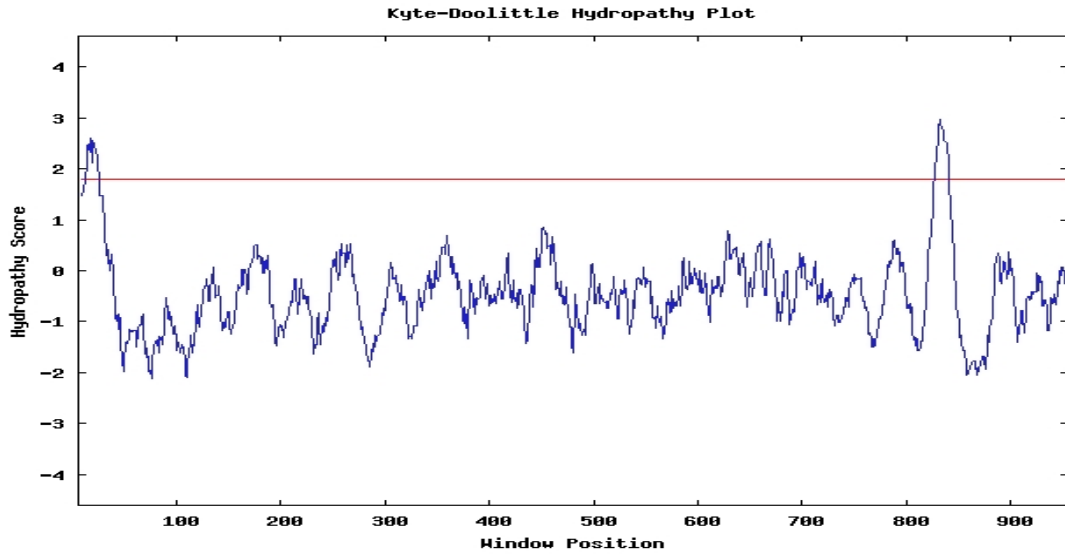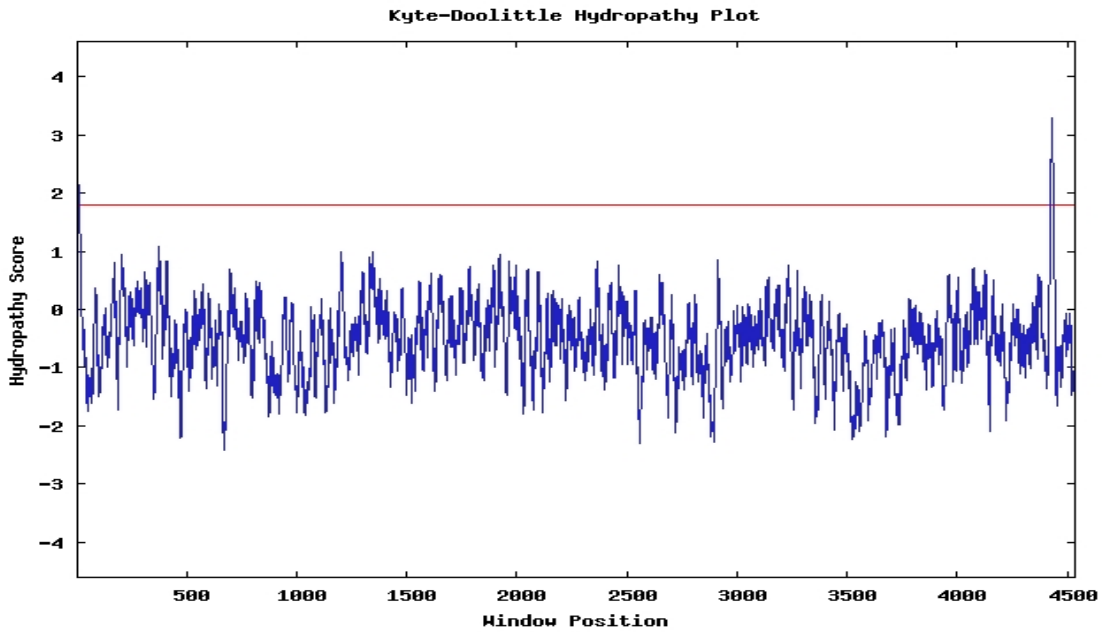