African Journal of Biotechnology Vol. 12(16), pp. 2051-2057, 17 April, 2013
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB10.1032
ISSN 1684–5315 ©2013 Academic Journals

Full Length Research Paper

In silico modeling of lipase H

Ammara Jabeen[1], Asif Mir[2], Nyla Jabeen[3], Jabar Zaman Khan Khattak[2] and Naveeda Riaz[3]*

[1]Department of Biosciences, COMSATS Institute of Information Technology, Sahiwal, Pakistan.
[2]Department of Biotechnology and Bioinformatics, International Islamic University (Male campus), Islamabad, Pakistan.
[3]Department of Biotechnology and Bioinformatics, International Islamic University (Female campus), Islamabad, Pakistan.

Accepted 30 September, 2011

LAH 2 is a type of autosomal recessive hypotrichosis that affect hairs, eyebrows, scalp and eyelashes. Mutations in Lipase H gene result in LAH 2. Changes that result from mutation on physiochemical properties, post-translational modifications, functional sites, secondary structure and tertiary structure lipase H gene (LIPH) at molecular level were analyzed in the current study. Results indicate that the 3rd motif of LIPH deletes as a result of mutation. The number of alpha helices and beta sheets become varied in normal and abnormal protein. Tertiary structure of LIPH was predicted through homology modeling. Mutations were then inserted to reveal the difference between normal and abnormal structure. Ligands to target LIPH are also retrieved. ASP178, ASP207 and HIS248 constitute the active site of LIPH. Missense mutations in *LIPH* also brought drastic changes at molecular level that led to imbalanced function. All these changes can be studied through bioinformatics without going for expensive laborious and time consuming experimental techniques.

Key words: LAH2, LIPH, tertiary structure, secondary structure, physiochemical properties, ligands, active site.

INTRODUCTION

Autosomal recessive hypotrichosis (AH) is a form of hypotrichosis (an inherited hair ailment) that is characterized by a condition of having no hair growth and total baldness of the affected area that remains unchanged throughout the life of an affected individual. It affects an individual right from birth and usually continues throughout their lives. Clinical features of autosomal recessive hypotrichosis are diffuse and progressive hair loss, which usually begins in early childhood (Pasternack et al., 2008; Al Aboud et al., 2002), hair on eyebrows, scalp, body and eyelashes are sparse or absent at all, absence of axillary hair, and affected males usually have sparse beards (Wali et al., 2007; Al Aboud et al., 2002). While in LAH 2, normal beard is present in affected males (Aslam et al., 2004). More also, in some of the cases disease is characterized by lack of normal hair follicle structures and comedo-like leftovers of the hair follicle. The leftovers of the hair follicle infundibulum show hyper-keratinization.

Morphologically, sebaceous glands appear to be normal but lost acquaintances to the leftovers of the hair follicle infundibulum (Azeem et al., 2008*)* and hair on legs and arms are absent in males affected by LAH 2. (Aslam et al., 2004).

AH results from mutations in lipase H gene (LIPH), located at chromosome 3q27.2, it is now renamed as LAH2 (Wali et al., 2007). LIPH produces lysophosphatidic acid (LPA) that interacts with P2RY5 to control hair growth. Total 2 missense mutations in *LIPH* gene have been reported, which are c.322T > C (p.W108R) and c.2T > C (p.M1T) (Naz et al., 2009). There are three allelic variants of *LIPH* that are homozygous exon 4 deletion (Kazantseva et al., 2006), c.346-350delATATA in exon 2 (Ali et al., 2007), c.659-660delTA in exon 5 (Jelani et al., 2008). Missense mutations do affect primary structure of protein as a consequence, 3D structure of the protein changes that will ultimately affects function of the protein. Due to mutations physiochemical properties of a protein can also be changed. Physiochemical properties are therefore important for determination of ligand that will bind to that protein (Kahraman, 2007). Isoelectric point is

*Corresponding author. E-mail: naveedariaz@hotmail.com. Tel: +92-51-9258016. Fax: +92519257954.

also crucial in understanding enzyme substrates interactions. Domains folds independently into a tertiary structure (Chothia, 1992), hence any change in domain will definitely affect tertiary structure of the protein and that change can be a consequence of missense mutation.

