JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Genome Annotation of Brain Cancer and Structure Analysis by applying Drug Designing Technique

Vinita Kukreja[1], Uma Kumari[2]*

[1]Trainee at BPRI,

*Senior Bioinformatics Scientist (Bioinformatics Project and Research Institute), Noida - 201301, India

ABSTRACT: Cancer, one of the leading causes of deaths globally, in which majority of central nervous system (CNS) tumors are originates from glial cells leading to gliomas. Glioblastoma Multiforme (GBM) is an aggressive type tumor, arise from astrocytes cells. GBMs are classified into primary and secondary, these alters the core signaling pathways inducing cell proliferation and intensifying cell survival. To analyse structure properties domain and function of GBM, we examined the protein that has major role in the therapy resistance. Using computational approach we were able understand the molecular interaction between receptor and ligand. Graphical representation shows the similarities between query and database sequence. Using molecular graphic program we visualize our target protein in different form shapely, structure, spacefill which includes amino acid residues, protein secondary structure and size dependent atoms respectively. Molecular Docking of target protein determines the conformation and binding energy between ligand and receptor. Bioinformatics and virtual screening methods to help in identify drugs that can bind to specific targets, so as to promote the research and development of brain cancer drugs.

KEYWORDS: - Brain tumor, glioma, gene, targeted therapy, precision medicine, Docking, Structure Analysis, Panorama view.

INTRODUCTION

Cancer is the second most leading destructive disease of a mankind. In 2018, 18.1 million cancer cases were reported and 9.6 million people were died due to cancer worldwide. [1] In year 2020, approximately 19.3 million new cases and 10 million deaths were reported. [2] The growing rate of mortality due to cancer all over the world is a matter of concern. The greatest challenge regarding cancer is the cell variation leading to specific diagnosis and effectiveness of treatment. [3] The most common cancer types are skin, breast, lung, prostate, stomach and colon. Hallmark of cancer has six traits that helps for understanding the remarkable diversity of tumours. These traits are: 1) Self-sustaining growth signals, 2) Avoiding growth suppressors, 3) Tissue invasion and metastasis, 4) Allowing replicative immortality, 5) Inducing angiogenesis and 6) Evasion of cell death. [4] Brain tumor mostly originate from glial cell and it also occur because of metastatic tumor cells spreading throughout the body. [5] Worldwide, 308,102 new cases and 251,329 deaths were reported due to brain cancer in 2020. [2]

Glioblastoma Multiforme (GBM), the grade IV astrocytoma, is an aggressive type of tumour arising from astrocytes cell. The cause of glioblastoma is still unknown but it is the most malignant and chronic. [6-7] Glioblastoma can classified into primary (de novo) and secondary (arises from low-grade gliomas), they are histologically identical but are different in genomic profiling. While studying the cancer type Glioblastoma, The Cancer Genome Atlas construct a set of activated core signaling pathways: 1) Receptor tyrosine kinase (RTK)/Ras/phosphoinositide 3-kinase (PI3K) signaling pathway, 2) p53 pathway and 3) Retinoblastoma (RB) pathway. [8-11] Most of the genetic aberrations because of primary and secondary GBMs shows alteration in these core signaling pathways (i.e., RTK/RAS/PI3K alters almost 88% of GBMS, RB pathway alters 78% and p53 pathway alters almost 87%) inducing cell proliferation and increasing the chance of cell survival. [12] Primary GBM has particular alterations of genes which includes overexpression of epidermal growth factor receptor (EGFR), mutation in tumor suppressor gene phosphate and tensin homologue (PTEN), and chromosome 10q deletion. Genetic alterations in secondary GBM includes mutation in tumor suppressor gene p53 and retinoblastoma (RB), isocitrate dehydrogenase 1 and 2 (IDH1/2) alteration and loss of chromosome 19q. [13-14]

Based on molecular alterations and gene expression GBM classified into four subtypes that are: 1) Classical: Tumors have higher expression of EGFR gene, tumor protein 53 gene is rarely mutated and amplification of chromosome 7 has been seen in this subtype. 2) Mesenchymal: In this subtypes number of alterations are high in neurofibromin 1 (NF1) gene and less in EGFR gene. 3) Proneural: Tumors belongs to this subtype has high rates of mutation in a-type platelet-derived growth factor receptor (PDGFRA) gene, p53 gene and IDH1 gene. 4) Neural: This subtype represented by the expression of neuron marker but during clinical evaluation they represent themselves as normal cells. [15] IDH gene mutations were also predicted as a biomarkers of GBM. Isocitrate dehydrogenase 1 and 2 enzymes convert isocitrate into alpha-ketoglutarate, NAPDH and CO_2. IDH1 R132H mutation is point mutation of IDH1 amino acid 132, where arginine is replaced by histidine, this point mutation in IDH1 leads to the conversion of alpha-ketoglutarate to 2-hydroxyglutarate. These IDH1 alterations results in gene expression changes in human gliomas. [10, 16-17]

