Journal of Chemical and Pharmaceutical SciencesISSN: 0974-2115

In-silico modelling of DTNBP1, NRG1 and AKT1 gene target of Schizophrenia for protein structure analysis

Neema Tufchi*, Kumud Pant, Syed Mohsin Waheed
Department of Biotechnology, Graphic Era University, Dehradun, Uttarakhand, India
***Corresponding author: E-Mail: neematufchi@gmail.com**

ABSTRACT

Schizophrenia (SZ) is a chronic mental and heritable disorder mainly recognized by neurophysiological impairment and neuropsychological abnormalities and is strongly associated with three genes viz; dysbindin (DTNBP1), neuregulin (NRG1) gene and AKT1 (v-akt murine thymoma viral oncogene homolog1). The three genes are present on different chromosomes besides other genes associated with Schizophrenia, but the exact cause for this disease is not known. In the present work, through bioinformatics techniques the analysis of the above three genes was performed. The primary analysis (molecular weight, pI etc.) of proteins obtained from these genes was done through Protparam server, for secondary and tertiary structural analysis SOPMA, SOSUI, TMHMM were used. Homology modelling was done using Phyre 2 server with structural validation through Rampage server. These 3D models can further be used for drug development against this particular disease.

KEY WORDS: Schizophrenia, Dysbindin, Neuregulin, AKT1, SOPMA, SOSUI, TMHMM.

1. INTRODUCTION

Schizophrenia is a severe type of mental illness which is increasingly gaining importance. It is also called as split personality disorder affecting about 1% of the population ranging from 15-35 years. A recent survey analysis estimated that the risk of Schizophrenia in males is 40% higher than in females. The patients with schizophrenia have several symptoms including delusions, hallucinations, disorganized speech, social withdrawal, lack of energy, incoherence and physical agitation, absence of normal emotions. The inclination of Schizophrenia is affected by several environmental factors, polygenic components, and their interactions. The main cause of the disease is still unknown but there are some evidences that show genetic factors have a role in developing Schizophrenia. In previous studies it has been found that this disease causes abnormalities in brain structure including enlargement of the ventricles. Neuroleptics like Chlorpromazine, Melperone, Quetiapine, Haloperidol etc. are used to treat Schizophrenia, which acts by blocking the action of the neurotransmitter dopamine at the D2 receptor. As there is a higher concentration of D2 in the schizophrenic patients, so the antipsychotic drugs mainly target the D2 receptor. The three putative Schizophrenia susceptibility genes have been identified viz; dysbindin (DTNBP1), neuregulin (NRG1) gene and AKT1 (v-akt murine thymoma viral oncogene homolog 1), also known as protein kinase B. Dysbindin-1 is a protein which encodes the dystrobrevin-binding protein 1 gene and is situated in synaptic sites of brains in human within chromosome 6p22.3. However their roles are still unknown but its location suggested that the genetic variations in this gene has greater effect in human cognitive abilities and effects on postsynaptic structure and function. Reduced expressions of this gene have been reported in the hippocampus and prefrontal cortex (PFC) of schizophrenic patients also variations in the gene affect mRNA expression and processing. Dysbindin is involved in intracellular protein trafficking involving lysosomes and related organelles and it is vital for synaptic homeostasis. Neuregulin 1 (NRG1) is another gene responsible for causing schizophrenia which is present on chromosome 8p. NRG1 is a growth factor which is important in nervous system development and functioning including synaptogenesis, neuronal migration, neuron-glia interaction, myelination and neurotransmission. NRG1 is involved in glutamatergic signaling and plays an essential role in neural development. The genetic variation in NRG1 is uncertain as SNPs associated with the disease are noncoding but it is possible that these non-coding SNPs are functionally associated with the disease, and altering the gene expression. The increased NRG1 expression is found in person having schizophrenia.

AKT1 gene is linked with many cellular pathways and its deregulation in the pathway leads to many incurable human diseases and several neurodegenerative brain disorders, and is present on chromosome 14q32.32. The lower concentration of this gene is responsible for causing Schizophrenia in many cases. AKT1 is known for cell proliferation, synaptic development, and protein trafficking. The function of AKT1 is co regulated by serotonin and dopamine type-2 receptor which works on the same mechanism as that of kinase-3 beta and ß-arrestin-2. These pathways establish targets for mood altering and antipsychotic drugs. AKT1 promotes cell proliferation and survival by organizing many genes of apoptosis and cell survival.