Moreover, polypeptide that is formed after translation might not be the final protein. Post-translational modifycations occur in a protein to make it functional. Glycosylation pattern is important to predict as it affects protein folding, protein localization, protein trafficking, protein solubility, antigenicity, biological activity, half-life and cell-cell interactions (Gupta et al., 2004). Phosphorylation is important to cellular regulation; cellular signal pathways, metabolism, growth, differentiation and membrane transport (Wong et al., 2007). Prediction of phosphorylation at serine, threonine and tyrosine site is important as it affects multitude of cellular signaling processes (Blom et al., 1999). Topology can be thought of as a critical intermediate stage between amino acid sequence and 3D structure – secondary structure (Heijne, 2006). It involves linking and arrangement of 2D structure elements into 3D space. It is more convenient and accurate than resolving 2D struc-ture through x-ray crystallography because membrane proteins are difficult to crystallize (Tusnaady and Simon, 1998). Secondary structure is crucial to predict as it determines tertiary structure of a protein. Any alteration in secondary structure element contributes to changed tertiary structure. Protein 3D structure is very crucial as it determines the function that protein has to perform. In order to carry out proper function, protein must be carried to designated location (Duan et al., 2001); hence, it is indispensable to predict effect of mutations on physiochemical properties, post-translational modifications, domains, secondary structure and tertiary structure in order to have better therapeutic strategies for a particular genetic disease.

Biological data are available in bulk over the internet. The most practical way to model a protein is to use bioinformatics tools to access data that is present in different databases. In this study, we have predicted the effect of mutations on physio-chemical properties, post-translational modifications, domains, secondary structure and Tertiary structure of *LIPH* gene to highlight the changes brought out by the missense mutations on LIPH. The 3D structure of LIPH is not determined yet. Experimental ways to resolve 3D structures are x-ray crystallography and nuclear magnetic resonance spectroscopy, but these methods are expensive, laborious and time consuming. Due to these reasons, proteins structure information is still limited. However, the protein sequencing methods are very fast and as a result, there is a large gap between the number of known protein sequences and the number of known three-dimensional protein structures (Bernasconi and Segre, 2000). Bioinformatics computational methods and molecular dynamic simulations offer the most attractive way to solve this problem and serve as a substitute for protein structure prediction (Liang and Hsu, 2005). In this study, we predicted the normal and mutated structures of LIPH using bioinformatics tools. In addition, ligands of LIPH were also retrieved.

MATERIALS AND METHODS

Sequence of LIPH was taken from NCBI (Accession: ABM67095.1 GI: 122892055). Reported mutations were retrieved from literature in Pubmed database. InterPro, 3DID, ScanProsite, motif scanner, NetNGlyc, KinasePhos 2.0 and NetCGlyc, were used to predict functional sites of LIPH. Physiochemical properties were envisaged through ProtParam. HNN was used to predict secondary structure, while tertiary structure was predicted by homology modeling approach using MODELLER. In the first step homologs were searched through PSI-BLAST. LIPH sequence was search against PDB database. Two PDBs were selected as template. Sequence to structure alignment was performed with the help of MODELLER. Ten models were generated for each template; model with lowest objective function value was selected as the best model for both templates. Procheck and Prosa web analysis were then used to evaluate the quality of the two predicted models, and model with lowest Z-score was selected as the best model. Mutations were then inserted in the LIPH sequence and predicted structure and results for mutated LIPH were retrieved by same tools. Active sites were predicted by giving input of predicted tertiary structure of LIPH to PAR3D and ligands were retrieved from KEGG ligand database.

RESULTS

Domains

After inserting mutations in LIPH sequence, 3rd motif of TAGLIPASE domain was deleted, and as a result the function that it performed was suppressed. This therefore had an overall effect on the function of protein.

Post-translational modifications

No effect of mutation was predicted in N-glycosylation, cAMP- and cGMP-dependent protein kinase phosphorrylation site, protein kinase C phosphorylation site and Casein kinase II phosphorylation sites in LIPH.

Physiochemical properties

There are total of 451 amino acids in LIPH. Molecular weight of LIPH is 50859.3, theoretical isoelectric point is 7.15 and there are total 7108 atoms. The pI value of LIPH with mutation p.W108R was 7.53 and with mutation p. M1T, it was 7.11, respectively. Molecular weight of LIPH with mutation p.W108R was 50829.3 and with mutation p. M1T molecular weight was 50829.2. Hence, theoretical isoelectric point and molecular weight of LIPH were changed

as a consequence of mutations.