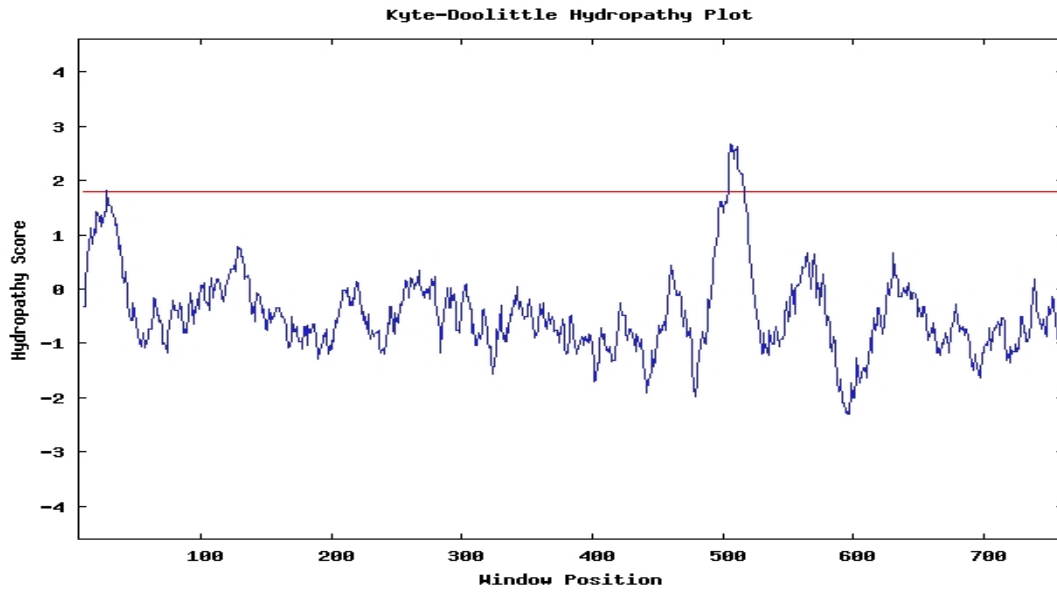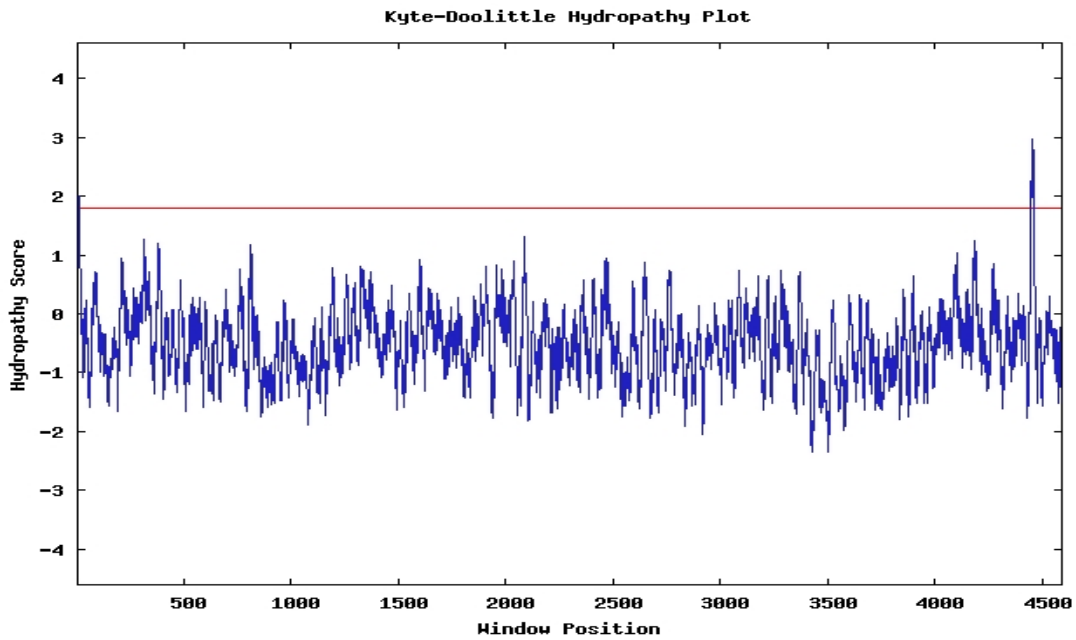

