Scientific paper

# **Assessment of Interaction of Human OCT 1-3 Proteins and Metformin Using Silico Analyses**

# Faruk Berat Akçeşme,<sup>1,\*</sup> Nail Beşli,<sup>2</sup> Jorge Peña-García<sup>3</sup> and Horacio Pérez-Sánchez<sup>3</sup>

 $^1$  Department of Biostatistics and Medical Informatics, Faculty of Medicine, University of Health Sciences, Istanbul, Turkey

<sup>2</sup> Department of Medical Biology, Faculty of Medicine, University of Health Sciences, Istanbul, Turkey.

<sup>3</sup> Bioinformatics and High Performance Computing Research Group (BIO-HPC), Computer Engineering Department, Universidad Católica de Murcia (UCAM), Guadalupe, 30107, Spain.

\* Corresponding author: E-mail: farukberat.akcesme@sbu.edu.tr Tel: +090-537- 502 2634

Received: 05-13-2020

#### **Abstract**

Metformin, a drug frequently used by diabetic patients as the first-line treatment worldwide, is positively charged and is transported into the cell through human organic cation transporter (hOCT 1-3) proteins. We aimed to mimic the cellular uptake of metformin by hOCT1-3 with various bioinformatics methods and tools. 3D structure of OCT1-3 proteins was predicted by considering the structures and function of these proteins. We predicted functional regions (active and ligand binding sites) of OCT1-3 and performed comparative bioinformatics analysis. The predicted structure of hOCT1-3 was then analyzed in the Blind Docking server and the results were confirmed with predicted binding site residues and conserved domain regions. We simulated the OCT1-3 and metformin docking and also validated the docking procedure with other substrates of HOCT1-3 proteins. We selected the best poses of metformin docking simulations as per binding energy (-5.27 to -4.60 kcal/mol). Lastly, we validated the static description of protein-ligand (OCT-Metformin) interactions by performing molecular dynamics simulation. Eventually, we obtained stable simulation of OCT-metformin interaction.

Keywords: Organic cation transporters; metformin; protein structure prediction; model Validation; docking

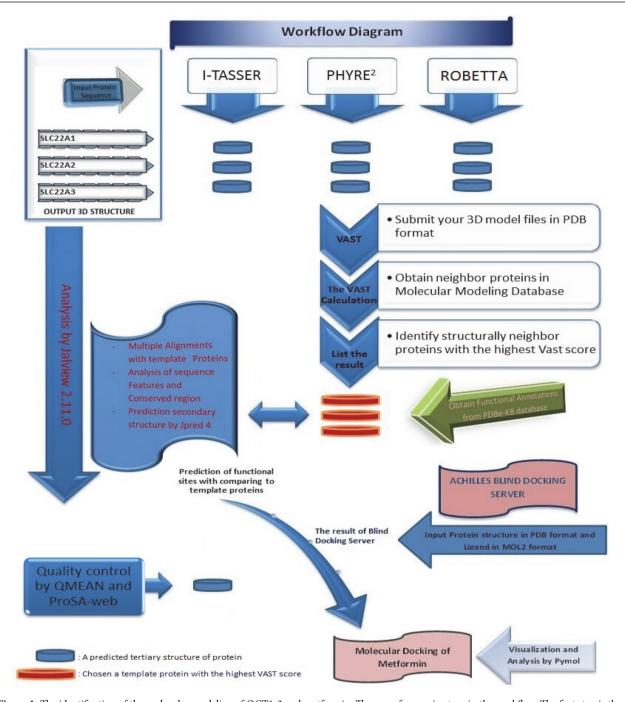
#### 1. Introduction

Membrane proteins are associated with the prominent functions in the cell, approximately responsible for 30% genes in the human genome<sup>1</sup> and currently possessing 50% of pharmaceutical drug discovery.<sup>2</sup> Membrane proteins perform a broad variety of particular roles during cellular events.<sup>3</sup> Due to its structural and physicochemical properties, the plasma membrane has a selective permeability for organic and inorganic substances including cation and anion compounds. Hence, it assists to sustain the unique content both inside and outside of the cell.

One of the protein families that provide translocation of cationic organic and inorganic compounds that localized in the cell membrane is SLC (Solute carrier) family from the MFS superfamily. The SLC family is a 22-membered cell membrane transporter. A subfamily of the SLC family is SLC22A1,2 and 3 (cd17379: MFS\_SLC22A1\_2\_3).

