Computational Biology Approach for Therapeutic Intervention of Alexander Disease by Post Transcriptional Gene Silencing

Tammanna R. Sahrawat1, Sanya Taneja1, Gurpreet Kaur1
1 Centre for Systems Biology and Bioinformatics, UIEAST, Panjab University, Chandigarh

ABSTRACT

Alexander disease (AxD) primarily affects the white matter of the central nervous system (CNS). It is an astrogliopathy in which Astrogial cells involved in maintenance homeostasis and providing defence to the brain are affected. Therefore their dysfunction has been implicated in a number of neurological, neuropsychiatric and neurodegerative disorders. GFAP (Glia fibrillary acidic protein) is the major intermediate filament protein present in astrocytes whose heterozygous missense mutations have been reported to be a cause of AxD. In the absence of any effective therapeutic intervention of AxD, in the present study PTGS (Post transcriptional gene silencing) approach to knock down mutant \textit{gfap} gene. Various mutations causing AxD were checked for their pathogenicity using various \textit{in silico} tools and 13 mutations were shortlisted based on their pathogenicity and probability of occurrence. Thereafter siRNA were designed against the mutant genes to silence them and thereby preventing the accumulation of mutant \textit{gfap} that causes the pathophysiology of AxD.

How to Cite:
Tammanna R. Sahrawat et. al. Computational Biology Approach for Therapeutic Intervention of Alexander Disease by Post Transcriptional Gene Silencing. IJCB.2017; Volume 6 (Issue 1): Page 18-24.
The activity of specifically target genes can be altered/terminated using the technique of gene silencing which is based on a natural process that is used to degrade mRNA transcribed from a specific gene, thereby preventing expression of gene and synthesis of corresponding protein. Therefore to silence gene expression Post Transcriptional Gene Silencing (PTGS) has been used to cause effective and specific inhibition of disease-causing genes [9]. siRNA molecules based therapy has been successfully used to target a number of mutant genes such as SOD1 associated with ALS [10], viral gene expression in cancerous cells, knockdown of host receptors and coreceptors for HIV [11], silencing of hepatitis A [12], hepatitis B genes [13], influenza gene expression [14], and measles viral replication [15]. ‘Genetic medicines’ have been developed for the treatment of nearly 1,800 known monogenic hereditary disorders [15]. PTGS has also been used to treat asthma and metabolic diseases such as chronic obstructive pulmonary disease (COPD) and cystic fibrosis [16]. It has also been reported that bacterial diseases can be prevented by targeting the host genes involved in the immune response caused by infection or those that mediate entry of bacteria into host cells [17-18].

Alexander disease (AxD) is a very rare, progressive and fatal neurological disorder that was first described by WS Alexander in 1949 characterized by genetic disorder of astrocytes that mainly affects young children [19]. The pathophysiology of the disease is decreased myelin formation, which is the fatty, insulating wrapping sheath around nerve fibers and constitutes the white matter of brain. Therefore AxD has been included among leukodystrophies [20].

The characteristic feature of AxD is the accumulation of abnormal protein aggregates of glial fibrillary acidic protein (GFAP) in astrocytes which are referred as “Rosenthal fibers” dispersed in the brain and spinal cord [21]. AxD has been reported to be caused due to heterozygous, sporadic and missense mutations in gfap that codes for GFAP present on chromosome 17q21 [22]. Brenner et al. (2001) performed sequence analysis of DNA from patients representing different AxD phenotypes, reported that most cases were associated with nonconservative mutations in the coding region of the gfap gene [23] and nearly one-third involve either of two amino acids- ARG79 or ARG239 [24]. ARG79 and ARG239 (R79 and R239) positions hot spots have produced the greatest variety of substitutions: C, H, L, G and S for ARG79 (R79) (only P is missing among possible single nucleotide mutations), and C, H, L and P for ARG239 (R239) (only G and S are missing). Arginine codons are recognized as particularly prone to mutation, presumably due to methylation of the CpG dinucleotide [25].

