



Received on 16 March 2019; received in revised form, 12 June 2019; accepted, 16 November 2019; published 01 December 2019

A NEW TARGET FOR DIABETES: HOMOLGY MODELING AND IDENTIFICATION OF POTENTIAL LEADS THROUGH VIRTUAL SCREENING

Kotla Naveen Kumar ^{*1}, N. V. Sreedevi ² and P. Spandana ³

Department of Medicinal Chemistry ¹, National Institute of Pharmaceutical Education and Research (NIPER), Medicinal chemistry, Hyderabad - 500037, Telangana, India.

Drug Safety Associate ², Accenture, Bengaluru - 560029, Karnataka, India.

G. Pulla Reddy College of Pharmacy ³, Hyderabad - 500028, Telangana, India.

Keywords:

Homology modeling,
Virtual screening, Monocarboxylate
transporter 11 (MCT11), SLC16A11,
Autodock Vina, PyRx

Correspondence to Author:

Kotla Naveen Kumar

Department of Medicinal Chemistry,
National Institute of Pharmaceutical
Education and Research (NIPER),
Medicinal Chemistry, Hyderabad -
500037, Telangana, India.

E-mail: naveenkotla.medchem@gmail.com

ABSTRACT: Monocarboxylate transporter 11 (MCT11) is a transporter protein encoded by the SLC16A11 gene. MCT11 protein is found to play an important role in many biological activities in different species. MCT11 helps in hepatic lipid metabolism by catalyzing the transport of pyruvate molecules across the membrane. It is also found to be involved in regulation of Diabetes. Therefore, MCT11 protein is taken into consideration for this study. Moreover, a rapid analysis through different protein databases revealed that there is no protein crystal structure available. Therefore, Homology modeling is adopted to generate the model using SWISS-MODEL and validated using PROCHECK. More specifically, this is the first Homology model developed for the MCT11 protein. The sequence analysis studies such as physicochemical properties such as molecular weight (47790.99 Da), Theoretical isoelectric point (8.67), Total number of positively (Asp+Glu) and negatively (Arg+Lys) charged residues (+R/-R) were 19 and 22 respectively. The grand average of hydropathicity (GRAVY) indicating a value of 0.687 was computed. The secondary structure prediction and conserved domain sequence were computed. Finally, the modeled protein subjected to virtual screening to identify the possible leads using PyRX software, which uses the Autodock Vina module for virtual screening. About 1943 compounds were screened virtually and 10 compounds were found to be extremely effective having binding energies range from -11.0 Kcal/mol to -11.60 Kcal/mol.

INTRODUCTION: Diabetes mellitus is a non-communicable, pandemic metabolic disorder characterized by intolerance to glucose and alterations in lipid and protein metabolism ¹. The major problem arises when the pancreas is inefficient to produce insulin or body unable to use the produced insulin, which is also known as “Insulin resistance” ².

Type 2 diabetes occurs when these two conditions amplified and lead to increased levels of glucose in the blood, which is a hallmark of Type 2 diabetes ³. Chronic diabetic patients who are untreated suffer from complications such as retinopathy, nephropathy, and peripheral neuropathy. Acquiring other types of risks such as cardiovascular disease and cancer are also higher in diabetic patients ⁴.

According to the recent 2017 reports of the International Diabetes Federation, over 425 million diabetic patients are present worldwide, and it is estimated that about 325 million people are at risk of developing diabetes ⁵. The major symptoms of Type 2 diabetes include excessive thirst and dry

	<p style="text-align: center;">DOI: 10.13040/IJPSR.0975-8232.10(12).5397-03</p>
	<p style="text-align: center;">This article can be accessed online on www.ijpsr.com</p>
<p>DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.10(12).5397-03</p>	

mouth, recurrent fungal infections in the skin, frequent and abundant urination, slow healing wounds, blurred vision, lack of energy and extreme tiredness⁶. The radical risk factors include primarily the family history, overweight, unhealthy diet, physical inactivity, increasing age, high blood pressure, impaired glucose tolerance (Impaired glucose is higher than normal blood glucose but below the threshold for diagnosis), and poor nutrition during pregnancy⁷.