Besides many essential cation molecules for the cell, the SLC transporters are the target of drugs with high pharmacological value. The human body constitutes more than 400 important SLC transporters for a broad range of tasks including drug metabolism as absorption, distribution, and excretion. Hence, there is a growing interest in the effects of the drug on the development and progression of interactions with these transporters.<sup>5</sup>

Metformin, categorized as an anti-diabetic medication, is uptaken by the cell via SLC transporter proteins encoded by SLC22A1-3 (also named OCT1-3) genes.<sup>6</sup> OCT1-3 membrane proteins are expressed at different levels in several tissues, to name a few, the renal expression level of OCT-2 is high, whereas OCT-3 is most commonly expressed in skeletal muscle,<sup>7</sup> and OCT-1 is expressed primarily in hepatocytes.<sup>8</sup> Metformin is hydrophilic (logD –6.13 pH 6.0) and its pKa (physiological pH) is 12.4.<sup>9</sup> Functional elimination of OCT-1 in pri-



**Figure 1.** The identification of the molecular modeling of OCT1-3 and metformin. There are four main steps in the workflow. The first step is the prediction of the 3D structure of OCT1-3 proteins and quality control of the model protein structures. The second one is the identification of the template proteins with the VAST. The third one is the analysis of the sequence data by Jalview 2.11.0 while the fourth is the molecular docking by Achilles Blind Docking Server.

mary mouse hepatocyte culture and OCT-1 has been demonstrated to play an important role in metformin response in vivo mouse model.<sup>10</sup> Following the entry into the cell via HOCT1-3, metformin exhibits its anti-diabetic properties in several ways. It is broadly believed that the blood-glucose-lowering impact of metformin is mediated chiefly through the repression of hepatic glucose production by decreasing gluconeogenesis and blocking gluca-

gon-mediated signaling in the liver.<sup>11,12</sup> Furthermore, some mechanisms have been suggested that metformin can activate AMPK (AMP-activated protein kinase) by the upstream liver kinase B1,<sup>11</sup> enhanced AMP/ATP rate hereby the restraint of mitochondrial respiratory chain complex I.<sup>13</sup> Metformin improves the activity of the insulin receptor and IRS-2 (insulin receptor substrate 2) and boosts glucose uptake through enhanced translocation of

glucose transporters, such as GLUT-1, to the plasma membrane.  $^{14}$ 

In the current study, we aimed to predict the three-dimensional structure of human OCT1-3 using various computational approaches through consideration of the current authenticated/trusted bioinformatics tools. The molecular docking was also performed under the inspiration of the in vitro and in vivo foundings. We also performed Molecular dynamics (MD) simulation of metformin and hOCT1-3 dockings at the atomic level for validation. The computational approaches to OCT1-3 proteins, the particular structure prediction of the proteins and the simulations using MD, have been broadly implemented for investigating their dynamic actions.

Given the pharmacological importance of SLC22A1,2 and 3 proteins in humans, determination of the structure of these proteins, the estimation of their active sites, and the definition of how the transport mechanism works have aroused great interest. In the present study, we have reported and defined the ligand-dependent interactions of hOCT1-3 with metformin and the other ligands utilizing computational approaches and explored the found interactions through comparative analysis, homology modeling, and molecular dynamic studies. This study is the first attempt to demonstrate OCT1-3 interaction with metformin. This interaction has characterized by docking analysis and the results were validated with MD simulations. This is an important study that uses the predicted structure of OCT proteins to stimulate the interaction which stays highly stable throughout the MD analysis.

#### 2. Materials and Methods

### 2. 1. Computational Structural Modeling of OCT1-3

#### 2. 1. 1. Prediction of Secondary and Tertiary Structure of OCT1-3 Proteins

We retrieved OCT1, 2, and 3 (Accession no: AAI26365.1, NP\_003049.2, and NP\_068812.1, respectively) from GenBank in FASTA format, predicted the secondary structure of the proteins using JPred4,1<sup>5</sup> which is the latest version of the JPred online prediction server supplying by the JNet algorithm.