Fig. 2. A hydropathy plot showing the hydrophobic (+), hydrophilic (-) and transmembrane (red line) region in LRP6

297
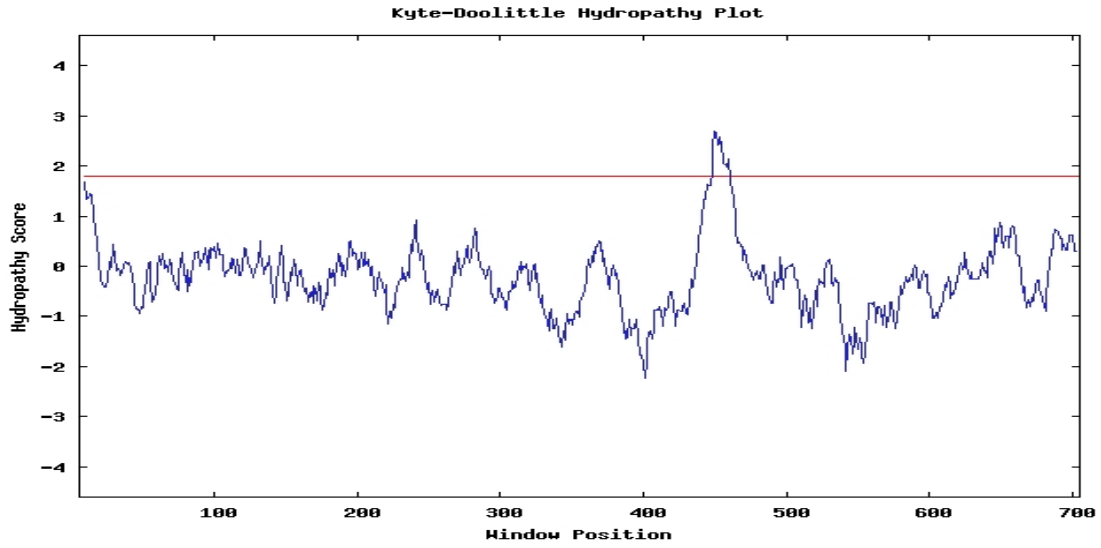
Fig. 3. A hydropathy plot showing the hydrophobic (+), hydrophilic (-) and transmembrane (red line) region in LRP10
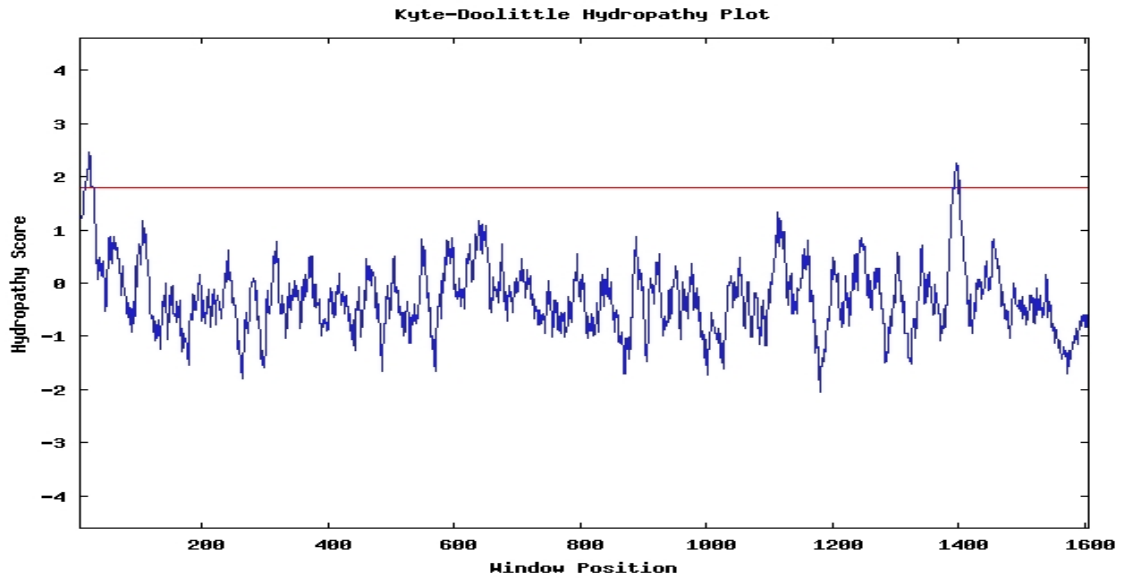
Fig. 4. A hydropathyplot showing the hydrophobic (+), hydrophilic (-) and transmembrane (red line) region in LRP5