Many classes of medications are available to target different cellular mechanisms to treat this condition. The two main treatment approaches include the stimulation of insulin release⁸ and insulin-stimulated glucose uptake⁹. However, Therapeutic response varies from person to person, and it requires multiple medications to maintain blood glucose levels. On the other hand, all these drugs have moderate to serious adverse effects¹⁰. Therefore, there is a need to identify the new therapeutic agents which target specifically to a selective target protein that plays a role in disease strategy. It has been uncovered that diabetes does not only occur due to the environmental factors but also due to the heredity¹¹. Therefore, a deep understanding of the genetic contributions involved in disease incidence might help to focus it more specifically and effectively.

A study conducted by Rusu *et al.*, suggested that chromosome 17p13 is associated with diabetes risk factors in Mexican adults, a population with a high prevalence of Type 2 diabetes¹². "They showed that 12 non-coding variants span the SLC16A11 gene locus and that there are four missense variants in the SLC16A11 gene itself". The gene SLC16A11 is a member of the SLC16 gene family and estimated that it is linked to an increase in 30% of risk rate in this population¹². Monocarboxylate transporter 11 (MCT11) is a proton-coupled transporter protein that is encoded by SLC16A11, which plays an important role in transporting the monocarboxylates between the cell membrane. A knockdown study conducted on SLC16A11 revealed that a significant increase in intracellular acylcarnitines, diacylglycerols, and triacylglycerols and extracellular triacylglycerols. These above changes strongly suggest that fatty acid and lipid metabolism is affected by Monocarboxylate transporter 11 (MCT11), proteins. Therefore,

targeting the MCT11 might lead to personalized treatment of diabetes¹³.

Herein, we identified a Monocarboxylate transporter 11 (MCT11) protein as possible therapeutic target to treat type-2 diabetes. However, there is no three-dimensional protein crystal structure deposited in the Protein data bank. The three-dimensional structure of a protein is crucial for understanding the molecular function and its interactions with the biological process¹⁴. Often, proteins play a major role in cell signaling pathways and act as therapeutic targets. Although a great trend is directed towards the structure resolving and deposition in PDB, Some new identified targets structure is still a question due to the risks involved in the crystallization techniques¹⁵. In other hand, due to the exploration of human genome project new sequences tangled up every day. Hence, a bioinformatics approach is needed to model the proteins whose structures are unknown.

Thus, we adopted a Homology modeling method to model the unknown protein structure based on sequence alignment and model building using known identical protein template sequence¹⁶. Homology modeling plays a crucial role in predicting protein structure¹⁷. For the model development, we utilized an online server SWISS Model and generated the model. In homologous complexes, it is been observed that interacting residues are often conserved. Hence, SWISS-MODEL is used in this study to build the homologous model of the protein. This is an automated homology modeling server with increased accuracy for generated models¹⁸.

MATERIALS AND METHODS:

MCT11 Protein Sequence Analysis: The MCT11 protein sequence originated from Homo sapiens was retrieved from UniProtKB database bearing accession no Q8NCK7. The sequence was saved in FASTA format and further used it for analysis of protein¹⁹.

Physicochemical Properties: Physicochemical properties such as Number of amino acids, molecular weight (Mol. wt.), amino acid composition, theoretical isoelectric point (pI), total number of positive (Arg+Lys) and negative (Asp+Glu) residues (+R/-R), Atomic composition,

extinction coefficient (EC), estimated half-life, instability index (AI), and grand average of hydropathicity (GRAVY) of MCT11 transporter protein was analyzed by searching protein sequence in the Expasy's ProtParam server (<http://web.expasy.org/protparam/>)²⁰.

Secondary Structure Prediction: Several online Bioinformatics tools were adopted to analyze the conserved domains of MCT11 transporter protein. The NCBI conserved domains database (CDD) is used to retrieve the protein annotation. "CDD is a protein annotation resource that consists of a collection of well-annotated multiple sequence alignment models for ancient domains and full-length proteins²¹. These are available as position-specific score matrices (PSSMs) for fast identification of conserved domains in protein sequences *via* RPS-BLAST". PSIPRED is a secondary structure prediction server, which is simple and works based on neural network analysis developed by the UCL department of computer science²². InterProScan, a tool developed by The European Bioinformatics Institute (EBI). "This provides a functional analysis of proteins sequence by classifying into families and predicting the presence of domains and important sites"²³.