Secondary structure

Mutations also affected secondary structure of LIPH. Number of alpha helices and beta sheets changed, thus leading to a change in tertiary structure. There were a total of 117 alpha helices and 92 extended strands in the secondary structure of the normal LIPH. However, with mutation p. W108R, the number of alpha helices remained the same but extended strands increased to 93. While with mutation p. M1T, the number of alpha helices reduced to 116, while the number of extended strands increased to 93.

Tertiary structure

Tertiary structure of LIPH was predicted through homology modeling approach. Z-score for model predicted by using 1BU8 as a template was -1.156 and for the model predicted by using 1W52 as a template was -2.473. Since 1W52 has less Z-score, the 3D structure that was predicted using 1W52 as a template was therefore selected. Predicted tertiary structure of LIPH in a ribbon view is shown in Figure 2. This view of protein has been got in pmol. Normal structure with tryptophan at position no.108 is shown in Figure 3(a), while mutated structure with arginine at position 108 is shown in Figure 3(b). Also, normal structure with methionine at position no. 1 is shown in Figure 4(a), while mutated structure with threonine at position no. 1 is shown in Figure 4(b). To highlight the difference, mutated amino acid is shown in sphere view, while the rest of the protein is shown in cartoon view. In normal LIPH, amino acid is shown in red color, while in mutated LIPH, amino acid is shown in pink color.

Active sites and ligands

LIPH has metal binding site that is constituted of ASP178, ASP207 and HIS248 (Figure 5). Moreover, ligands of LIPH include aliphatic amide, aliphatic alcohol, aliphatic aldoxime and aliphatic hydroxy acid (Figure 6).

DISCUSSION

LIPH has lipase and tag lipase domains. Lipases are basically lipolytic enzymes that are involved in catalyzing hydrolysis of ester linkages of triglycerides. Lipase domain interacts with colipase and PLAT domain. As a result of mutation, one of the motifs of tag lipase domains disappeared, resulting in loss of a functional site, and as a consequence, functionality of *LIPH* is altered. LIPH contains N-glycosylation site, cAMP- and cGMP-dependent protein kinase phosphorylation site, protein kinase C phosphorylation site, casein kinase II phosphorylation site, N-myristoylation site, cell attachment sequence and leucine zipper pattern. LIPH is glycosylated at positions 50, 58, 66 and 357; serine at positions 14, 29, 97, 121,122, 144, 206, 262, 272, 287, 324 and 361 is phosphorylated; threonine at positions 21, 81, 124, 320 and 348 is phosphorylated; while tyrosine at position 149, 213, 223, 273 and 276 is phosphorylated. LIPH also has cAMP-dependent protein kinases phosphorylation sites that are present at 14, 16, 60, 97 and 196 residue number. Residues 16, 56, 97, 124, 144, 206, 258, 272, 287, 320 and 359 are the casein kinase II phosphorylation sites. Secondary structure of LIPH comprises of 117 alpha helices, 92 extended strands and it contains 242 random coils. Elements of secondary structure are highlighted in Figure 8. This figure depicts that how folding of secondary structure elements occurs in 3D structure of LIPH.

Mutation p.W108R results in conversion of tryptophan at position number 108 to arginine. Due to this mutation, threonine at position number 113 that was not a part of secondary structure in normal LIPH, now participates in the formation of an extended strand at 108 locations instead of 109 and was 6 residues long instead of five. This extended strand of 6 thus affected the tertiary structure. Moreover, mutation p.M1T results in conversion of methionine at position number 1 to threonine. This mutation results in decrease in alpha helices within protein structure (Figure 4a and b). Cystine at position number 12 makes alpha helix in normal LIPH, but due to mutation it now make random coil.