For the treatment of AxD the current standard focuses on control of seizure, nutrition, and maintenance of pulmonary function. Only three reports describe attempts at alternative forms of therapy such as bone marrow transplantation which was given to one patient who died after 4.5 months at the age of 1 year [26]. Due to absence of any specific therapeutic intervention of Alexander Disease, the present study was undertaken to target the disease at genetic level by designing siRNAs to knock down the mutant gfap gene thereby preventing expression of mutant protein to obviate pathophysiology of disease resulting from accumulation of Rosenthal fibers.

2. RESEARCH METHOD

2.1 Sequence Retrieval:
The nucleotide sequence for gfap gene (Homo sapiens) was retrieved from NCBI (Accession No. - NM_002055).

2.2 Mutation Analysis

2.2.1 Databases
The variations reported to cause AxD were collected from the database OMIM.

2.2.2 Prediction of pathogenic mutations
To check the pathogenicity of these reported variations in the GFAP protein, they were analyzed using in silico mutation analysis tools to predict their damaging effect on the protein function.

2.2.2.1 UMD Predictor (Universal Mutation Database Predictor)
UMD predictor was used for assessment of pathogenic missense and synonymous mutations and follows a combinatorial approach. The algorithm score ranges from 0 to 100 and is divided into four classes - <50 polymorphism, 50-64 probable polymorphism, 65-74 probably pathogenic mutation, >74 pathogenic mutation [27].

2.2.2.2 Provean (Protein Variation Effect Analyzer)
It is a web server which supports high-throughput analysis for human and mouse variants at both the genomic and protein levels and also provides accelerated analysis of protein variants from any organisms. If the PROVEAN score is equal to or below a pre-defined threshold (e.g. -2.5), the protein variant is predicted to have a "deleterious" effect otherwise the variant is predicted to have a "neutral" effect [28].

2.2.2.3 PolyPhen-2 (Polymorphism Phenotyping v2)

The effect of a single-nucleotide change is predicted based on various features formulated from sequence annotations, multiple sequence alignments, and, where available, 3-D structures. PolyPhen-2 predicts change of a single nucleotide between numerical score ranging from 0.0 (benign) to 1.0 (damaging) [29].

2.3 Post Transcriptional Gene Silencing

2.3.1 DESIGN OF siRNA

Whitehead siRNA imbrute the design of short oligonucleotides twenty one nucleotide in length to prevent the expression of target genes. This tool uses various rules to design effective siRNAs such as thermodynamic stability of the double stranded RNA duplex, GC content, presence of SNPs [30].

2.3.2 Self-Complementary analysis of antisense strands

2.3.2.1 RNAfold tool (http://rna.tbi.univie.ac.at/)

Secondary structure can be predicted from a single sequence or the consensus for a set of aligned sequences using Vienna RNA secondary structure server. The server predicts minimum free energy (mfe) structure of a single sequence via classic algorithm of Zuker and Stiegler [31] and the centroid structure [32].

3. RESULTS AND ANALYSIS

Alexander disease typically results from dominant mutations in gfap that arise de novo. Rosenthal fibers that are a prominent feature in AxD are formed due to accumulation of mutant glial fibrillary acidic protein (GFAP) [23]. Rosenthal fibers are also found to occur in some forms of cancer. Nam et al. (2015) has reported that most of the gfap mutations that have been identified in AxD are heterozygous, sporadic and missense mutations [33].

To identify the validated variations that are known to cause the disease, the database OMIM was used. From the list of hits gene entry for Homo sapiens was selected having cytogenetic location -17q21.31 and the genomic coordinates- 17:44,905,625-44,915,551 [34]. The mutations causing the disease were explored using “Allelic Variants” option of OMIM and thirteen mutations in gfap (Table 1) were obtained. Therefore the present study was focused on these mutations.