Alignment of Sequence, Model Building, and Model Validation: MCT11 protein template sequence is searched against the known protein structures sequence and model building is performed using the SWISS-MODEL an online server. In this step, the server compares the unknown sequence with known protein structures deposited in the Protein Data Bank (PDB)²⁴. The generated model was validated using different tools to ensure the accuracy of model. Ramachandran plot was generated using PROCHEK tool²⁵.

Virtual Screening: Protein-ligand docking is one of the best computational methods to predict the ligand orientation and conformation in the binding site. AutoDock Vina is an open-source software currently used by many scientists for molecular docking and virtual screening. AutoDock Vina uses empirical scoring function for prediction of protein-ligand affinity. Virtual screening works on molecular docking algorithm to predict the binding energies and interactions. To perform virtual screening we incorporated PyRx software in this

work²⁶ which works on Autodock Vina module. PyRx is a flexible, versatile and user-friendly application to perform virtual screening using AutoDcok 4 and Autodock Vina²⁷.

RESULTS AND DISCUSSION:

Physicochemical Properties: Expasy's ProtParam server was used to compute the physicochemical properties of the MCT11 protein. A total of 471 amino acid residues were predicted. The molecular weight of 47790.99 Da and the isoelectric point (pI) of protein of 8.67 indicating the slightly basic nature of the protein. The pI values function in protein purification by isoelectric point focusing on a polyacrylamide gel. A total number of positively (Asp+Glu) and negatively (Arg+Lys) charged residues (+R/-R) were 19 and 22 respectively. The Instability index (II) values of 43.33 indicating it an unstable protein and an aliphatic index of 112.14. The grand average of hydropathicity (GRAVY) indicating a value of 0.687. The physicochemical properties were listed in **Table 1**.

TABLE 1: THE PHYSICOCHEMICAL PROPERTIES OF MCT11 ANALYZED USING EXPASY'S PROTPARAM TOOL

S. no.	Physicochemical properties	MCT11
1	No of amino acids	471
2	Molecular weight	47790.99
3	Theoretical PI	8.67
4	Total number of negatively charged residues (Asp + Glu)	19
5	Total number of positively charged residues (Arg + Lys)	22
6	Total number of atoms	6837
7	Formula	C ₂₂₁₃ H ₃₄₅₇ N ₅₇₁ O ₅₈₄ S ₁₂
8	Ext. coefficient	68660
9	Abs 0.1% (=1 g/l): assuming all pairs of Cys residues form Cystines	1.437
10	The instability index (II) is computed to be	43.33
11	Aliphatic index	112.14
12	Grand average of hydropathicity (GRAVY)	0.687

Secondary Structure Prediction: To predict the secondary structure, protein functionality annotation and conserved domains of MCT11 transporter protein sequence, a wide range of tools were incorporated. NCBI conserved domain analysis suggests that the MCT 11 transporter

protein sequence has conserved domain sequence with the cd06174, which was a major facilitator family of large and diverse secondary transporter

family. The other conserved domains were found in the oxalate/formate antiporter family.