Fig. 5. A hydropathy plot showing the hydrophobic (+), hydrophilic (-) and transmembrane (red line) region in LRP8
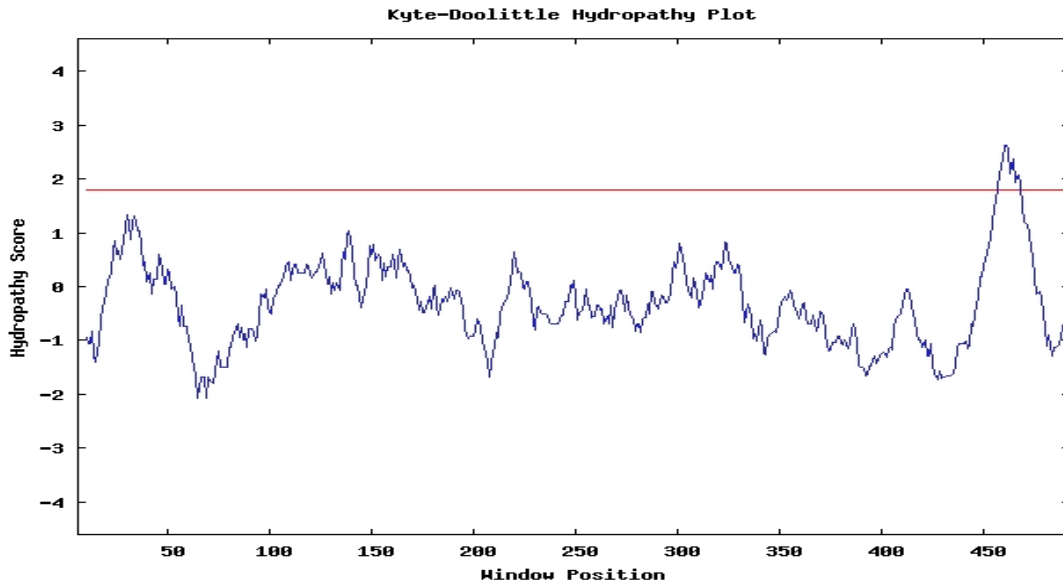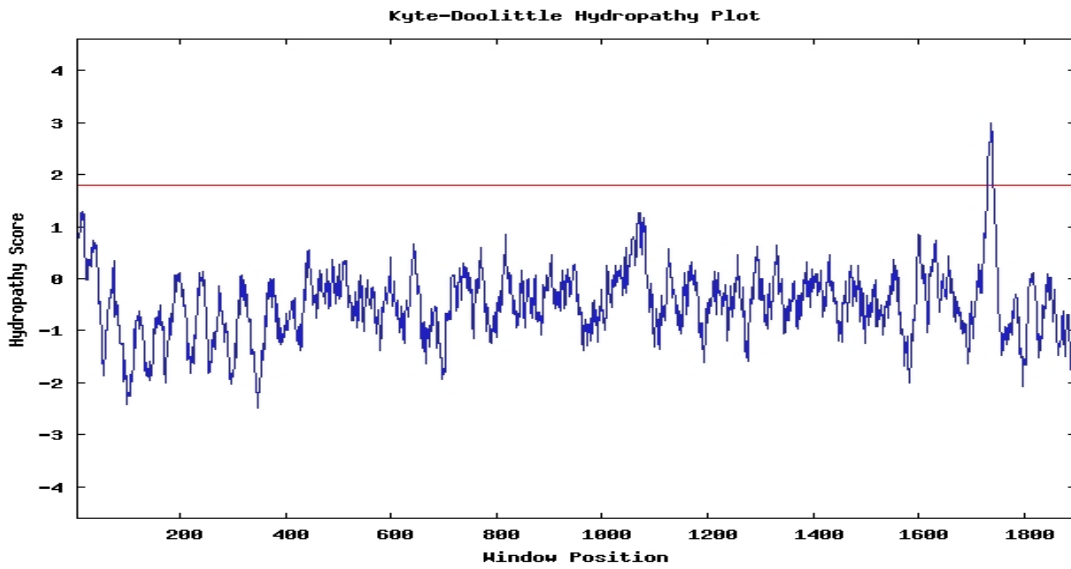