Each of OCT1-3 protein structures was predicted on PHYRE<sup>2</sup>, <sup>16</sup> Robetta<sup>17,18</sup> and I-TASSER<sup>19,21</sup> (protein structure prediction servers). In these prediction tools, homology modeling (or comparative modeling) was used to compare experimentally determined proteins as templates. To control the quality of the model proteins, we performed the local structural quality of transmembrane protein models analysis using (QMEANBrane)<sup>22</sup> and ProSAweb.<sup>23</sup>

We compared each of the obtained models to all PDB proteins in MMDB (Molecular Modeling Database) to find 3D similar structures in VAST (Vector Alignment Search Tool).<sup>24</sup> The VAST analysis was contributed to the following proteins with the highest scores, PDB Id: 4zw9\_A, 4zwc\_A, and 5c65\_A. Each of the OCT1-3 protein structures obtained from the Robetta server was then selected as a model since its neighboring proteins have the highest VAST scores and %Ids, compared to the other prediction tools (See Fig1).

The Computed Atlas of Surface Topography of proteins (CASTp)<sup>25</sup> 3.0 was utilized to predict the surface of the binding pocket of the model proteins for the interaction with their substrates.

#### 2. 1. 2. Sequence Analysis

We performed a pairwise in BLASTP<sup>26</sup> and multiple sequence alignment in Clustal OMEGA<sup>27</sup> for each of the OCT1, 2, and 3 proteins with the selected template proteins (4zw9, 4zwc, and 5c65) obtained from the VAST. Parameters for the alignment with the Clustal OMEGA were set as –GAPEXT:0.1, ENDGAPS: 0.5, GAPDİST: 1, GAPOPEN:10, and MATRIX: BLOSUM62. We analyzed and interpreted the results in the Jalview 2.11.<sup>28</sup>

To analyze the feature of the sequences, functional annotations of template proteins were retrieved from the PDBe-KB database<sup>29</sup>, followed by the comparison of these proteins with the model proteins to identify the conserved regions- the sequence features. In this way, we assigned the predicted functional sites, predicted PTM sites, and predicted ligand binding sites, and ligand binding sites, and interaction interfaces for our model proteins.

#### 2. 1. 3. Visualization

The visualization of the primary and secondary structures of the proteins was performed using the Jalview 2.11. The PyMOL<sup>30</sup> software was utilized to represent and analyze the atomic structure of proteins.

#### 2. 1. 4. Molecular Docking Simulations

One of the most essential steps in this study is the molecular simulation as given in the workflow in Fig1. For the preparation of the docking process, hOct1-3 proteins were downloaded from the Robetta server in PDB format. All ligands (Metformin, Phenformin, and Norepinephrine) of hOct1-3 were retrieved within the SDF format from PubChem.<sup>31</sup> We removed the water, added the polar hydrogen to the model proteins. Then, we charged the model proteins and the ligands by the computation of Gasteiger before converting to pdbqt format using the AutoDock Vina.<sup>32</sup> Prior to the docking process, the 3D model protein structure was subjected to the energy minimization method in chimera 1.1433 with the default parameters: the steepest descent:100 with 0.02 step sizes, without fixing any atoms, followed by 10 steps of conjugate gradient steps with 0.02 step size (Å) minimization.

The docking study was carried out under the ACHIL-LES BLIND DOCKING SERVER3<sup>4</sup> protocol (https://bio-hpc.ucam.edu/achilles/). The figures were prepared using the PyMOL.

#### 2. 1. 5. Molecular Dynamics (MD) Simulations

To evaluate the structural constancy and validate the static description of the protein-ligand (hOCT-metformin) interactions, we ran an MD simulation using the Desmond Software.<sup>35</sup> The dynamic nature of protein-ligand interactions has been studied and atomic-level interactions were investigated.