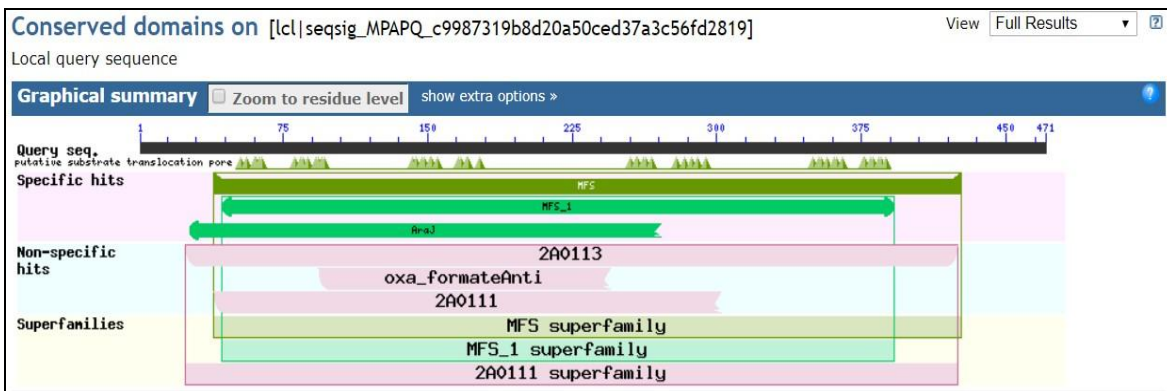


FIG. 1: REPRESENTS THE CONSERVED DOMAIN SEQUENCE OF MCT11 TRANSPORTER PROTEIN USING NCBI-CDD

A 471 chain length of MCT11 protein sequence was submitted to the InterProScan Beta tool, which analyses sequence by classifying them into families and anticipates the important sites and finally provides the functional analysis of protein

sequence. PSIPRED was secondary structure prediction server, which was simple and works based on neural network analysis developed by the UCL department of computer science.



FIG. 2: A: INTERPROSCAN ANALYSIS OF MCT11 TRANSPORTER PROTEIN. B: PSIPRED SECONDARY STRUCTURE PREDICTION OF MCT11

Alignment of Sequence and Model Building:

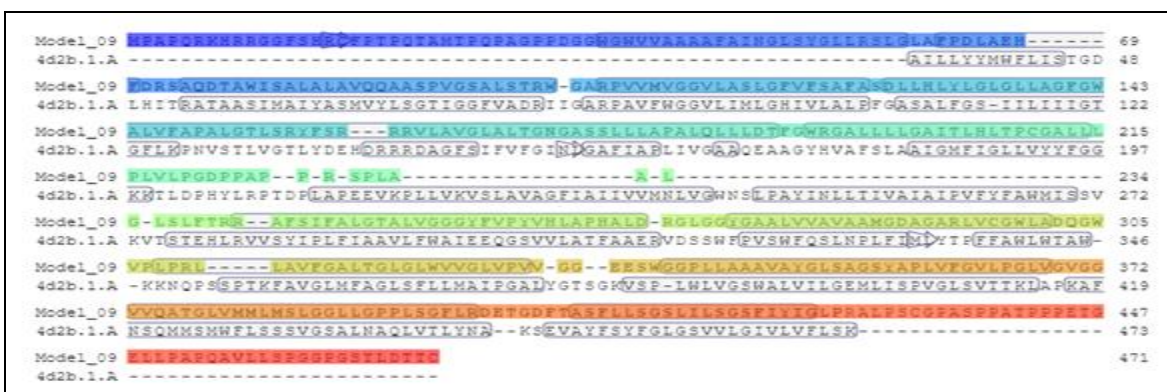


FIG. 3: SEQUENCE ALIGNMENT OF MCT11 TRANSPORTER SEQUENCE WITH THE 4d2b PDB PROTEIN TEMPLATE

Model Validation: Validation of the model is crucial step to ensure the generated model has reliability and properly aligned to the template of structure. Therefore, we validated our model using different methods and tools. The alignment of 3D structures of generated model was verified by aligning with the actual crystal structure.

The calculated Root mean square deviation (RMSD) of the alignment of both models is 0. Indicating that generated model is most similar to template.

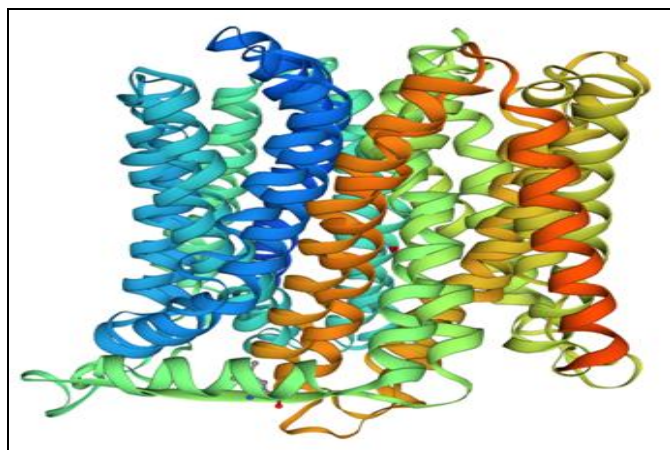


FIG. 4: 3D CARTOON STRUCTURE ALIGNMENT OF MCT11 GENERATED MODEL AND CRYSTAL STRUCTURE OF 4d2b PDB PROTEIN

Ramachandran plot was generated using the PROCHECK tool, which was “one method to analyze the energetically allowed regions for backbone dihedral angles against amino acid residues in protein”²⁸. This plot showed a distribution of 367 residues of MCT 11 transporter protein with 92.0% of residues in most favored region having 264 residues in core region, 7.0% in allowed region with 20 residues, only 1.0% in generously allowed region with 3 residues, and 0.0% in disallowed region with no residues.