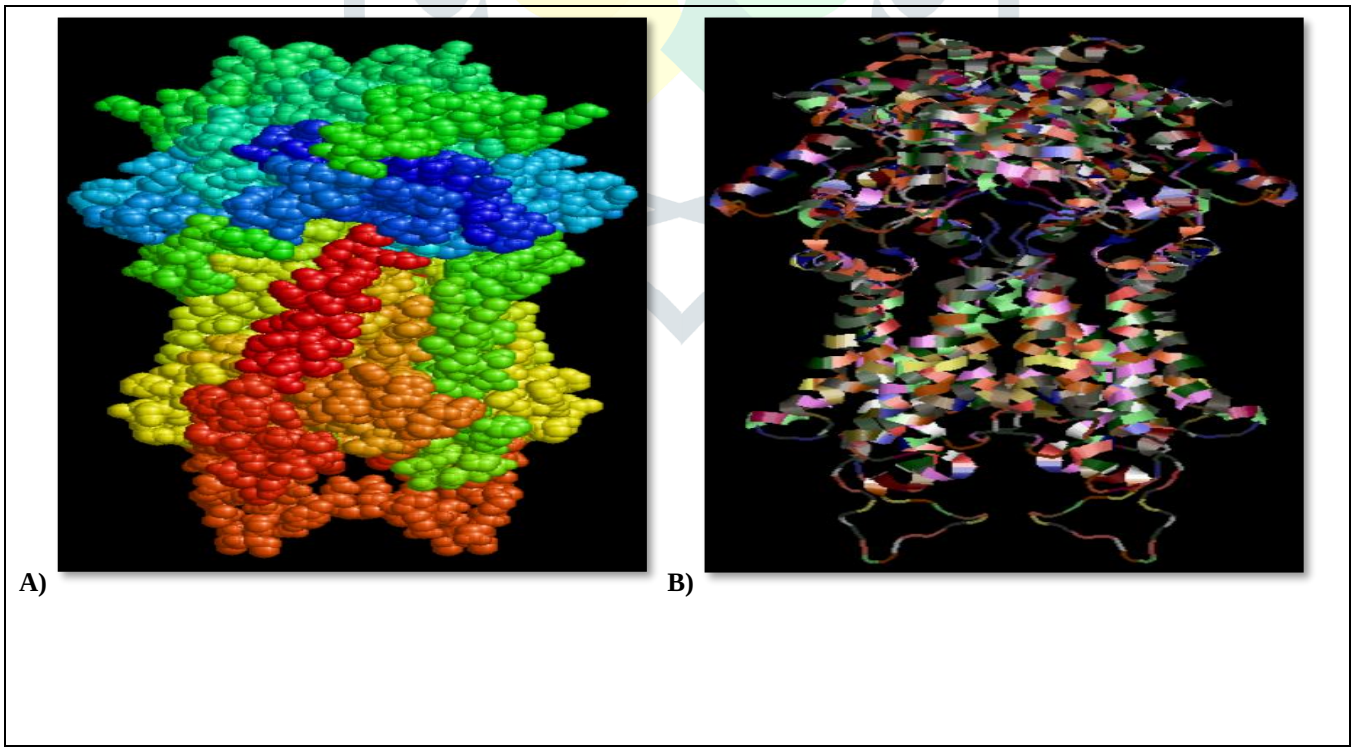
With the help of computational approach and virtual screening methods, we were able to identify the drugs that bind to our target protein and further understanding the developments of brain cancer drugs.