Fig. 6. A hydropathy plot showing the hydrophobic (+), hydrophilic (-) and transmembrane (red line) region in LRP1

Fig. 7. A hydropathy plot showing the hydrophobic (+), hydrophilic (-) and transmembrane (red line) region in LRP3
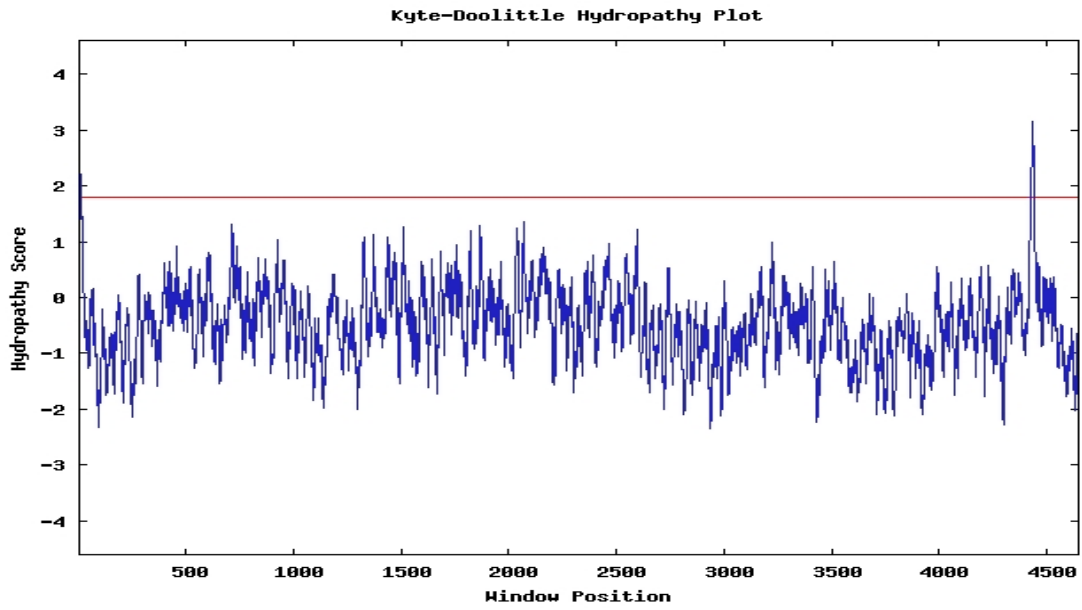

