

## Modelling Approach to Design The Structure of cln8 and Targeting it for Drug Binding, A Means to treat Epilepsy

<sup>1</sup>Subhashree Hota \*, <sup>2</sup>Pritam Kumar Panda, <sup>3</sup>Suman Mishra, <sup>4</sup>Samaya Chakra Pradhan, <sup>5</sup>Biswajit Satapathy, <sup>6</sup>Piyus Pritam  
<sup>1,2,3,4,5,6</sup>BJB Autonomous College, Utkal University  
Email ID: \*mail2subhashree143@gmail.com

Received: March 3<sup>rd</sup> 2014, Accepted: March 12<sup>th</sup> 2014, Published: May 1<sup>st</sup> 2014

### Abstract:

Epilepsy is one of the genetic diseases that is existing from the past. Though it is less observed statistically, it has a very bad effect on the victim. There are several approaches to resolve certain symptoms of the disease. Drug targeting is one of the routinely used approaches for the same. The current work involves the study of the protein CLN8, one of the major causes of epilepsy. The Sequence, Structural and functional annotation of the protein was performed using several insilico tools. As the structure of the CLN8 protein is not available in the literature, the work included modeling approach using modeler to design the appropriate structure for the protein. The protein thus designed is validated and is used as a receptor in docking. The chemicals that can act against the complications of Epilepsy were screened from CHEBI and Pubchem. The selected chemicals were filtered according to their Druglikeness, Toxicity and Ambiguity. The final list of selected chemicals was allowed to dock with the receptor using Hex Software followed by energy evaluation.

**Key Words:** *Ceroid lipo fuscinos neural 8(CLN8 Protien); Computational Analysis; Homology Modelling; Rasmol Drug designing.*

### Introduction:

#### **EPILEPSY:**

Epilepsy is a brain disorder in which clusters of nerve cells, or neurons, in the brain sometimes signal abnormally. Neurons normally generate electrochemical impulses that act on other neurons, glands, and muscles to produce human thoughts, feelings, and actions. In epilepsy, the normal pattern of neuronal activity becomes disturbed, causing strange sensations, emotions, and behaviour, or sometimes convulsions, muscle spasms, and loss of consciousness. During a seizure, neurons may fire as many as 500 times a second, much faster than normal. In some people, this happens only occasionally; for others, it may happen up to

hundreds of times a day. Epilepsy is not contagious and is not caused by mental illness or mental retardation. Some people with mental retardation may experience seizures, but seizures do not necessarily mean the person has or will develop mental impairment. Many people with epilepsy have normal or above-average intelligence. Famous people who are known or rumored to have had epilepsy include the Russian writer Dostoyevsky, the philosopher Socrates, the military general Napoleon, and the inventor of dynamite, Alfred Nobel, who established the Nobel Prize. Several Olympic medalists and other athletes also have had epilepsy. Seizures sometimes do cause brain damage, particularly if they are severe. However, most seizures do not seem to have a detrimental effect on the brain. Any changes that do occur are usually subtle, and it is often unclear whether these changes are caused by the seizures themselves or by the underlying problem that caused the seizures. The major gene having a vital role in epilepsy was found to be CLN8. CLN8 (ACCESSION NO: NP\_061764). The official name of this gene is “ceroid-lipofuscinosis, neuronal 8 (epilepsy, progressive with mental retardation).”

**NORMAL FUNCTION OF GENE:** The CLN8 gene provides instructions for making a protein whose precise function is not known but is thought to play a transport role within cells. Specifically, the CLN8 protein likely helps to transport materials in and out of a cell structure called the endoplasmic reticulum, which is involved in protein production, processing, and transport. Based on the structure of the CLN8 protein, it may also be involved in regulating the levels of fats (lipids) in cells. Neuronal ceroid lipofuscinosis 8; CLN8 and Northern epilepsy : autosomal recessive inheritance, age of onset 2-7 years, clinically characterized by psychomotor deterioration, epileptic seizures, myoclonus, ataxia, loss of vision and premature death ,clinically characterized by psychomotor deterioration, epileptic seizures, myoclonus, ataxia, loss of vision and

premature death, accumulation of characteristic storage material in cells, Northern epilepsy: an allelic disorder caused by a Finnish founder mutation in *CLN8*; age of onset 5-10 years, characterized by progressive epilepsy and mental retardation, motor clumsiness, behavioral problems and a protracted disease course.