Tertiary structure of LIPH was predicted through homology modeling (Figures 3a and b). Chain X, crystal structure of a proteolyzed form of pancreatic lipase related protein 2 from horse (PDBID: 1W52X) with E-value 6e-56 and query coverage 89%, and chain A, rat pancreatic lipase related protein 2 (PDBID: 1BU8A) with E-value 3e-55 and query coverage 72% were selected as templates. Through MODELLER tertiary structure of *LIPH* was predicted. Model predicted by using 1W52 as a template was selected as the best model on the basis of Z-score calculated by Prosa. Normal and mutated structures were superimposed on each other by using Swiss PDB Viewer to highlight the differences between the two structures. Difference is highlighted in yellow color (Figure 7a and b). LIPH interacts with PLA-1, EDG7, PDNP2, STCH, LCAT, GP1B, PNPLA2, SOAT1, LIPF and CEL proteins (Figure 1).

PLA-1 shows sequence-specific DNA binding and transcription factor activity. EDG7 acts as G-protein coupled receptor and as a transducer. PDNP2 is involved in hydrolysis of lysophospholipids to form lysophosphatidic acid (LPA) in extracellular fluids, with lysophosphatidylcholine being its key substrate. STCH has peptide-independent ATPase activity. LCAT mediates

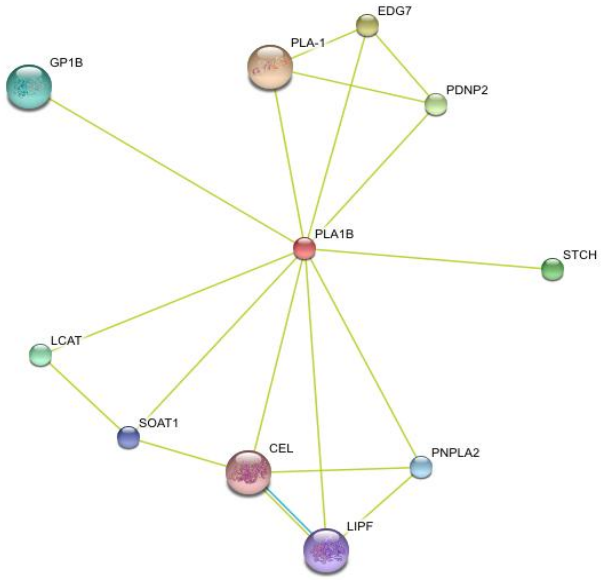


Figure 1. Interaction pathway of LIPH; all of the proteins LIPH interacts.

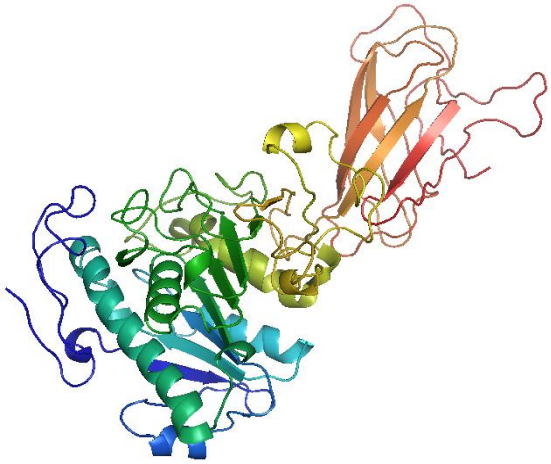


Figure 2. Tertiary structure of LIPH (ribbon mode) with 12 predicted Helices and 13 strands.

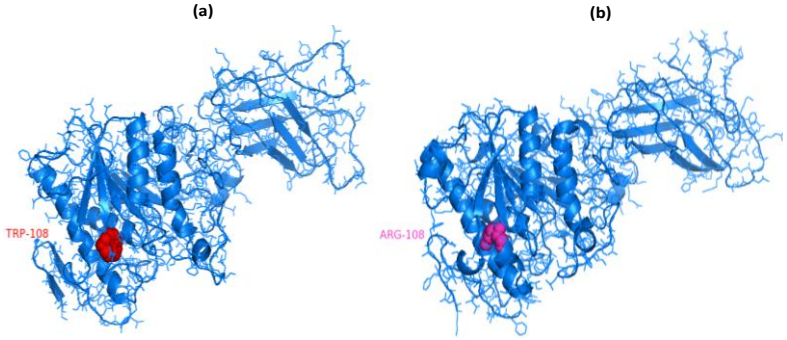


Figure 3. (a) Normal structure with tryptophan at position 108 (red color); (b) Mutated structure with arginine at position 108 (pink color).