MATERIAL AND METHODS

Drugs that are used in this study have been taken from PubChem. PubChem is an open chemistry database that collects information on chemical structures, physical and chemical properties, bioactivities, identifiers, toxicity and many more. We have taken our protein of interest from Protein Data Bank (PDB). PDB has information about 3D structures of nucleic acids, proteins and complex molecules. PDB was established in 1971 as a joint venture between Cambridge Crystallographic Data Centre, UK and Brookhaven National Laboratory, US. ABCG2 protein (PDB ID **6HBU)**, is a transporter protein in humans belongs to ATP-binding cassette (ABC) family. ABCG2 protein used to transport drugs across the membranes having ATP molecules bound to the nucleotide-binding domains (NBDs) present at N-terminal as an energy source. It mainly transport drugs through the blood-brain-barrier (BBB) and maternal-foetus barrier. It also has major role in phenotypic drug resistance of some cancer cell lines. [18-19] Structural visualization of ABCG2 protein carried out by RasMol. RasMol is molecular graphic program used to study the visualization of biomolecules where "Display menu" is used to change the appearance of molecules and "Colour command" is used to identify different objects by their specific colour. Molecular modeling database (MMDB) holds information about 3D structure of biomolecules acquire from PDB. MMDB is used to analyze structure and identify developmental and functional relationship among them, to visualize 3D structure based on sequence-structure relationships, to identify specific chemical bound structure their interactions and active sites and many more. To obtain record of our protein, we entered PDB ID **6HBU** in query box of MMDB which shows the summary of structure. Various display options are present in summary page in which gene symbol leads to molecular information of ABCG2 protein to acquire the sequence of 6HBU. To analyze protein sequence we used Basic Local Alignment Search Tool (BLAST), it is used to find the similarities between sequences and to understand the evolutionary and functional relations between the sequences. To study the structure prediction, homology and phylogenetic analysis we align multiple sequence using Constraint-based Multiple Alignment Tool (COBALT). COBALT is used for the alignment of multiple protein sequence which sustained by collection of pairwise constraints using information about conserved domain and sequence similarity. The interaction between ABCG2 protein and selected ligands can be achieved by molecular docking, it also helps to identify binding affinity and conformational change in protein. Docking can also predict in which position a ligand binds to protein. Molecular graphics viewer tool is used to visualize real time high quality 3D image of ligand and protein, used to view and analyze molecular data and features the suitable binding sites for drugs.

RESULT AND DISCUSSION

Using RasMol we have seen different appearance of ABCG2 protein like in Figure 1: A) Spacefill command display atoms in van der waals sphere with the size depend upon the type of atom. B) Shapely colour command represent each amino acid residue in its given unique colour. C) Group colour command display residue colour based on their position in macromolecular chain i.e. N-terminal of protein is red in colour and C-terminal of protein is blue in colour. D) Structure colour scheme colours the molecule by secondary structure of protein therefore, alpha-helix represent as magenta, beta-sheets represent as yellow in colour, turns are pale blue and all the other residues are white in colour.