The pathogenicity of these 13 variations were analyzed using in silico mutation analysis tools PolyPhen 2, UMD and Provean. For analysis in UMD predictor, multiple analyses were performed taking all the known mutations as the input. The tool allows the pathogenic assessment of the mutations using a combinatorial approach and gives results in a tabulated form where they are categorized according to the score that is generated (Table 1).

| S.No | Substitution at nucleotide level | Substitution at the amino acid level | PolyPhen Score | UMD-predictor Score | Provean score |
|------|----------------------------------|-------------------------------------|----------------|---------------------|---------------|
| 1.   | C>T                              | R239C                               | 1.000          | 99                  | -7.37         |
| 2.   | G>A                              | R239H                               | 1.000          | 69                  | -4.60         |
| 3.   | C>T                              | R416W                               | 0.997          | 100                 | -5.99         |
| 4.   | G>A                              | R79H                                | 0.999          | 78                  | -4.84         |
| 5.   | C>T                              | R79C                                | 0.466          | 96                  | -6.99         |
| 6.   | C>T                              | R88C                                | 0.999          | 96                  | -7.40         |
| 7.   | C>A                              | R88S                                | 0.978          | 96                  | -5.49         |
| 8.   | C>T                              | L76F                                | 0.993          | 78                  | -3.86         |
| 9.   | A>T                              | N77Y                                | 1.000          | 93                  | -7.73         |
| 10.  | G>C                              | E362D                               | 1.000          | 69                  | -2.71         |
| 11.  | G>T                              | R276L                               | 1.000          | 93                  | -6.51         |
| 12.  | T>C                              | L352P                               | 1.000          | 87                  | -6.40         |
| 13.  | C>A                              | D78F                                | 0.985          | 69                  | -3.87         |
The number of variants that were obtained from UMD predictor was 79, greater than the number of submitted variants due to the existence of variants in multiple transcripts. Out of these 62 were pathogenic, 11 were probably pathogenic, 3 were polymorphism and 3 each showed polymorphism and probable polymorphism.

The 13 pathogenic variations obtained from OMIM were also analyzed with two other in silico tools i.e. Provean and PolyPhen 2. The results obtained from all the in silico tools are tabulated in Table 1 to compare and assess the accuracy of prediction of mutations and shortlist the possible variants resulting in synthesis of mutant GFAP protein. The variations of gfap which are reported to be pathogenic in the OMIM databases were also found to be damaging or deleterious when analyzed using various in silico tools.

Previous studies on the disease revealed that the mutations involving amino acid substitutions from arginine had the most common occurrence with a percentage of about 50.5% in the reported cases [35]. In the present study the mutations ARG239CYS (R239C), ARG416TRP (R416W), ARG88CYS (R88C) and ARG276LEU (R276L) were found to be most damaging and deleterious mutations. Nam et al have previously reported that ARG239CYS (R239C), ARG416TRP (R416W), ARG88CYS (R88C) and ARG276LEU (R276L) mutations occurring at nucleotide positions 729, 1246, 328 and 841[36-39] respectively are responsible for the pathophysiology of AxD and the formation of Rosenthal fibers thereby validating our observations.

In the absence of any therapy/commercially available drug for treatment of AxD or alleviation of its neuropathological feature of aggregation of Rosenthal Fibers, the present study was undertaken to prevent the expression of mutant gene, which may be targeted using Post Transcriptional Gene Silencing by design of siRNA constructs.

siRNA were designed for the entire 13 gene mutations previously shortlisted using Whitehead siRNA web server and 10 constructs were obtained. As per siRNA design guidelines if GC content of one strand is greater than 60% it cannot become a functional siRNA because bond between guanine and cytosine are much stronger compared to A-U or A-T double bonds and each strand of siRNA can have a tendency to form loops by itself. Therefore, constructs C2, C3 and C4 were rejected as they form self-complimentary structures (Table 2) which may prevent the constructs from binding to their target mRNA and have effective PTGS role.