In this study three genes for Schizophrenia were selected and further analysed using many computational tools. *In-silico* techniques is an emerging field that has been now-a-days used to retrieve, analyse, and uploads data from different sources. In other context, it utilizes the information for designing many computational models or simulations that can be further used for suggesting hypothesis, making predictions and finally provides many advances in therapeutics and medicines.

2. MATERIALS AND METHODS

Databases: The fasta format of all the three genes responsible for Schizophrenia was retrieved from National Center of Biotechnology Information (NCBI). The NCBI is one of the most used database and an important resource for bioinformatics tools and services. The sequences were then searched for Expressed Sequence Tag (ESTs) in NCBI. ESTs are the mRNA sequence fragments or conserved regions which can be derived from cDNA libraries. Uptill now 11, 37, 27 ESTs have been developed from Dysbindin, Neuregulin, and AKT1 gene respectively.

Table.1. ESTs of Genes for Schizophrenia

Accession no.	Protein Name	Length
AF394226.1	Dysbindin (DTNBP1)	1349bp
NM_001165972.1	Neuregulin (NRG1)	3689 bp
gbKR710120.1	AKT1	1572bp

Nucleotide conversion: The ESTs were first converted into protein using expasy translation tool for further analysis. Translate is a tool which translates a nucleotide (DNA/RNA) sequence into a protein sequence.

Primary structure prediction: The primary structures of all the three translated proteins were computed using Protparam server. Protparam server allows the user to compute physical and chemical parameters of the given protein sequences. The parameter for primary prediction includes amino acid composition, atomic composition, molecular weight, extinction coefficient, theoretical pI, aliphatic index, instability index, and grand average of hydropathicity.

Domain Identification: The EST sequences of all the three genes were analyzed for domain identification. Domains are the conserved part of protein which can function independently and is one of the important levels to understand protein function. Prosite is a webserver which identifies the conserved portion of protein sequences. This webserver identifies these regions by using two signatures viz; generalized profiles and regular expressions.

Secondary Structure Prediction: The secondary structures of the proteins were predicted using online server SOPMA (Self-Optimized Prediction Method with Alignment). The server requires fasta format as input and estimates various parameters like alpha-helix, beta sheet and coil. SOPMA method predicts 69.5% of amino acids from a database containing 126 chains of non-homologous proteins.

The proteins were then predicted for their functional characterization using SOSUI and TMHMM tools. These tools are used to characterize the nature of protein whether soluble or transmemberane. The accuracy of transmembrane regions in SOSUI's is very high (>90%). Transmemberane proteins and soluble proteins, both have different physiochemical properties. There are very less PDB templates for membrane proteins but the energy required in threading of proteins are derived from globular proteins and there is no role of membrane proteins in threading.

TMHMM is a protein membrane topology method which works on hidden Markov model and it can differentiate between soluble and membrane proteins. More than 4000 protein sequences can be submitted in fasta format. $\alpha + \beta = \chi.$ (1)

Tertiary Structure Prediction: Modelling was done using automated server Phyre 2, which is a suite of tools for prediction and analysis of protein structure as well as their functions and mutations. It involves advanced homology detection methods for building 3D models, also predicts the ligand binding sites and study the effect of amino acid for a given protein sequence. The modeled structure was then visualized and analysed using visualization software Rasmol.

Model Validation: Model validation is an important step in In-silco protein analysis. The models were validated using Rampage server. The PDB file of the protein were uploaded and checked for Ramachandran Plot. The obtained results were then saved for analysis.

3. RESULTS AND DISCUSSION

The ESTs of three genes which are known to cause schizophrenia were retrieved from NCBI database and for further analysis these nucleotide sequences were converted into protein sequences. The translate tool is used for this purpose, which converts DNA/RNA sequences to protein sequences. The figures 1 & 2 shows the ESTs of AKT1 gene and its translated form.

The protein sequences of all the three genes were then analysed for domain identification using Prosite. Table.1 illustrates the functional region of the protein sequence with their score.