**Tools and databases:**

**NCBI:** (URL: [www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/))

The **National Center for Biotechnology Information (NCBI)** is one of the most widely used primary database that collects the information regarding genes, proteins, domains, homologues, related literatures etc. The most commonly employed use of NCBI is in the collection of the sequence and functions of genes and proteins.

**GENECARDS:** (URL: <http://www.genecards.org/>)

**GeneCards** is a database of human genes that provides genomic, proteomic, transcriptomic, genetic and functional information on all known and predicted human genes.<sup>[1][2][3][4]</sup> Gene cards can be used for the collection of the complete gene information and its annotation that includes structural, functional, mutational, expression related data etc. Being a very user friendly database it has direct connection to several other databases.

**BLAST** (URL: [www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)):

In bioinformatics, Basic Local Alignment Search Tool, or BLAST, is an algorithm for comparing primary biological sequence information, such as the amino-acid sequences of different proteins or the nucleotides of DNA sequences. A BLAST search enables a researcher to compare a query sequence with a library or database of sequences, and identify library sequences that resemble the query sequence above a certain threshold.

**CLUSTAL-W** ([www.ebi.ac.uk/Tools/msa/clustalw](http://www.ebi.ac.uk/Tools/msa/clustalw))

**Clustal** is a widely used multiple sequence alignment programme. The algorithm would align two or more sequences under comparison and would check for the homologous regions among them. Based on the homology existing among the sequences their evolutionary distances are calculated and depicted as a dendrogram or tree. The tree is mostly used to trace the identity of an unknown sequence based on its homology with the known sequences.

**SDSC BIOLOGY WORKBENCH** (URL: <http://workbench.sdsc.edu/>)

The Biology WorkBench is a web-based tool for biologists that can enable the complete sequence, structural and alignment annotation of the gene and protein sequences. It is a free web based platform for the end users. The WorkBench allows biologists to search many popular protein and nucleic acid sequence databases. Database searching is integrated with access to a wide variety of analysis and modeling tools, all within a point and click interface that eliminates file format compatibility problems. The tools that are used from the workbench in the current work include TEXSHADE, BOXSHADE, CLUSTALDIST, CLUSTALW etc to annotate the query sequence.

**SMART** : (URL: <http://smart.embl-heidelberg.de/>)

**SMART:** This is a free online tool that has inbuilt search system to run internal Blast and analyze the Domains, Motifs etc present in a given input sequence. It also provides information like the total number of domains, their location and each of their function in the cell.

**STRING** (URL: <http://string-db.org/>)

STRING is a database of known and predicted protein interactions. The interactions include direct (physical) and indirect (functional) association. STRING quantitatively integrates interaction data from these sources for a large number of organisms, and transfers information between these organisms where applicable. The database currently covers 5'214'234 proteins from 1133 organisms.

**Primary Structure Prediction Tool:**

**PROTPARAM** :

(URL: <http://web.expasy.org/protparam/>)

ProtParam (References / Documentation) is a tool which allows the computation of various physical and chemical parameters for a given protein stored in Swiss-Prot or TrEMBL or for a user entered sequence. The computed parameters include the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) (Disclaimer).

## Secondary and Tertiary Structure Prediction Tools:

**SOPMA** (URL: [http://npsa-pbil.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_sopma.html](http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html))

**CPH** (URL: <https://www.cph.org>)

To analyze the structural conformations of a protein inside the cell several online tools can be employed. The secondary structural classification can be made with the help of a tool SOPMA wherein the details like the concentration of the alpha helix, beta sheets etc can be identified in the given sequence. The tertiary structure of the protein can be detected using the tool CPH. This tool provides the list of PDB id's among which the best id must be selected. The structure can later be **Visualized Using Various Softwares**.

## SAVESSERVER

(URL: [nihserver.mbi.ucla.edu/SAVES/](http://nihserver.mbi.ucla.edu/SAVES/))

SAVES is a server for analyzing protein structures for validity and assessing how correct they are. It utilizes 6 programs for doing this:

- PROCHECK
- WHAT\_CHECK
- ERRAT
- VERIFY\_3D
- PROVE
- SFCHECK

This server is designed to run all of the programs and present the user with all of the data on a single page.

## Softwares Used:

### A. MODELLER:

**MODELLER** is a computer program used in producing homology models of protein tertiary structures as well as quaternary structures (rarer).<sup>[1][2]</sup> It implements a technique inspired by nuclear magnetic resonance known as *satisfaction of spatial restraints*, by which a set of geometrical criteria are used to create a probability density function for the location of each atom in the protein.