Identification of Active Site: The active site was identified by Tripos Sybyl x 2.0 through site ID find pockets methods. The active amino acid residues in the pocket are Ala 74, Gln 75, Thr 77, Ala 78, Pro 261, Tyr 262, Val 263, Leu 265, Ala 266, Pro 267, Asp 271, Val 282, Val 283, Val 285, Ala 286, Leu 342, Ala 346, Tyr 349, Ser 395, Leu 398, Arg 399, Asp 400, Pro 64, Asp 65, Leu 61, and Glu 68. These active pocket residues are used as binding residues in the virtual screening method to screen the molecules.

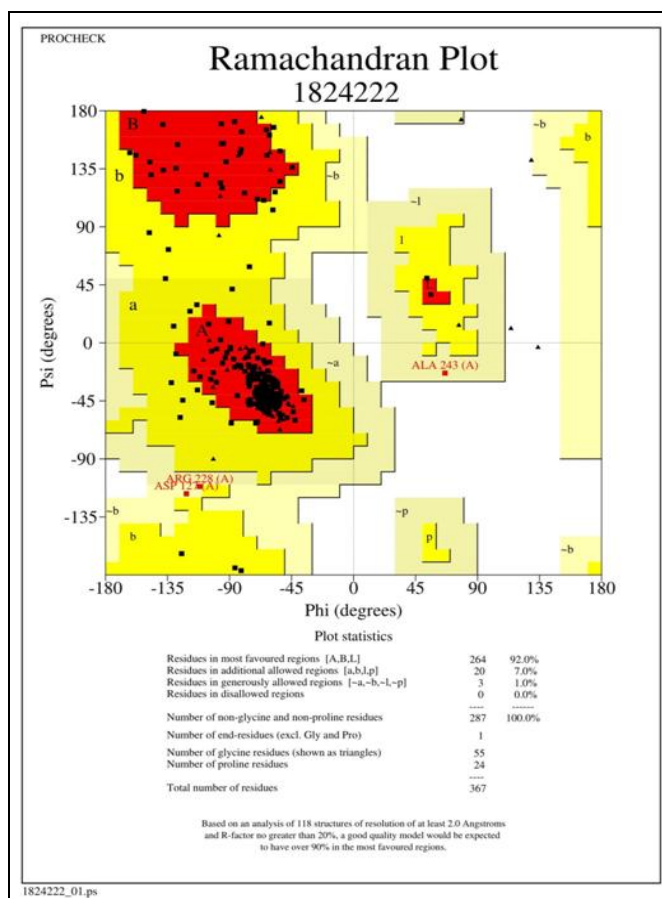


FIG. 5: RAMACHANDRAN PLOT OF THE MODELLED MCT11 TRANSPORTER PROTEIN EXHIBITED 92.0% AMINO ACID RESIDUES IN MOST FAVORED REGION

Virtual Screening: In this study, we used anti-hyperglycemic agents downloaded from NCBI-Pubchem Bioassay database. A total of 2103 compounds were downloaded from a project entitled in PubChem Bioassay database “Broad Institute HTS compounds Modulating PGC-1 α Acetylation and Oxidative Metabolic function Inhibitor Probe Project” as SDF format and subjected for energy minimization. Three-dimensional conformer generation and duplicate molecules deletion provided a total of 1943 compounds. The homology modeled protein was used for virtual screening; the protein and all the ligands were prepared using PyRx and Open Babel interfaces respectively. The grid box is setup by considering all the active residues and making sure all the residues are occupied by the grid box. Finally, virtual screening job was performed in PyRx using AutoDock Vina wizard interface. Virtual screening of 1943 compounds generated good binding energies, out of which 75 compounds have binding energy in range of -10.0 Kcal/mol to -

11.60 Kcal/mol. The compounds (10) having highest binding energies were shown in **Table 2** and all the 10 best compounds were superimposed in binding pocket with their best binding pose which was represented in **Fig. 6**.