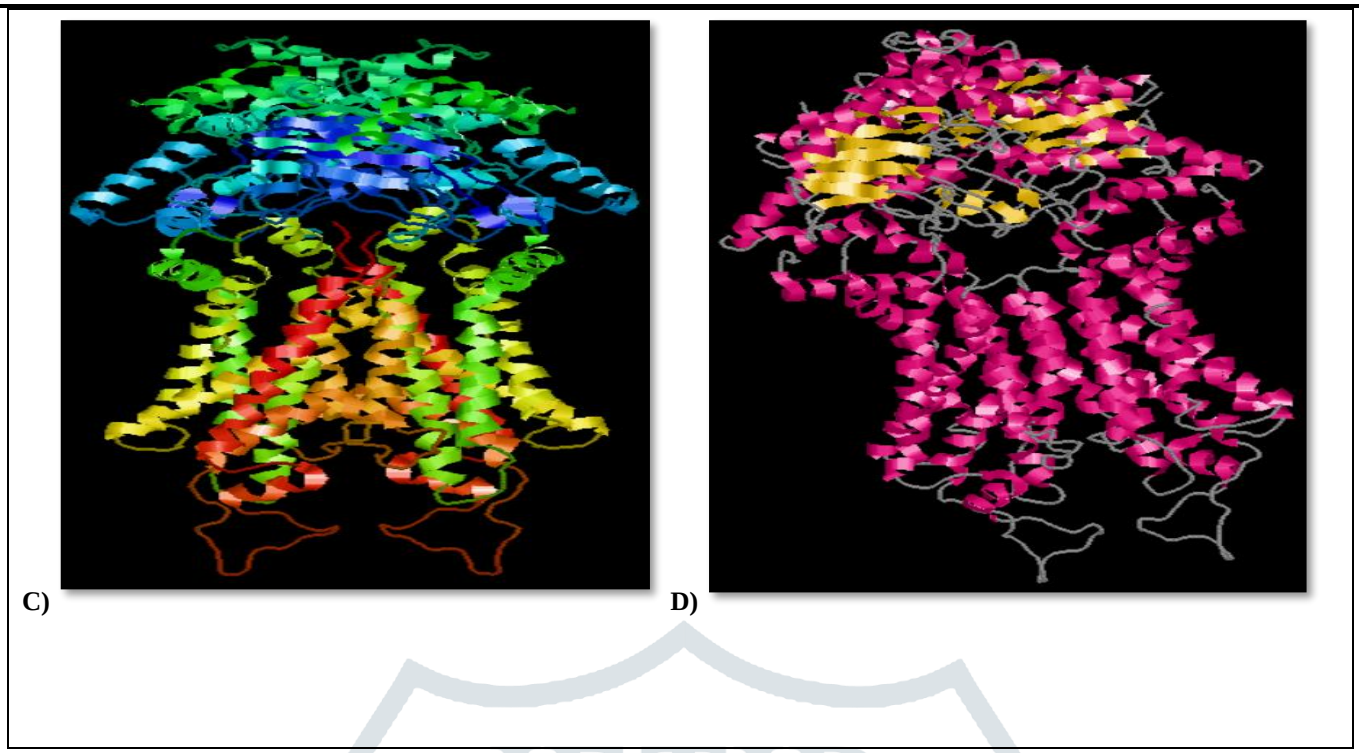


Figure 1: Structure visualization using RasMol

BLAST is used to search databases for distinguished local alignment to query sequence. Figure 2 shows the summary of databases that are similar to ABCG2 protein. Higher scoring hits are present at the top of the list, maximum score represent highest alignment score between the protein and database sequences. Data representation in BLAST tool is depend upon the E-value, if e-value is zero it gives good alignment and e-value more than zero shows gap in the alignment. The aligned databases has 0.0 E-value. Percentage identity shows the similarity between protein and aligned sequences.

Descriptions | Graphic Summary | Alignments | Taxonomy

Sequences producing significant alignments Download Select columns Show 100

select all *100 sequences selected* GenPept Graphics Distance tree of results Multiple alignment MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Cryo-EM structure of the ABCG2 E211Q mutant bound to ATP and Magnesium [Homo sapiens]	Homo sapiens	1319	1319	100%	0.0	100.00%	655	6HBU_A
broad substrate specificity ATP-binding cassette transporter ABCG2 isoform 1 [Homo sapiens]	Homo sapiens	1318	1318	100%	0.0	99.85%	655	NP_001335914.1
Cryo-EM structure of the ABCG2 E211Q mutant bound to estrone 3-sulfate and 5D3-Fab [Homo sapiens]	Homo sapiens	1318	1318	99%	0.0	100.00%	664	6HCO_A
Chain A, Broad substrate specificity ATP-binding cassette transporter ABCG2 [Homo sapiens]	Homo sapiens	1316	1316	99%	0.0	99.85%	665	7OJ8_A
ATP-binding cassette protein ABCG2 [Homo sapiens]	Homo sapiens	1316	1316	100%	0.0	99.69%	655	AAO14617.1
Structure of an ABC transporter: complete structure [Homo sapiens]	Homo sapiens	1316	1316	99%	0.0	99.85%	664	5NJ3_A
breast cancer resistance protein [Homo sapiens]	Homo sapiens	1315	1315	100%	0.0	99.69%	655	AAC97367.1
ABC transporter ABCG2 [Homo sapiens]	Homo sapiens	1315	1315	100%	0.0	99.69%	655	AAG52982.1
mutant ATP-binding cassette sub-family G (WHITE) member 2 [Homo sapiens]	Homo sapiens	1314	1314	100%	0.0	99.69%	655	AAQ92941.1

Figure 2: Summary of similar database sequence to ABCG2 protein

Figure 3 shows the graphical overview of multiple protein sequence alignment. Sequences that matches with canonical sequence are gray in colour and the mismatch sequence are coloured in red. Blue colour represents insertion in the sequence.