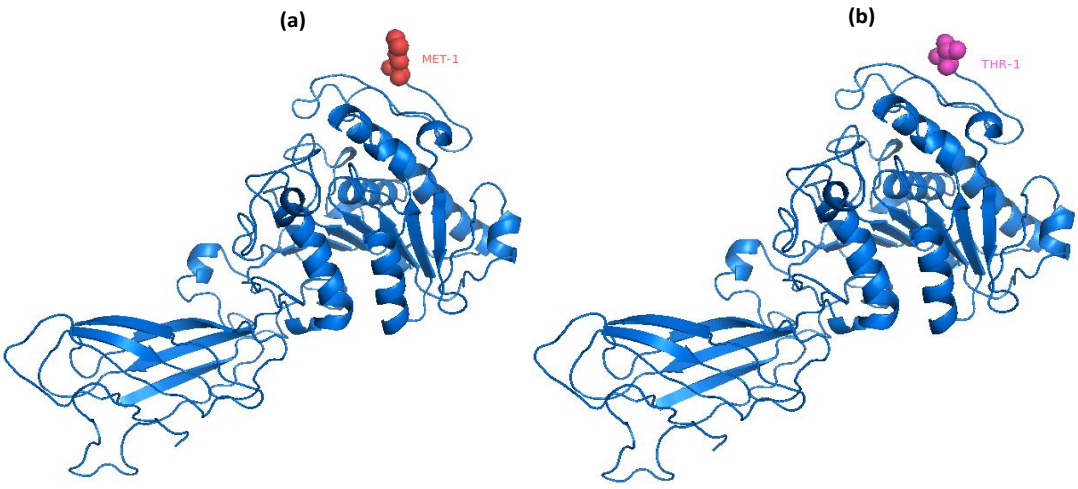


Figure 4. (a) Normal structure with methionine at position 1 (red color); (b) Mutated structure with threonine at position 1 (pink color).

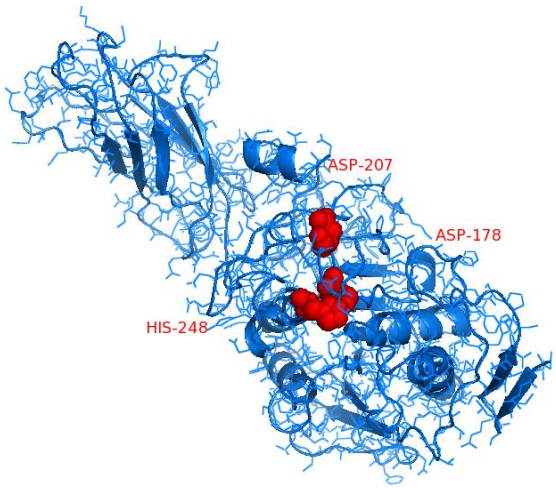


Figure 5. Active sites in *LIPH* (red color).