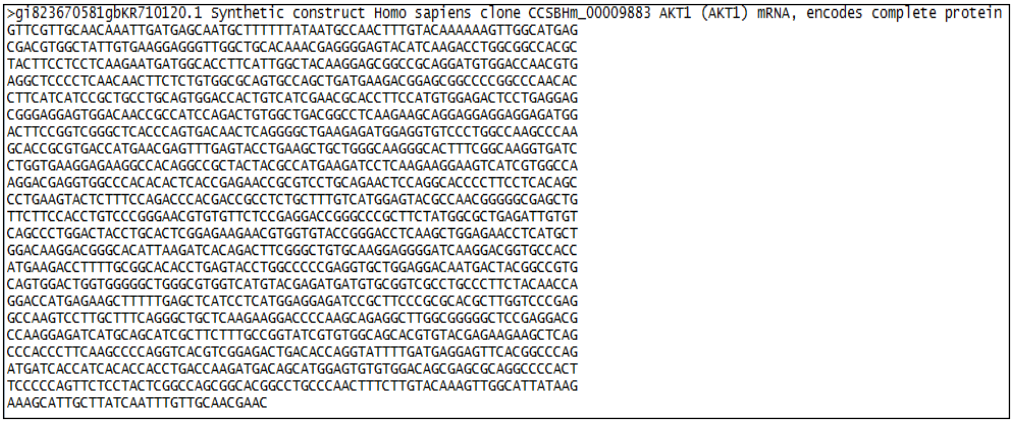

Fig.1. Screenshot of Nucleotide sequence of AKT1 gene

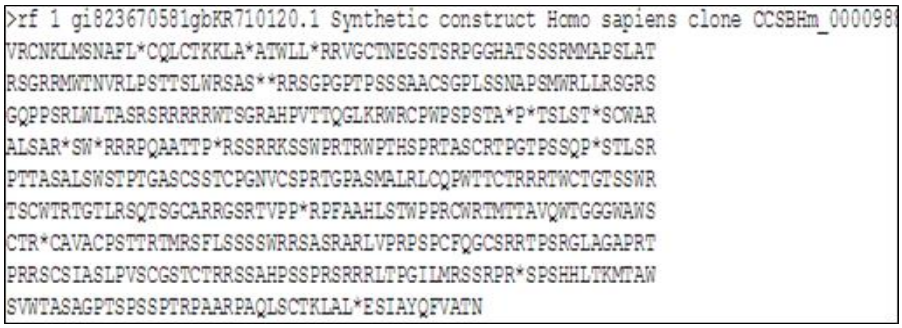

Fig.2. Screenshot of Translated protein of AKT1 gene

Table.2. Domain Identification using Prosite

Gene Name	Domains	Score
Dysbindin	No hits found	0.0
Neuregulin 1 (NRG1)	HFKPCRdkdlAYCLNDGECFVIETltGSHKHCRCKEGYQGVRCD	13.501
AKT1	FEYLKLLGKGTFGKVILVKEKATGRYYAMKILKKEVIVAKDEVAHTL TENRVLQNSRHPFLTALKYSFQTHDRLCFVMEYANGGELFFHLSRER VFSEDRARFYGAEIVSALDYLHSEKNVVYRDLKLENLMLDKDGHIKI TDFGLCKEGIKDGATMKTFCGTPEYLAPEVLEDNDYGRAVDWWGLG VVMYEMMCGRLPFYNQDHEKLFELILMEEIRFPRTLGPEAKSLLSGLL KKDPKQRLGGGSEDAKEIMQHRF	49.135

The domain of the dysbindin was not found, illustrating that there are no conserved region found in this sequence. Primary structures of these proteins were studied using Protparam server. The primary structure information analysis is essential for determining/ predicting properties of protein in detail and is the first step for protein analysis.

Table.3. Primary analysis of Protein Sequences

Protein Name	Mol wt.	pI	Instability index	Aliphatic index	GRAVY
Dysbindin	50174.2	4.88	45.84	74.34	-0.651
Neuregulin1 (NRG1)	134413.9	9.27	51.56	78.96	-0.157
AKT1	61463.3	6.00	37.03	73.24	-0.499

In the above table it is revealed that molecular weights of three proteins are 50174.2 (Dysbindin), 134413.9 (Neuregulin1), and 61463.3 (AKT1). The pI of two proteins was less than 7 which indicated that they are acidic in nature and one protein was greater than 7 showing its basic nature. The instability index (II) of dysbindin and AKT1 was found to be less than 50 showing that protein is stable but the instability index of gene Neuregulin1 was 51.56 and thus is unstable in nature. The aliphatic index of genes ranged between 73.24 to 78.76. The range of GRAVY (Grand Average of Hydropathicity) of Schizophrenia genes was found to be -0.651 to -0.157. The lowest value of Gravy indicates the possibility of better interaction with water.