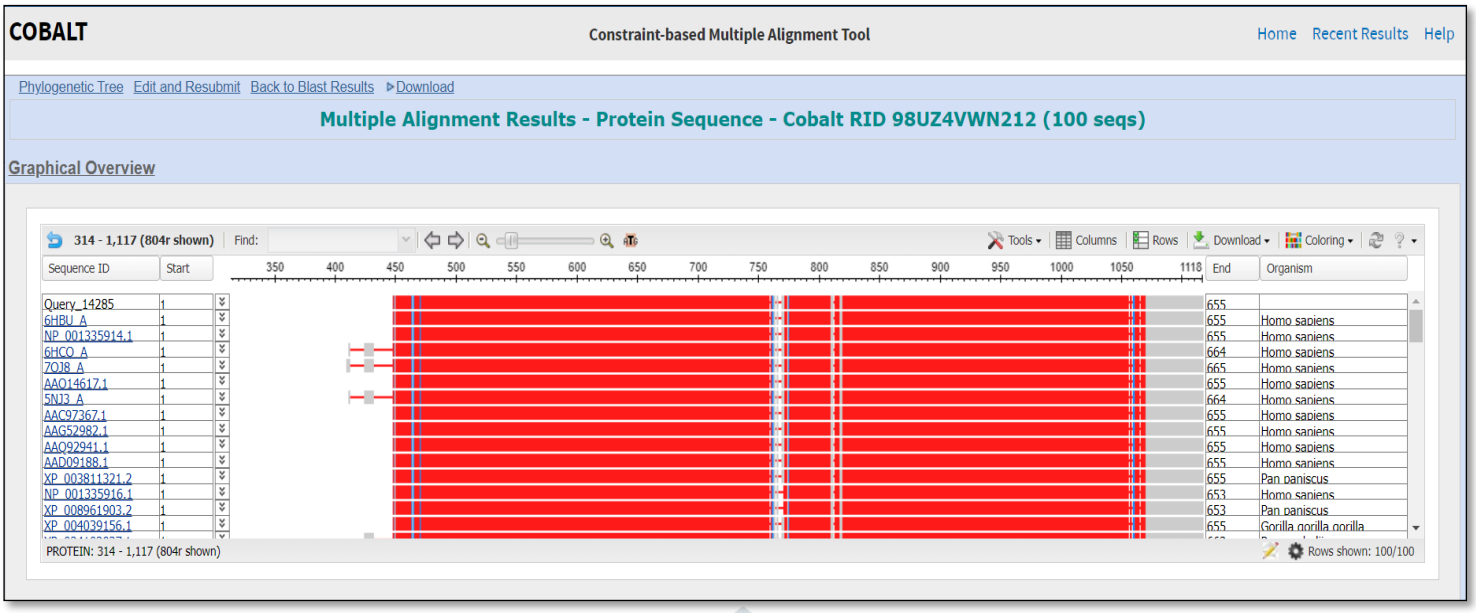


Figure 3: Multiple alignment representation of cobalt

Block substitution matrix (BLOSUM) is a scoring matrix which compares protein sequences to determine the quality of alignment. Sequences that shows good match with canonical sequence are blue in colour while the sequence that show worst match are green in colour.

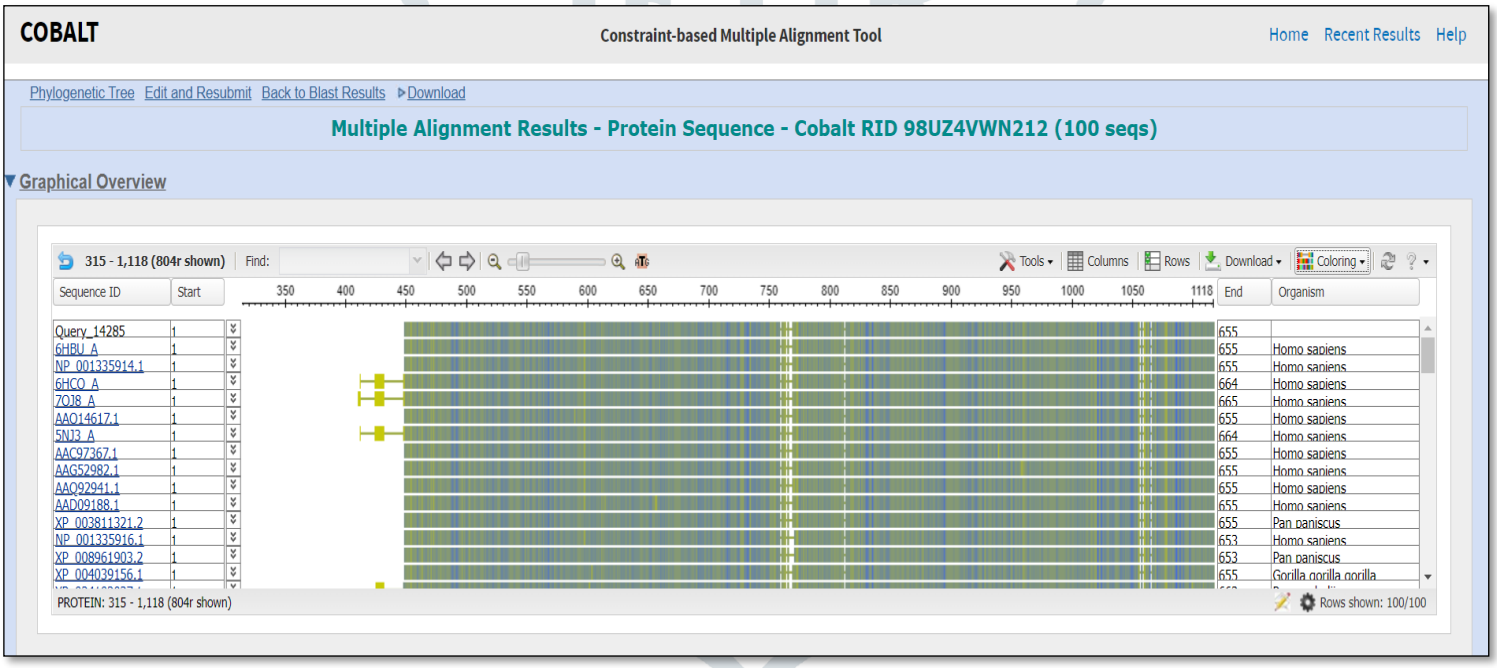


Figure 4: Graphical overview of BLOSUM62

Table 1: Bioactivity score of the synthesized compounds against regular human receptors.