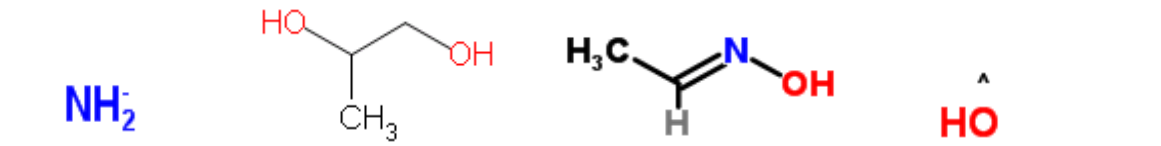
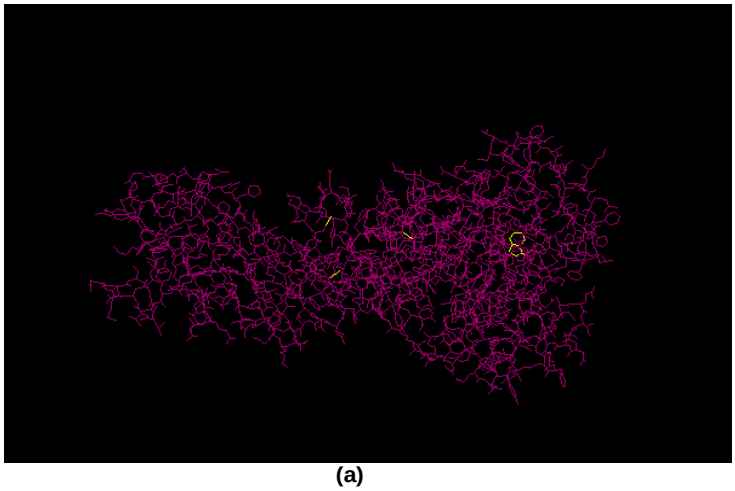
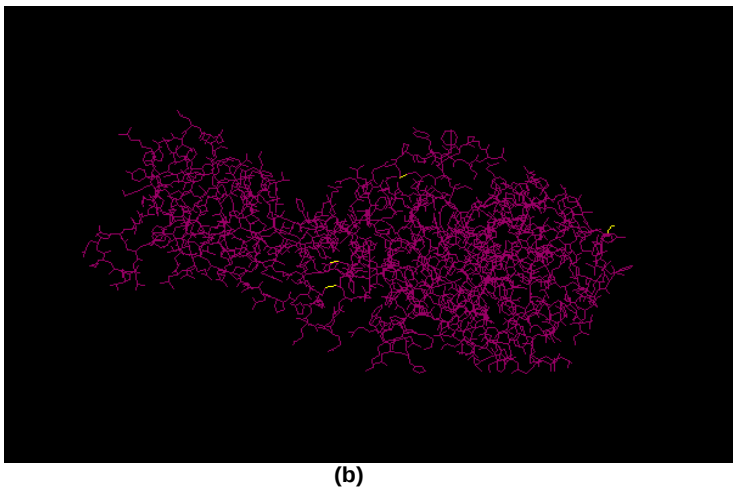


Figure 6. Ligands for LIPH.

mediates electro-neutral potassium-chloride co-transport when activated by cell swelling. GP1B catalyzes the initial step in triglyceride hydrolysis in adipocyte and non-adipocyte lipid droplets. It has acylglycerol transacylase activity, which may act coordinately with LIPE/HLS within the lipolytic cascade, regulates adiposome size and involved in the adiposomes degradation. It also has a role in energy homeostasis, response of an organism against starvation and in enhancing hydrolysis of triglycerides and providing free fatty acids to other tissues to be oxidized in situations of energy depletion. SOAT1 is an enzyme that catalyzes the reaction for the formation of fatty acid-cholesterol esters and is involved in assembly of lipoprotein and dietary cholesterol absorption. It also



(a)



(b)

Figure 7. (a) Difference in mutated (p.W108R) and normal structure; (b) Difference in mutated (p.M1T) and normal structure.

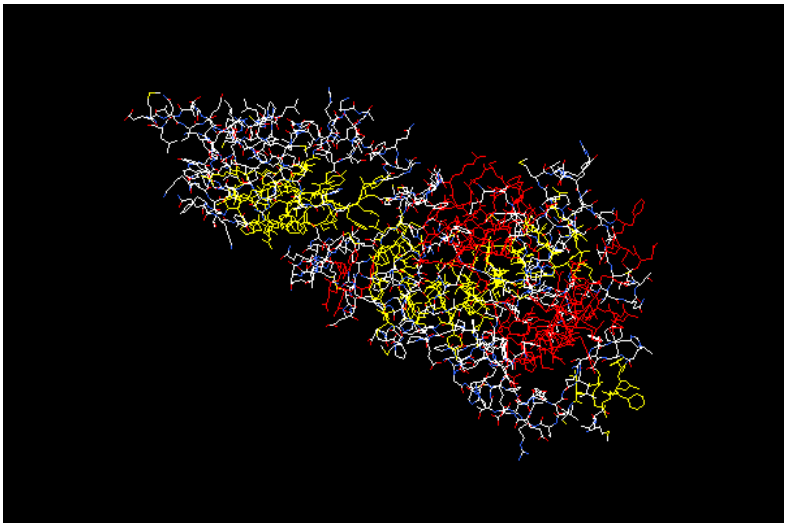


Figure 8. Secondary structure elements; helices are highlighted by red color, while strands are highlighted by yellow color.

exhibits acyltransferase activity and act as a ligase. LIPF has lipid-binding and triacylglycerol lipase activity. While CEL is an enzyme that is responsible for catalyzing fat and vitamin assimilation and acts in combination with colipase and pancreatic lipase for the absolute digestion of dietary triglycerides.