The secondary structure analysis of schizophrenia genes were predicted by SOPMA server which revealed the helical structure of the protein including alpha helix, beta turn, and random coil. In all the three genes alpha helix shows highest value followed by random coil, extended strand and beta turn.

Table.4. Secondary prediction of the Proteins

Secondary structures	Dysbindin	Neuregulin 1	AKT1
Alpha helix	56.79%	26.63%	38.42%
310 helix	0.00%	0.0%	0.00%
Pi helix	0.00%	0.00%	0.00%
Beta bridge	0.00%	0.00%	0.00%
Extended strand	9.35%	21.10%	19.59%
Beta turn	6.24%	8.16%	9.98%
Bend region	0.00%	0.00%	0.00%
Random coil	27.62%	44.11%	32.02%
Ambiguous states	0.00%	0.00%	0.00%
Other states	0.00%	0.00%	0.00%

SOSUI and TMHMM predicted that AKT1, and dysbindin were soluble protein, whereas Neuregulin1 was found to be a membrane protein with 23 residues (Table.5 and Figure.3).

Table.5. Transmembrane Region predicted by SOSUI

Protein	N terminal	Trans membrane Region	C terminal	Type	Length
Neuregulin1	81	VVPLFIGFIGLGLSLMLLKWIVV	103	Primary	23
	371	QVLSISCIIFGIVIVGMFCAAF	392	Primary	22
	939	SGGAFLSFCVPVCTCVCERVCVW	961	Primary	23
	970	CKHMCSLCVCVCMCARITLAKIC	992	Primary	23
	1004	IISLLDLFTLNFSIK	1018	Secondary	15
	1090	TVTICFLIILITIYFIASLLCVD	1112	Primary	23
	1142	TVNCVVCSFGIKQYFYHCGELKP	1164	Secondary	23
	1183	ISVHISSLFPIIGGVELLCLHCF	1205	Primary	23

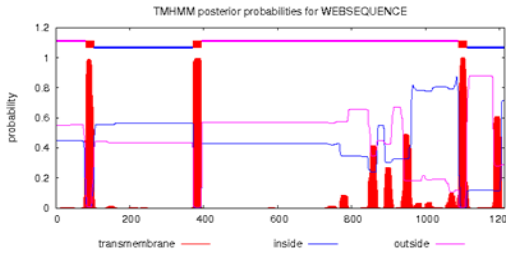

Fig.3. TMHMM showing the transmembrane region of Neuregulin1.

The tertiary structures of three proteins were modelled using Phyre2 (Figure.4). The proteins were visualized and analysed with the help of Rasmol.

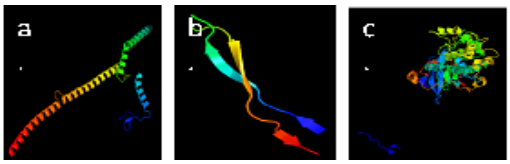

Fig.4. Modelled structures of all the three proteins (a) Dysbindin (b) Neuregulin (c) AKT1

For validation of modeled protein structure Ramachandran plot (RM plot) was drawn using Rampage server. This plot shows the relationship between phi and psi angles (dihyderal angles) of a particular protein. The protein structure is said to be stable if more than 95% of its residues fall in the favored region.

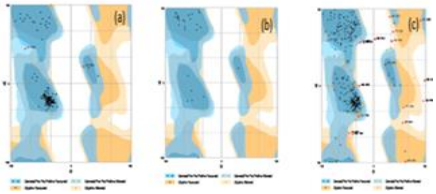

Fig.5. Ramachandran plot of the proteins (a)Dysbindin (b)Neuregulin (c) AKT1

In above figures it is indicated that Dysbindin, Neuregulin and Akt1 have 98.4%, 100%, and 96.3% of residues lying in favored region. Thus these modeled structures are validated and hence can be used for drug development and designing.

4. CONCLUSION

The cause behind schizophrenia is unknown, but it has a major genetic component so there is an urgent need to develop some therapeutics against this disease. In the present study three genes susceptible in Schizophrenia were

selected, translated, and analyzed. Their primary, secondary and tertiary structure predictions were determined using various bioinformatics tools which provide very useful and basic information about the genes and also the modeling of proteins was performed. The protein models thus generated can be utilized for performing various studies including discovery and development of drugs against Schizophrenia.