Fig. 8. A hydropathy plot showing the hydrophobic (+), hydrophilic (-) and transmembrane (red line) region in LRP1B

Fig. 9. A hydropathy plot showing the hydrophobic (+), hydrophilic (-) and transmembrane (red line) region in LRP11
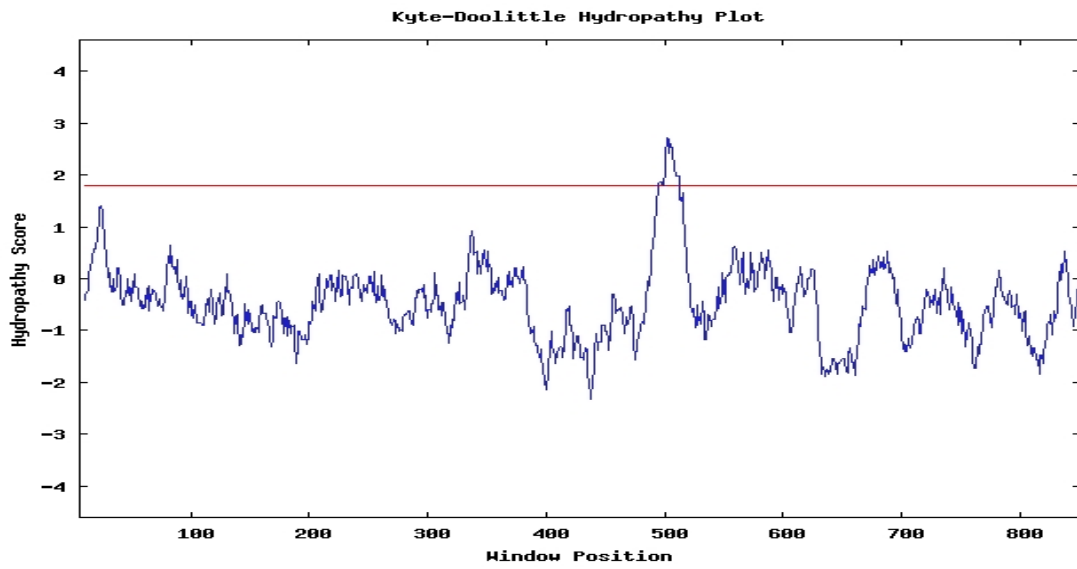

Fig. 10. A hydropathy plot showing the hydrophobic (+), hydrophilic (-) and transmembrane (red line) region in LRP4

301

Fig. 11. A hydropathy plot showing the hydrophobic (+), hydrophilic (-) and transmembrane (red line) region in LRP2

Fig. 12. A hydropathy plot showing the hydrophobic (+), hydrophilic (-) and transmembrane (red line) region in LRP12

4. CONCLUSION

Relative study of functional, structural, phylogenetics and physico-chemical properties of protein family could be extensively achieved with the availability of several computational tools, as detail information on structure, function and evolutionary relationship of a candidate protein with other closely related family members could be retrieved. Owing to the characterization of the human LRP family, a new insight has been achieved. It herefore predict that LRP 3, 12 and 8 as a potential protein cluster showing similarities in many properties. Little is presently known about these three proteins, but this study shows that they share a close evolutionary relation, with both having almost similar functional domains and LRP8 already implicated in diseases like Myocardial infarction1-a condition defined by the irreversible necrosis of heart muscle secondary to prolonged ischemia. Thus, LRP 3 and 12 molecules might be possible candidates in heart related disorder. However our findings can be of significant assistance to researchers whose interest is on LRP protein family to possibly focus on this protein as a proposed contributor in diseases condition.