Drugs	GPCR ligand	Ion channel modulator	Kinase inhibitor	Nuclear receptor ligand	Protease inhibitor	Enzyme inhibitor
Mifamurtide	-3.84	-3.92	-3.92	-3.93	-3.75	-3.84
Dichloroacetic acid	-3.72	-3.70	-3.73	-3.65	-3.52	-3.51
Carmustine	-0.93	-0.59	-1.05	-1.44	-1.09	-0.35
Cirazoline	0.42	0.52	-0.44	-0.71	-0.22	0.16
Barbexaclone	-0.30	-0.01	-0.46	-0.50	-0.12	-0.13
Tiazofurin	0.38	-0.05	-0.06	-0.90	-0.05	0.81
D-Asparagine	-1.60	-1.11	-2.27	-2.35	-1.13	-1.06

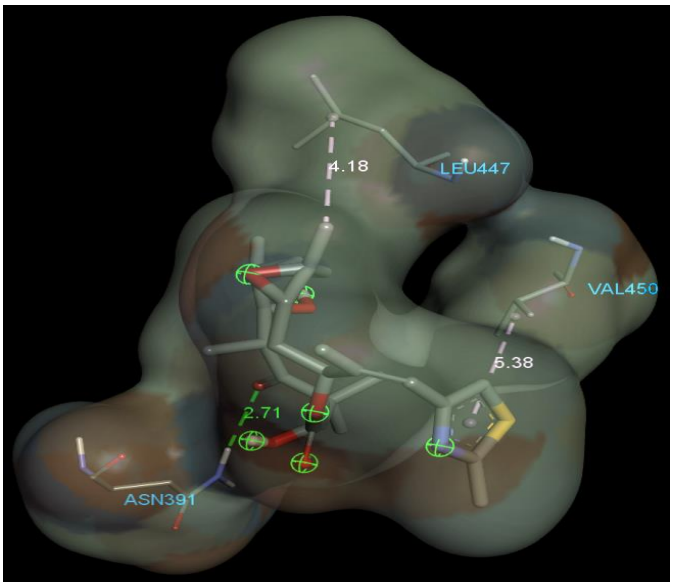
Midomafetamine	-0.20	-0.09	-0.71	-1.05	-0.49	-0.21
Chlorphenesin	-0.44	-0.13	-0.66	-0.55	-0.80	-0.19
Patupilone	0.11	-0.17	-0.28	0.50	0.26	0.65
Niaprazine	0.26	0.06	0.15	-0.23	0.11	-0.03
2-deoxyglucose	-0.62	0.07	-1.00	-0.76	-0.17	0.31
Vapreotide	-3.72	-3.85	-3.83	-3.87	-3.65	-3.78
Zorubicin	-0.27	-1.23	-0.77	-0.67	0.12	-0.24
9,10-Deepithio-9,10-Didehydroacanthifolicin	-1.51	-2.56	-2.50	-1.94	-1.03	-1.51
SR-9009	0.02	-0.18	-0.25	-0.24	0.07	-0.12

To evaluate druglikeness, we did ADME (Absorption, Distribution, Metabolism and Excretion) analysis to examine the effect of drugs in regards to safety and potency of a living organism. Out of 16 drugs, 4 drugs does not follow Lipinski's rule of 5.

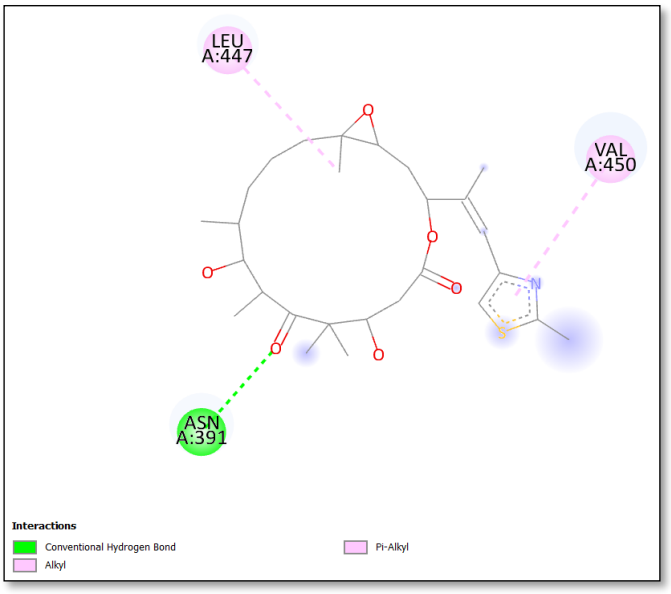
Table 2: Docking result showing binding energy of drugs.