5. ACKNOWLEDGEMENT

The authors express the deep sense of gratitude to the Department of Biotechnology of Graphic Era University for all the support, assistance, and constant encouragements to carry out this work.

REFERENCES

Bagchi P, Kar A, and Vinobha C.S, Establishing an *in-silico* ayurvedic medication towards treatment of Schizophrenia, Int J Sys Biol, 1, 2009, 46-50.

Balu D.T, Carlson G.C, Talbot K, Kazi H, Hill-Smith T.E, Easton R.M, Birnbaum M.J, and Lucki I, Akt1 Deficiency in Schizophrenia and Impairment of Hippocampal Plasticity and Function, PMC, 22, 2012, 230–240.

Bhargava K, Nath R, Seth P.K, Pant K.K, and Dixit R.K, Molecular Docking studies of D2 Dopamine receptor with Risperidone derivatives, Bioinformation, 10, 2014.

Emamian E.S, AKT/GSK3 signaling pathway and schizophrenia, Front. Mol. Neurosci., 2012.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M.R, Appel R.D, Bairoch A, Protein Identification and Analysis Tools on the ExPASy Server, in John M. Walker, The Proteomics Protocols Handbook, Humana Press, 2005, 571-607.

Geourjon C, and Deleage G, SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments, Bioinformatics, 11, 2006, 681-684.

Harrison P.J, and Law A.J, Neuregulin 1 and Schizophrenia: Genetics, Gene Expression, and Neurobiology, Biological Psychiatry, 60, 2006, 132-140.

Hirokawa T, Boon-Chieng S, and Mitaku S, SOSUI: classification and secondary structure prediction system for membrane proteins, Bioinformatics, 14, 1998, 378-379.

Joachim Hallmayer, Getting our AKT together in schizophrenia?, Nature Genetics, 36, 2004.

Kelley L.A, The Phyre2 web portal for protein modeling, prediction and analysis, Nature Protocols, 10, 2015, 845-858.

Law A.J, Lipska B.K, Weickert C.S, Hyde T.M, Straub R.E, Hashimoto R, Harrison P.J, Kleinman J.E, and Weinberger D.R, Neuregulin 1 transcripts are differentially expressed in schizophrenia and regulated by 5 SNPs associated with the disease, PNAS, 103, 2006, 6747-6752.

Lovell S.C, Davis I.W, Arendall III W.B, De Bakker P.I.W, Word J.M, Prisant M.G, Richardson J.S, and Richardson D.C, Structure validation by Calpha geometry: phi,psi and Cbeta deviation, Proteins: Structure, Function & Genetics, 50, 2002, 437-450.

Numakawa T, Yagasaki Y, Ishimoto T, Okada T, Suzuki T, Iwata N, Ozaki N, Taguchi T, Tatsumi M, Kamijima K, Straub R.E, Weinberger D.R, Kunugi H, and Hashimoto R, Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia, Hum. Mol. Gen., 13, 2004, 2699–2708.

Owen J.M, Williams N.M, and Donovan M.C.O, Dysbindin-1 and schizophrenia: from genetics to neuropathology, J. Clin. Invest., 113, 1255–1257.

Papaleo F, and Weinberger D.R, Dysbindin and Schizophrenia: It's dopamine and glutamate all over again, Biol. Psychiatry, 69, 2011, 2-4.

Parkinson J, and Blaxter M, Expressed Sequence Tags: An Overview, Molecular Biology, 533, 2009, 1-12.

Sehgal S.A, Mannan S, Kanwa S, Naveed I, and Mir A, Adaptive evolution and elucidating the potential inhibitor against schizophrenia to target DAOA(G72) isoforms, Drug Design, Develop and Therapy, 2015.

Sigrist C.J.A, Castro E.D, Cerutti L, Cuche B.A, Hulo N, Bougueleret B.A, and Xenarios I, New and continuing developments at PROSITE, Nucleic Acids Res, 6, 2012.

Stefansson H, Steinthorsdottir V, Thorgeirsson T.E, Gulcher J.R, and Stefansson K, Neuregulin 1 and Schizophrenia, Ann Med., 36, 2004, 62-71.