For gene silencing those siRNAs can be effective which do not form secondary structure (internal hairpin and stem loops). Therefore, siRNA constructs C1, C5 and C6 were good candidates for gene silencing of gfap. Further the binding energy of constructs C1 and C6 were highly negative indicating highest specificity followed by C5 with nucleotide binding positions of 14-36, 1345-1367 and 1319-1341 in the gfap gene respectively.

### Table 2. siRNA secondary structures obtained using RNAfold server.

| No. | Possible secondary structure (mfe) | Possible secondary structure (css) | Position | Free energy (kcal/mol) |
|-----|----------------------------------|----------------------------------|----------|-----------------------|
| C1  | CCCACTCCTCCTCATAAAGGCCCTCG       |                                  | 14-36    | -0.06                 |
| C2  | TGCCCTATAGACAGGAAGCAGATG         |                                  | 576-598  | 2.2                   |
| C3  | TCCAAGTTTGAGCAGACCTGAGA          |                                  | 841-863  | 1.4                   |
| C4  | GACACCAAGTCGTGTCAGAAGG           |                                  |          |                       |

In the absence of any therapy/commercially available drug for treatment of AxD or alleviation of its neuropathological feature of aggregation of Rosenthal Fibers, the present study was undertaken to prevent the expression of mutant gene, which may be targeted using Post Transcriptional Gene Silencing by design of siRNA constructs.
4. CONCLUSION
In the present study, a computational biology approach has been used to target mutant gene gfap reported to be associated with AxD. siRNA were designed to prevent synthesis of mutant protein which results in formation of abnormal deposits known as Rosenthal Fibers. Three siRNA constructs designed in the present study i.e. C1, C6 and C5 which may be further tested for their efficacy in PTGS of mutant gfap to therapeutically alleviate symptoms of AxD by preventing the synthesis of mutant protein.

ACKNOWLEDGEMENTS
The authors would like to thank their parent institute Panjab University Chandigarh, India for providing the infrastructure for carrying out the research.

REFERENCES
[1] Sahrawat TR, Taneja S. Investigation of mutant GFAP protein associated with Alexander Disease and its Therapeutic Intervention: Structure Based Drug Design, Journal of Progressive Research in biology. 2016 Oct; 3(2):184-191.
[2] Werner T. Next generation sequencing in functional genomics. Briefings in Bioinformatics. 2010; 11(5):499-511.
[3] Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, et al. The complete genome of an individual by massively parallel DNA sequencing. Nature. 2008 04/17/print; 452(7189):872-6.
[4] Kadakkuzha BM, Puthanveettil SV. Genomics and proteomics in solving brain complexity. Molecular BioSystems. 2013; 9(7):1807-21.
[5] Pillay K, Gene mutations: Understanding the significance using in silico analysis, Journal of the Practice of Cardiovascular Sciences. 2015; 1(1):58-59.
[6] Leong IUS, Stuckey A, Lai D, Skinner JR, Love DR., Assessment of the predictive accuracy of five in silico prediction tools, alone or in combination, and two metaservers to classify long QT syndrome gene mutations, BMC Medical Genetics, 2015; 16(1):34.
[7] Boillee S, VandeVelde C, Cleveland DW. ALS: A Disease of Motor Neurons and Their Non neuronal Neighbors, Neuron. 2006; 52(1):39-59.Cited in PubMed; PMID 17015226.
[8] Minde DP, Anvarian Z, Rüdiger SG, Maurice MM. Messing up disorder: how do missense mutations in the tumor suppressor protein APC lead to cancer? Molecular Cancer. 2011; 10:101 doi:10.1186/1476-4598-10-101. Cited in PubMed; PMID 21859464.
[9] Yang D, Goga A, Bishop JM. RNA interference (RNAi) with RNase III-prepared siRNAs. Methods in molecular biology (Clifton, NJ). 2004; 252:471-82. PubMed PMID: 15017072. Epub 2004/03/16. eng.
[10] Geng CM, Ding HL. Double-mismatched siRNAs enhance selective gene silencing of a mutant ALS-causing allele. ActaPharmacol Sin. 2008; 29(2):211-216. Cited in PubMed; PMID 18215350.
[11] Martinez MA, Gutierrez A, Arman MU, Blanco J, Parera M, Gomez J, et al. Suppression of chemokine receptor expression by RNA interference allows for inhibition of HIV-1 replication, AIDS. 2003; 17: S103-S105. Cited in PubMed; PMID 15080188.