### B. HEX:

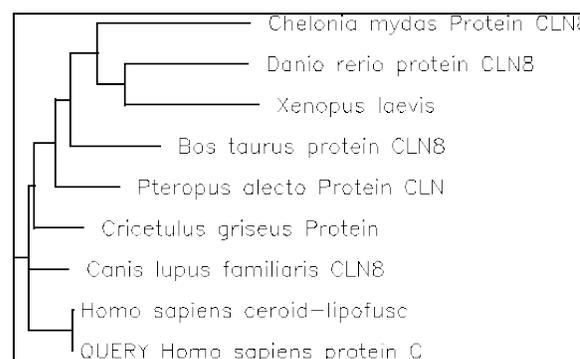
*Hex* is an interactive **Protein docking and molecular Superposition** program, used for the protein protein docking and protein ligand docking. The quality of the docking can be analyzed based on the docking energy values obtained.

## Result and discussion:

Sequence retrieval from Gene Cards: The sequence of the protein has been retrieved from the uni prot database and the basic annotation was performed using the gene cards data base. The complete name of the protein was found to be ceroid-lipofuscinosis, neuronal 8, the gene is found to encode a transmembrane protein postulated to function in lipid synthesis, transport, or sensing. The protein localizes to the endoplasmic reticulum (ER), and may recycle between the ER and ER-Golgi intermediate compartment.

## Phylogenetic Analysis (using Clustal-W):

**Figure 1: The Dendrogram showing the evolutionary relationship among the selected organisms.**



From the above phylogenetic tree it can be inferred that the query sequence of CLN8 was found to be homologous to all the other sequences under comparison.

## Domain Analysis (USING SMART)

### DOMAIN SEQUENCE:

```
VFWDLAATRAVFGVQSTAAGLWALLGDPVLH
ADKARGQQNWCWFHITTATGFFCFENVAV
HLSNLI FRTFDLFLVIHHLFAFLGFLGCLVNLQA
GHYLA MTTLLLEMSTPFTCVSWMLLK
AGWSESLFWKLNQWLMIHMFHCRMVLT YHM
WWVCFWHWDGLVSSLYLPHLTLFLVGLALL
TLIINPYWTHKKTQQLLPVD
```

SMART has been used to identify the total number of domains in the protein and their functions. There was one domain present in the query with the sequence shown above.

**Fig 2: SMART result showing the Domian TLC of the selected query sequence:**



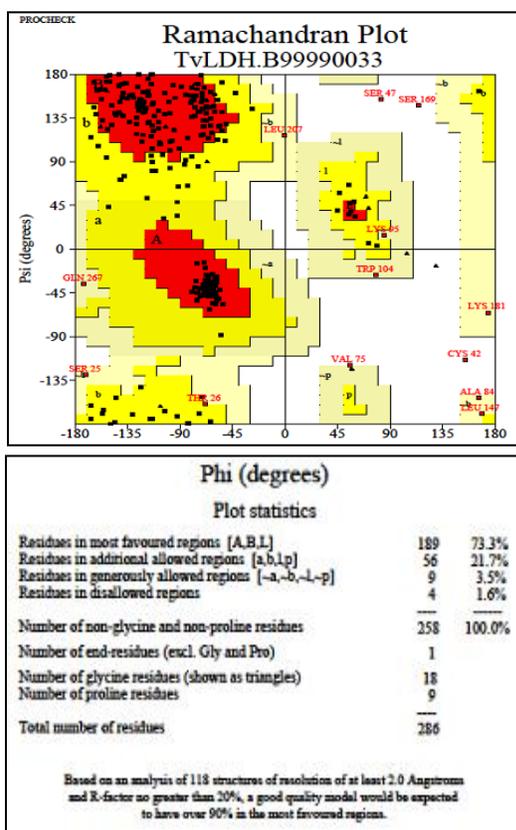
**Structure Designing Using Homology Modeling:**

The tertiary structure prediction tools like CPH, Hhpred, phyre could not provide the perfect structure for the query protein (CLN8). Thus homology modeling was opted for the modeling of the protein structure. MODELLER was used for the same.

**Multiple templates obtained for Modeller:**

5models were obtained with modeler after modeling the templates. The models were further validated using several approaches.

**Figure 3: Ramachandran Plot for assessing the quality of the Model obtained.**



From the above Ramachandran graph results it can be analyzed that the quality of the structure generated

was good with the allowed regions in 73% which indicates an acceptable model for the query sequence.