Drugs	Binding Energy
Dichloroacetic acid	-4.4
Carmustine	-4.4
Cirazoline	-5.8
Barbexaclone	-5
Tiazofurin	-6.1
D-Asparagine	-4.1
Midomafetamine	-5.1
Chlorphenesin	-5.6
Patupilone	-7.6
Niaprazine	-7.4
2-deoxyglucose	-4
SR-9009	-6.4

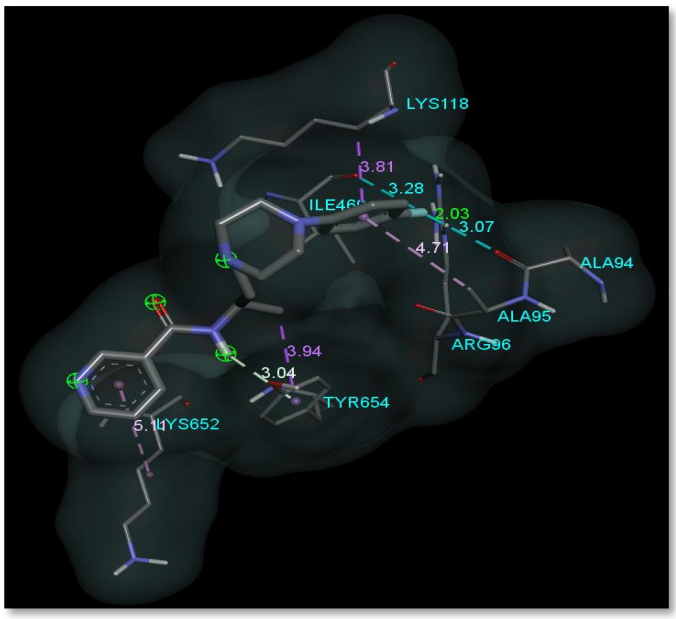
To identify the binding affinities of drugs we did molecular docking between ABCG2 protein and 12 drugs listed in Table 2. Out of 12 drugs, two of them shows good binding affinity with protein i.e., Patupilone has binding energy -7.6 and Niaprazine with -7.4 binding energy. Further, we did 3D structure visualization of ABCG2 protein with Patupilone and Niaprazine by converting them into pdbqt format. In Figure 5, docking image of 6HBU-Patupilone shows interaction of amino acid LEU447, VAL450, ASN391 with protein. These amino acid display different hydrogen-bonds between protein and ligand, to be precise Asparagine shows conventional H-bond with bond length 2.71Å whereas, Leucine and Valine shows alkyl bond with bond length 4.18 and 5.38 respectively. 6HBU-Niaprazine docking image shows interaction of amino acid ALA94, ALA95, ARG96, ILE460, LYS118, LYS652, TYR654 with protein. These amino acid Arginine shows conventional H-bond with 2.03Å bond length, Alanine and Isoleucine shows halogen bond with bond length 3.07Å and 3.28Å respectively, Tyrosine interact with protein through Pi-donor H-bond having 3.04Å bond length, with bond length 4.71Å and 5.11Å Alanine and Lysine respectively, shows Pi-alkyl bond.



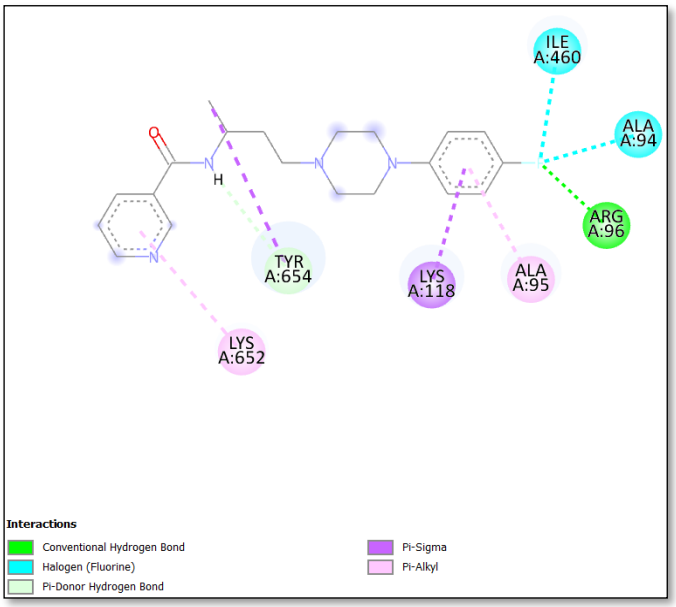
3D Structure of 6HBU-Patupilone



2D Structure of Patupilone



3D Structure of 6HBU-Niaprazine



2D Structure of Niaprazine

Figure 5: 3D structure visualization of protein and ligand

CONCLUSION

From our study we concluded that Patupilone and Niaprazine shows good binding affinity and can be used for further evaluation in the treatment of brain cancer. ABCG2 protein can be modified to study different novel therapy for the treatment of Glioblastoma Multiforme. In COBALT the colour reflects shows the average match over all the other residues in the column. Due to complex pathway GBMs are still incurable, continue research will eventually give the better treatment option for disease. Better understanding of molecular pathway in research and development of immunotherapeutic approach can increase the survival chance of patients.