[12] Kusov Y, Kanda T, Palmenberg A, Sgro J-Y, Gauss-Müller V. Silencing of Hepatitis A Virus Infection by Small Interfering RNAs, Journal of Virology. 2006; 80(11):5599-5610. doi:10.1128/JVI.01773-05. Cited in PubMed; PMID16699041.

[13] Jia F, Zhang YZ, Liu CM. A retrovirus-based system to stably silence hepatitis B virus genes by RNA interference, Biotechnology Letters, 2006; 28(20):1679-1685. Cited in Pubmed; PMID 16900331.

[14] Li YC, Kong LH, Cheng BZ, Li KS. Construction of influenza virus siRNA expression vectors and their inhibitory effects on multiplication of influenza virus. Avian diseases. 2005 Dec; 49(4):562-73. PubMed PMID: 16405000. Epub 2006/01/13. eng.

[15] O’Connor TP, Crystal RG. Genetic medicines: treatment strategies for hereditary disorders. Nat Rev Genet. 2006; 7(4):261-76.

[16] Leung RK, Whittaker PA. RNA interference: From gene silencing to gene-specific therapeutics, Pharmacology &Therapeutics, 2005; 107(2):222-239. Cited in PubMed; PMID 15908010.

[17] Lieberman J, Song E, Lee SK, Shankar P. Interfering with disease: opportunities and roadblocks to harnessing RNA interference. Trends in molecular medicine. 2003 Sep; 9(9):397-403. PubMed PMID: 13129706. Epub 2003/09/18. eng.

[18] Zaas DW, Duncan MJ, Li G, Wright JR, Abraham SN. Pseudomonas invasion of type I pneumocytes dependent on the expression and phosphorylation of caveolin-2. J Biol Chem 2005 Feb; 280(6):4864-72. PubMed PMID: 15545264. Epub 2004/11/17. eng.

[19] Rodriguez D, Gauthier F, Bertini E, Bugiani M, Brenner M, N'Guyen S, et al. Infantile Alexander disease: spectrum of GFAP mutations and genotype-phenotype correlation. Am J Hum Genet. 2001 Nov; 69(5):1134-40. PubMed PMID: 11567214. Pubmed Central PMCID: PMC1274357. Epub 2001/09/22. eng.

[20] Brenner M, Johnson AB, Boespflug-Tanguy O, Rodriguez D, Goldman JE, Messing A. Mutations in GFAP, encoding glial fibrillary acidic protein, are associated with Alexander disease. Nature genetics. 2001 Jan; 27(1):117-20. PubMed PMID: 11138011. Epub 2001/01/04. eng.

[21] Flint D, Li R, Webster LS, Naidu S, Kolodny E, Percy A, et al. Splice Site, Frameshift and Chimeric GFAP Mutations in Alexander Disease. Human mutation. 2012 04/30; 33(7):1141-8. PubMed PMID: PMC3674965.

[22] Messing A, Brenner M, Feany MB, Nedergaard M, Goldman JE. Alexander Disease. The Journal of Neuroscience. 2012; 32(15):5017.

[23] Brenner M, Johnson AB, Boespflug-Tanguy O, Rodriguez D, Goldman JE, Messing A. Mutations in GFAP, encoding glial fibrillary acidic protein, are associated with Alexander disease. Nature genetics. 2001 Jan; 27(1):117-20. PubMed PMID: 11138011. Epub 2001/01/04. eng.

[24] Messing A, LaPash Daniels CM, Hagemann TL. Strategies for treatment in Alexander disease. Neurotherapeutics: the journal of the American Society for Experimental NeuroTherapeutics. 2010 Oct; 7(4):507-15. PubMed PMID: 20880512. Pubmed Central PMCID: PMC2948554. Epub 2010/10/01. eng.