**DOCKING (Using HEX):**

The query protein CLN8 a cause of epilepsy was targeted in drug designing with the chemicals obtained from the pubchem database and were docked using HEX software to detect the binding affinity among the selected chemicals, a means to treat the disease symptoms. 9912049924

**Chemical Screening for Ligand Selection:**

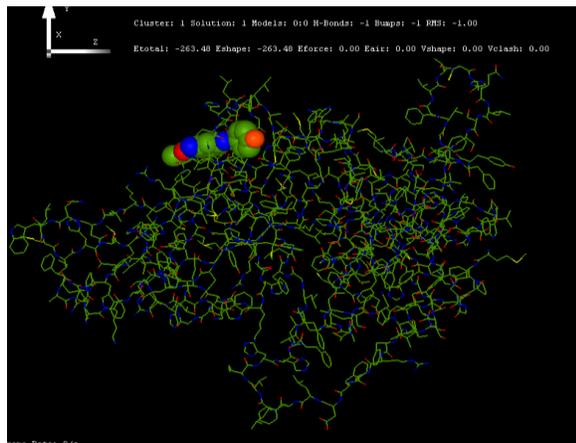
It is one of the major steps in drug design to screen the whole lot of chemicals related to the target protein that are available in various databases. Here in this work Pub Chem and ChEBI were screened for the chemicals that can act against the symptoms of epilepsy by binding to the query protein. The selected chemicals were screened for the properties like Druglikelinedd, ambiguity, Toxicity etc using various tools and softwares like Dundee proDrg Server, Moby ADME Server etc. The selected chemicals that passed the above filter tests were used as ligands to dock with the protein receptor selected. A total of FIVE Chemicals were used to dock with the CLN8 receptor whose energies were calculated using HEX Software. Based on the energy values the best ligand selected for docking with the CLN8 receptor was *Parampanel hydrate* With an e total of -263.48. Thus the above work proves that the ligand *Parampanel hydrate* can be used in the treatment of the symptoms of epilepsy by targeting the CLN8 receptor.

**Table 1: The summary of the selected chemicals after screening and their docking energy values with the selected CLN8 receptor**

S.No	Chemical	Docking Energy
1	<i>Ezogabine</i>	-248.99
2	<i>Parampanel hydrate</i>	-263.48
3	<i>Dipotassium clorazepate</i>	-248.98
4	<i>Clonazepam</i>	-261.92
5	<i>Clobazam</i>	-259.78

From the above table it can be inferred that the ligand 2, *Parampanel hydrate* can be selected as the best ligand due to its lowest e Value.

**Figure 4: Docking of the selected Ligand with the CLN8 Receptor:**



The above figure shows the docking of the selected ligand 2 with the receptor with an e total of -263.48, which is a good value for the binding affinity.

### Conclusion:

In the current work the major importance was given for the construction and validation of the protein structure by modeling using modeler approach. Once the structure was designed and optimized along with the quality evaluation, the structure was used for targeting with the selected ligand. To select the best drug for the receptor chemicals were collected from the PubChem and CHEBI databases. The collected chemicals were screened for the Drug likeliness, Toxicity and ambiguity check. The selected 5 chemicals were further used to dock with the pre designed receptor. After the docking analysis the evaluation of the energies was made based on which the best ligand selected was *Parampanel hydrate* with the e total of -263.48. The compound can be used as a drug for the treatment of epilepsy symptoms.

### Bibliography:

1. [www.ddtjournal.com/action/downloadaddoc.php?docid=574](http://www.ddtjournal.com/action/downloadaddoc.php?docid=574)
2. Miziak B, Błaszczyk B, Chrościńska-Krawczyk M, Danilkiewicz G, Jagiełło-Wójtowicz E, Czuczwar SJ.
3. “The problem of osteoporosis in epileptic patients taking antiepileptic drugs”
4. Schroeder AS, Homburg M, Warken B, Auffermann H, Koerte I, Berweck S, Jahn K,

Heinen F, Borggraefe I, “Prospective controlled cohort study to evaluate changes of function, activity and participation in patients with bilateral spastic cerebral palsy after Robot-enhanced repetitive treadmill therapy”

5. Walsh J, Thomas RH, Church C, Rees MI, Marson AG, Baker GA, “Executive functions and psychiatric symptoms in drug-refractory juvenile myoclonic epilepsy.”
6. [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)
7. [www.cmbi.ru.nl/edu/bioinf4/articles/homologymodeling.pdf](http://www.cmbi.ru.nl/edu/bioinf4/articles/homologymodeling.pdf)
8. Hex Protein Docking, [hex.loria.fr/](http://hex.loria.fr/)

\*\*\*\*\*