Conclusion

Due to missense mutations, isoelectric point of lipase H changes resulting in an alteration in the interaction pattern of proteins. There is also an alteration in secondary structure of a lipase H as a consequence of which tertiary structure that is directly related to functionality of protein also changes. *LIPH* can therefore be targeted to surmount LAH2.

REFERENCES

Al Aboud K, Al Hawsawi K, Al Aboud D, Al Githami A (2002). Hereditary hypotrichosis simplex: report of a family. Clin. Exp. Derm. 27:654-656.

Ali G, Chishti M, Raza S, John P, Ahmad W(2007). A mutation in the lipase H (*LIPH*) gene underlie autosomal recessive hypotrichosis. Hum. Genet. 121:319-325.

Aslam M, Chahrour M, Razzaq A, Haque S, Yan K, Leal S, Ahmad W (2004). A novel locus for autosomal recessive form of hypotrichosis maps to chromosome 3q26.33-q27.3. J. Med. Genet., 41: 849-852.

Azeem Z, Jelani M, Naz G, Tariq M, Wasif N, Naqvi S, Ayub M, Yasinzai M, Amin-ud-din M, Wali A, Ali G, Chishti M, Ahmad W (2008). Novel mutations in G protein-coupled receptor gene (P2RY5) in families with autosomal recessive hypotrichosis (LAH3). Hum. Genet.123:515-519.

Bernasconi A, Segre A (2000). Ab Initio Methods for Protein Structure Prediction: A New Technique based on Ramachandran Plots. ERCIM News 43.

Blom N, Gammeltoft S, Brunak S (1999). Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. J. Mol. Biol. 294:1351-1362.

Chothia C (1992). Proteins. One thousand families for the molecular biologist. Nature 357:543-544.

Duan Y, Kollman P (2001). Computational protein folding: From lattice to all atoms. IBM Systems. J. 40:297-309.

Gupta R, Jung E, Brunak S (2004). Prediction of N-glycosylation sites in human proteins. Retrieved June 2009 from *http://www.cbs.dtu.dk/services/NetNGlyc/abstract.php*.

Heijne G, (2006). Membrane-protein topology. Nature Rev. Mol. Cell Biol. 7:909-918.

Jelani M, Wasif, N, Ali G, Chishti M, Ahmad W (2008). A novel deletion mutation in LIPH gene causes autosomal recessive hypotrichosis (LAH2). Clin. Genet. 74:184-188.

Kahraman A, Morris R, Laskowski R, Thornton J (2007). Variation of geometrical and physicochemical properties in protein binding pockets and their ligands, BMC Bioinformatics 8:S8-S1.

Kazantseva A, Goltsov A, Zinchenko R, Grigorenko A, Abrukova P, Moliaka A, Kirillov Y, Guo A, Lyle Z, Ginter S, Rogaev E, Evgeny I (2006). Human hair growth deficiency is linked to a genetic defect in the phospholipase gene LIPH. Science 314:982-985.

Liang H, Hsu J (2005). Recent developments in structural proteomics for protein structure determination. Proteomics 5:2056-2068.

Naz G, Khan B, Ali G, Azeem Z, Wali A, Ansar M, Ahmad W (2009). Novel missense mutations in lipase H (LIPH) gene causing autosomal recessive hypotrichosis (LAH2). J .Dermatol Sci. 54:12-16.

Pasternack S, Von Kugelgen I, Al Aboud K, Lee Y, Ruschendorf F, Voss K, Hillmer A, Molderings J, Franz T, Ramirez A, Nurnberg P, Nothen M, Betz R (2008). G protein-coupled receptor P2Y5 and its ligand LPA are involved in maintenance of human hair growth**.** Nature Genet. 40:329-334.

Wali A, Chishti S, Ayub M, Yasinzai M, Ali G, John P, Ahmad W (2007). Localization of a novel autosomal recessive hypotrichosis locus (LAH3) to chromosome 13q14.11-q21.32. Clin. Genet. 72:23-29.

Wong Y, Lee T, Liang H, Huang C, Wang T, Yang C, Huang H, Tat M, Hwang J (2007). Kinase Phos 2.0: a web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns. Nucleic Acids Res. 35:588-594.