[25] Quinlan RA, Brenner M, Goldberg JE, Messing A, GFAP and its role in Alexander Disease. Experimental cell research. 2007 04/06; 313(10):2077-87. PubMed PMID: PMC2702672.

[26] Sahrawat, TR, Bhalla, S. Identification of Critical Target Protein for Cystic Fibrosis using Systems Biology Network Approach. Int. J. Bioautomation. 2013; 17: 227-240.

[27] Salgado D, Desvignes J, Rai G, et al. UMD-Predictor: A High-Throughput Sequencing Compliant System for Pathogenicity Prediction of any Human cDNA Substitution, Human Mutation. 2016; 37(5):439-446.doi:10.1002/humu.22965. Cited in PubMed; PMID 26842889

[28] Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function, Nucleic Acids Research. 2003; 31(13):3812-3814. Cited in PubMed; PMID:1284425.

[29] Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels, Bioinformatics. 2015; 31(16):2745-2747. doi:10.1093/bioinformatics/btv195. Cited in PubMed; PMID 25851949.

[30] Yuan B, Latek R, Hossbach M, Tuschl T, Lewitter F. siRNA Selection Server: an automated siRNA oligonucleotide prediction server. Nucleic Acids Research. 2004; 32(Web Server issue):W130-W134. doi:10.1093/nar/gkh366. Cited in PubMed; PMID 15215365.

[31] Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA Websuite. Nucleic Acids Research. 2008 July 1, 2008; 36(suppl 2):W70-W4.

[32] Ding Y, Chan CY, Lawrence CE. RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble. RNA 2005; 11:1157-1166.
[33] Nam TS, Kim JH, Chang CH, Yoon W, Jung YS, Kang SY, et al. Identification of a novel nonsense mutation in the rod domain of *gfap* that is associated with Alexander disease. European Journal of Human Genetics. 2015 Apr 23(1):72-78. PubMed PMID: PMC4266748.

[34] Bongcam-Rudloff E, Nister M, Betsholtz C, Wang JL, Stenman G, Huebner K, et al. Human glial fibrillary acidic protein: complementary DNA cloning, chromosome localization, and messenger RNA expression in human glioma cell lines of various phenotypes. Cancer Research. 1991 Mar 01;51(5):1553-60. PubMed PMID: 1847665. Epub 1991/03/01. eng.

[35] Li R, Messing A, Goldman JE, Brenner M. GFAP mutations in Alexander disease. International Journal of Developmental Neuroscience. 2002; 20(3–5):259-68.

[36] Brenner M, Johnson AB, Boespflug-Tanguy O, Rodriguez D, Goldman JE, Messing A. Mutations in GFAP, encoding glial fibrillary acidic protein, are associated with Alexander disease, Nat Genet. 2001; 27(1):117-120.Cited in PubMed; PMID 11138011.

[37] Perng MD, Wen SF, et al. The Alexander Disease-Causing Glial Fibrillary Acidic Protein Mutant, R416W, Accumulates into Rosenthal Fibres by a Pathway That Invokes Filament Aggregation and the Association of αβ-Crystallin and HSP27, American Journal of Human Genetics. 2006; 79(2):197-213. Cited in PubMed; PMID 16826512.

[38] Rodriguez D, Gauthier F, Bertini E, et al. Infantile Alexander Disease: Spectrum of *GFAP* Mutations and Genotype-Phenotype Correlation, American Journal of Human Genetics. 2001;69(5):1134-1140.Cited in PubMed; PMID 11567214.

[39] Namekawa M, Takiyama Y, Aoki Y, Takayashiki N, Sakoe K, Shimazaki H, et al. Identification of GFAP gene mutation in hereditary adult-onset Alexander's disease. Annals of neurology. 2002 Dec; 52(6):779-85. PubMed PMID: 12447932. Epub 2002/11/26. eng.