TABLE 2: BINDING ENERGIES OF 10 BEST COMPOUNDS

S. no.	PubChem substance ID	Binding Energy (Kcal/mol)
1	85791352	-11.0
2	14744900	-11.0
3	144190998	-11.1
4	131463813	-11.1
5	144191371	-11.2
6	131443643	-11.2
7	131442360	-11.2
8	89855186	-11.2
9	144191247	-11.5
10	131443642	-11.6

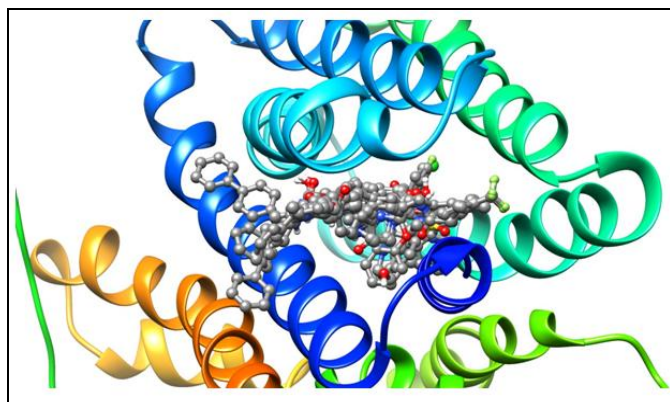


FIG. 6: SUPERIMPOSITION OF ALL THE 10 BEST ACTIVE MOLECULES IN THE BINDING POCKET

CONCLUSION: In this present study, Monocarboxylate transporter 11 (MCT11) protein sequence of Homo sapiens species was selected to identify the possible lead molecules. There was no crystal structure available for MCT11 protein. Therefore, homology modeling was adopted and model was generated using 4d2b known protein crystal structure sequence. Physicochemical and functional characterization of the protein was profoundly investigated. The three-dimensional models of protein were generated using SWISS-MODEL and validated by the accuracy of Ramachandran plot analysis using PROCHECK tool. The results suggested that proposed model is reliable and valid. This study not only provided information on the physiochemical characteristics, structural properties and molecular functions of Monocarboxylate transporter 11 (MCT11) protein, but also the best possible leads from virtual

screening. These lead molecules can be further tested in protein-specific bioassay and might help to treat diabetes as personalized medicine.

ACKNOWLEDGEMENT: Authors would like to thank the National Institute of Pharmaceutical Education and Research (NIPER), Medicinal chemistry, Hyderabad for support during the project.

CONFLICTS OF INTEREST: Authors declare no conflicts of interest

REFERENCES:

1. Jerrold JH, Blumberg B, Cave M and Machtinger R: Metabolism disrupting chemicals and metabolic disorders. *Reproductive Toxicology* 2017; 68: 3-33.
2. Wasserfall C, NickHS and Campbell-Thompson M: Persistence of Pancreatic Insulin mRNA Expression and Proinsulin Protein in Type 1 Diabetes Pancreata. *Cell Metabolism* 2017; 26: 568-75.
3. Hu F, Qiu X and Shizhong: Pancreatic islet dysfunction in type 2 diabetes mellitus. *Archives of Physiology and Biochemistry* 2018; 4: 1-7.
4. de Kort S, Masclee Ad AM and Sanduleanu S: Higher risk of colorectal cancer in patients with newly diagnosed diabetes mellitus before the age of colorectal cancer screening initiation. *Scientific Reports* 2017; 7: 46527.
5. Alberti KGMM, Zimmet P and Shaw J: International Diabetes Federation: a consensus on Type 2 diabetes prevention. *Diabetic Medicine* 2007; 24:451-63.
6. Liu X, Yuqian L and Linlin L: Prevalence, awareness, treatment, control of type 2 diabetes mellitus and risk factors in Chinese rural population: the Rural Diab study. *Scientific Reports* 2016; 6: 31426.
7. Vanesa B, Lazaros B, Ioanna T and Evangelos E: Risk factors for type 2 diabetes mellitus: An exposure-wide umbrella review of meta-analyses. *PLOS ONE* 2018; 13: e0194127.
8. Halperin F, Lopez X, Manning R, Kahn C and Kulkarni RN: Insulin augmentation of glucose-stimulated insulin secretion is impaired in insulin-resistant humans. *Diabetes* 2012; 61: 301-09.
9. Fujimoto B, Lee A and Wong AM: The exocyst complex in insulin-stimulated glucose uptake in skeletal muscle cells. *Diabetes* 2018; 67: 241-LB.
10. Arun C, Chitharanjan D and Reddy DVS: Clinical review of antidiabetic drugs: implications for Type 2 diabetes mellitus management. *Frontiers in Endo* 2017; 8: 1-12.
11. Ingelsson E and Mark IM: Human Genetics of Obesity and Type 2 Diabetes Mellitus. *Circulation, Genomic and Precision Medicine* 2018; 11: e002090.
12. Rusu V, Hoch E, Mercader JM, Tenen DE, Gymrek M and Hartigan CR: Type 2 diabetes variants disrupt function of *slc16a11* through two distinct mechanisms. *Cell* 2017; 170: 199-12.
13. Morris ME and Felmler MA: Overview of the proton-coupled MCT (SLC16A) family of transporters: characterization, function and role in the transport of the drug of abuse gamma-hydroxybutyric acid. *The AAPS Journal* 2008; 10: 311-21.
14. da Costa WLO, Araújo CL, Dias LM, Pereira LC, Alves JTC, Araújo FA and Folador EL: Functional annotation of