REFERENCE

1. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, Znaor A, Bray F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer. 2019 Apr 15;144(8):1941-1953. doi: 10.1002/ijc.31937. Epub 2018 Dec 6. PMID: 30350310.
2. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, Bray F. Cancer statistics for the year 2020: An overview. Int J Cancer. 2021 Apr 5. doi: 10.1002/ijc.33588. Epub ahead of print. PMID: 33818764.
3. Fisher R, Pusztai L, Swanton C. Cancer heterogeneity: implications for targeted therapeutics. Br J Cancer. 2013;108:479e485.
4. Hanahan, D.;Weinberg, R.A. Hallmarks of cancer: The next generation. Cell **2011**, 144, 646–674.
5. Butowski NA. Epidemiology and diagnosis of brain tumors. *Continuum*. (2015) 21:301–13. doi: 10.1212/01.CON.0000464171.50638.fa
6. Furnari FB, Fenton T, Bachoo RM, Mukasa A, Stommel JM, Stegh A, et al. Malignant astrocytic glioma: genetics, biology, and paths to treatment. Genes Dev. 2007;21:2683–710, doi: 10.1101/gad.1596707.
7. Loius DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, et al. The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol. 2007;114:97-109, doi: 10.1007/s00401-007- 0243-4.

8. Wilson TA, Karajannis MA, Harter DH. Glioblastoma multiforme: State of the art and future therapeutics. Surgical Neurology International. 2014; 5:64–62. DOI: 10.4103/2152-7806.132138 [PubMed: 24991467]
9. Ohgaki H, Kleihues P. The Definition of Primary and Secondary Glioblastoma. *Clin Cancer Res* (2013) 19(4):764–72. doi: 10.1158/1078-0432.CCR-12-3002
10. Parsons DW, Jones S, Zhang X, Lin JC, Leary RJ, Angenendt P, Kinzler KW. An integrated genomic analysis of human glioblastoma multiforme. Science. 2008; 321:1807–1812. DOI: 10.1126/science.1164382 [PubMed: 18772396]
11. Chen J, McKay RM, Parada LF. Malignant glioma: Lesions from genomics, mouse models, and stem cells. Cell. 2012; 149:36–47. DOI: 10.1016/j.cell.2012.03.009 [PubMed: 22464322]
12. Aldape K, Zadeh G, Mansouri S, et al. (2015). Glioblastoma: pathology, molecular mechanisms and markers. Acta Neuropathol, 129, 829-48.
13. Agnihotri S, Burrell KE, Wolf A, et al (2013). Glioblastoma, a brief review of history, molecular genetics, animal models and novel therapeutic strategies. AITE, 61, 25-41.
14. Cloughesy TF, Cavenee WK, and Mischel PS (2014). Glioblastoma: from molecular pathology to targeted treatment. Annu Rev Pathol, 9, 1-25.
15. Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, Alexe G, Lawrence M, O'Kelly M, Tamayo P, Weir BA, Gabriel S, Winckler W, Gupta S, Jakkula L, Feiler HS, Hodgson JG, James CD, Sarkaria JN, Brennan C, Kahn A, Spellman PT, Wilson RK, Speed TP, Gray JW, Meyerson M, Getz G, Perou CM and Hayes DN: An integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR and NF1. Cancer Cell 17: 98-110, 2010.
16. Frezza C, Tennant DA and Gottlieb E: IDH1 Mutations in gliomas: When an enzyme loses its grip. Cancer Cell 1917: 7-9, 2010.
17. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W, Kos I, Batinic-Haberle I, Jones S, Riggins GJ, Friedman H, Friedman A, Reardon D, Herndon J, Kinzler KW, Velculescu VE, Vogelstein B and Bigner DD: IDH1 and IDH2 mutations in gliomas. N Engl J Med 360: 765-773, 2009.
18. Manolaridis I, Jackson SM, Taylor NMI, Kowal J, Stahlberg H, Locher KP. Cryo-EM structures of a human ABCG2 mutant trapped in ATP-bound and substrate-bound states. Nature. 2018 Nov;563(7731):426-430. doi: 10.1038/s41586-018-0680-3. Epub 2018 Nov, 7. PMID:30405239 doi:http://dx.doi.org/10.1038/s41586-018-0680-3.
19. Zhang W, Mojsilovic-Petrovic J, Andrade MF, Zhang H, Ball M, Stanimirovic DB. The expression and functional characterization of ABCG2 in brain endothelial cells and vessels. FASEB J. 2003 Nov;17(14):2085-7. Epub 2003 Sep 4. PMID:12958161 doi:http://dx.doi.org/10.1096/fj.02-1131fje.