ACKNOWLEDGEMENT

COMPETING INTERESTS

Author has declared that no competing interests exist.

REFERENCES

1. Harris-White ME, Frautschy SA. Low Density Lipoprotein Receptor-Related Proteins (LRPs), Alzheimer's and Cognition. Current Drug Targets - CNS & Neurological Disorders. 2005;4:469-480.
2. Lillis AP, Van duyn LB, Murphy-ullrich JE, Strickland DK. LDL receptor-related protein 1: unique tissue-specific functions revealed by selective gene knockout studies. Physiol Rev. 2008;88:887–918.
3. Strickland DK, Gonias SL, ArgravesWS, Diverse roles for the LDL receptor family. Trends Endocrinol. vMetab. 2002;13:66.
4. Herz J, Clouthier DE, Hammer RE. LDL receptor-related protein internalizes and degrades uPA-PAI-1 complexes and is essential for embryo implantation. Cell. 1992;71:411.
5. Stricland DK, Kounnas MZ, Argraves WS. LDL receptor-related protein: a multiligand receptor for lipoprotein and proteinase catabolism, FASEB J. 1995;9:890-898.
6. Bacskai BJ, Xia MQ, Strickland DK, Rebeck GW, Hyman BT. The endocytic receptor protein LRP also mediates neuronal calcium signaling via N-methyl-D-aspartate receptors. Proc. Natl. Acad. Sci. U.S.A. 2000;97:11551.
7. Zhuo M, Holtzman DM, Li Y, Osaka H, DeMaro J, Jacquin M, Bu G. J. Role of tissue plasminogen activator receptor LRP in hippocampal long-term potentiation. J Neurosci. 2000;20:542-549.
8. Bairoch A, Apweiler R. The SWISS-PROT protein sequence data bank and its supplement TrEMBL. Nucleic Acids Res. 1997;25:31. [PMID: 9016499].

9. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press. 2005;571-607.

10. Kyte J, Doolitle R. A simple methods for displaying character of a protein. J. Mol. Biol. 1982;157:105-132.

11. Kella LA, Stemberg MJE. Protein structure prediction on wed: a case study using the Phyre server. Nature Protocols. 2009;4:363-371.

12. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance and Maximum Parsimony Methods. Molecular Biology and Evolution. 2011;28:2731-2739.

13. Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution. 1987;4:406-425.

14. Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. Edited in Evolving Genes and Proteins by V. Bryson and H.J. Vogel. Academic Press, New York. 1965;97-166.

15. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database: Nucleic Acids Research Database. 2012;40:290-301.

16. Peter Chou Y, GeraldFasman D. Structural and functional role of leucine residues in proteins. Journal of Molecular Biology. 1973;74(3):263–281.

17. Manisha N, Pracheta A, Aditi J, Aastha C, Astha J, Umang M, et al. Analysis of human collagen sequences. Bioinformation. 2012;8(1):026-033.

18. Becker WM, Kleinsmith LJ, Hardin J, Bertoni GP. The world of the cell, 7th edition. Pearson Education, Inc., publishing as Pearson Benjamin Cummings; 2009.

19. De Leeuw M, Reuveni S, Klafter J, Granek R. Coexistence of Flexibility and Stability of Proteins:An Equation of State. PLoS ONE. 2009;4(10):7296. Doi:10.1371/journal.pone.0007296.

20. Li Y, Lu W, Marzolo MP, Bu G. Differential functions of members of the low density lipoprotein receptor family suggested by their distinct endocytosis rates. J. Biol. Chem. 2001;276(21):18000–6.

21. Beffert U, Stolt PC, Herz J. Functions of lipoprotein receptors in eurons. J. Lipid Res. 2004;45(3):403–9.