- hypothetical proteins from the *Exiguobacterium antarcticum* strain B7 reveals proteins involved in adaptation to extreme environments, including high arsenic resistance. PLOS ONE 2018; 13: e0198965.
15. Arkhipova V, Guskov A and Slotboom DJ: Analysis of the quality of crystallographic data and the limitations of structural models. The J of General Physio 2017; 149: 1091.
 16. Dalton JAR and Jackson RM: An evaluation of automated homology modelling methods at low target–template sequence similarity. Bioinformatics 2007; 23: 1901-08.
 17. Vyas VK, Ukawala RD, Ghate M and Chintha C: Homology modeling a fast tool for drug discovery: current perspectives. Ind J of Pharma Sciences 2012; 74: 1-17.
 18. Waterhouse A, Bertoni M, Bienert S: SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Research 2018; 46: W296-W303.
 19. The UniProt C: a worldwide hub of protein knowledge, Nucleic Acids Research 2018; 47: D506-D515.
 20. Gasteiger E, Hoogland C and Gattiker A: Protein identification and analysis tools on the ExPASy server. In: Walker JM, ed. The Proteomics Protocols Handbook. Totowa, NJ: Humana Press 2005: 571-07.
 21. Marchler-Bauer A, Bo Y and Han L: CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res 2017; 45: D200-D203.
 22. Buchan DWA, Minneci F, Nugent TCO, Bryson K and Jones DT: Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Research 2013; 41: W349-W357.
 23. Jones P, Binns D and Chang HY: InterProScan 5: genome-scale protein function classification, Bioinformatics (Oxford, England) 2014; 30: 1236-40.
 24. Waterhouse A, Bertoni M and Bienert S: SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Research 2018; 46: W296-W303.
 25. Laskowski RA, MacArthur MW, Moss DS and Thornton JM: PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography 1993; 26: 283-91.
 26. Dallakyan S and Olson AJ: Small-molecule library screening by docking with PyRx. Chemical Biology 2015; 1263: 243-50.
 27. Trott O, Olson AJ: AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry 2010; 31: 455-61.
 28. Ramachandran GN, Ramakrishnan C and Sasisekharan V: Stereochemistry of polypeptide chain configurations. Journal of Molecular Biology 1963; 7: 95-99.

How to cite this article:

Kumar KN, Sreedevi NV and Spandana P: A new target for diabetes: homology modeling and identification of potential leads through virtual screening. Int J Pharm Sci & Res 2019; 10(12): 5397-03. doi: 10.13040/IJPSR.0975-8232.10(12).5397-03.

All © 2013 are reserved by the International Journal of Pharmaceutical Sciences and Research. This Journal licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.

This article can be downloaded to **Android OS** based mobile. Scan QR Code using Code/Bar Scanner from your mobile. (Scanners are available on Google Play store)