#### 3. Results and Discussion

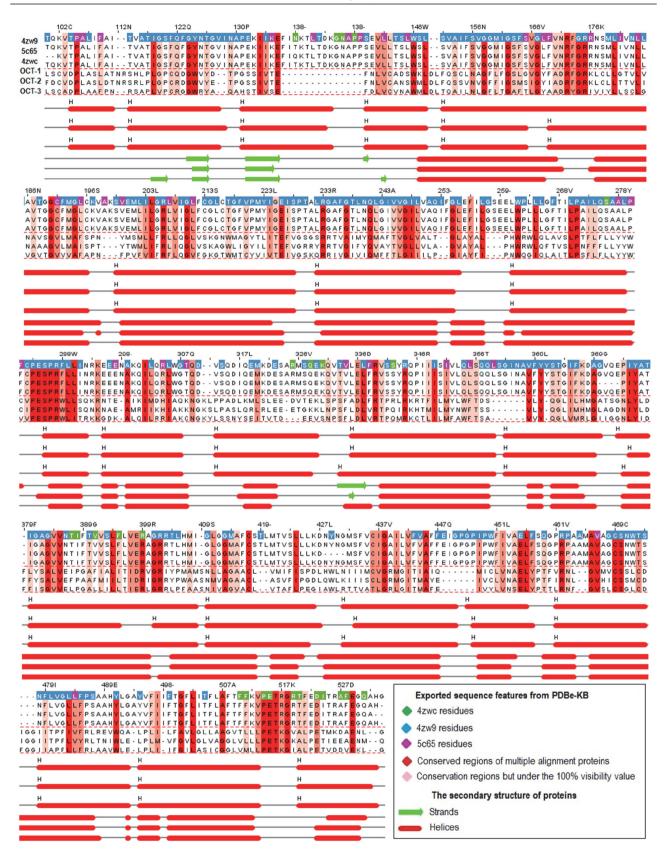
#### 3. 1. Alignments

We aligned a range of 146–445 aa of OCT-1 with 85–397 aa of 4ZW9 and 4ZWC as explained previously. The OCT-1 sequence has shown a 22.77% sequence identity with 4zw9 and 4zwc. We aligned a range of 146–540 aa of OCT-1 with of 63–477 aa of 5c65 by 22.51% of identity.

In the case of OCT-2, we aligned a range of 24–546 aa of OCT-2 with of 93–510 aa of 4ZW9 and 4ZWC, and 71–438 aa of 5c65 by the same percentage of identity ( 26.57% ).

**Table 1.** The list of the results of the sequence features analyzing multiple alignments. The residues of OCT 1-3 proteins that interact with metformin, phenformin, and norepinephrine and functional annotations of template proteins from PDBe-KB.

		Exported Functional Annotations of the templates proteins from PDBe-KB		
OCT-1				
Metformin	GLU <sup>137</sup> PRO <sup>481</sup> ARG <sup>488</sup>	Conserved Domain 4zw9; Predicted Ligand binding sites, 5c65; Ligand binding sites		
Phenformin	GLN <sup>152</sup> ASN <sup>156</sup> LYS <sup>214</sup> TRP <sup>354</sup> ASP <sup>357</sup> GLN <sup>362</sup> ILE <sup>446</sup>	4zw9; Predicted Ligand binding and functional sites, Ligand binding sites 4zw9; Ligand binding sites, Conserved Domain 4zw9; Predicted Ligand binding and functional sites 4zw9; Predicted Ligand binding sites, Conserved Domain 4zw9; Predicted Ligand binding and functional sites, Ligand binding sites 4zw9; Predicted Ligand binding and functional sites, Conserved Domain		
		OCT-2		
27.6	ASN <sup>157</sup>	4zw9; Predicted Ligand binding sites, Ligand binding sites, Conserved Domain 4zw9; Predicted Ligand binding sites, 4zwc; Predicted PTM sites,		
Metformin	CYS <sup>474</sup> ASP <sup>475</sup>	5c65; Ligand binding sites, Conserved Domain 4zw9; Predicted Ligand binding sites		
Phenformin	TYR <sup>37</sup> ASN <sup>157</sup> LYS <sup>215</sup> TYR <sup>245</sup> TYR <sup>362</sup> CYS <sup>474</sup> ASP <sup>475</sup>	4zw9; Predicted Ligand binding sites, Ligand binding sites, Conserved Domain 4zw9; Predicted Ligand binding and functional sites 4zw9; Predicted Ligand binding and functional sites, Ligand binding sites 4zw9; Predicted Ligand binding sites, Ligand binding sites, Predicted functional sites, 5c65; Predicted PTM sites, Conserved Domain 4zw9; Predicted Ligand binding sites, 4zwc; Predicted PTM sites, 5c65; Ligand binding sites, Conserved Domain 4zw9; Predicted Ligand binding sites		
		OCT-3		
Metformin	VAL <sup>37</sup> ASN <sup>162</sup> ARG <sup>212</sup> GLN <sup>366</sup>	4zw9; Predicted Ligand binding sites, Ligand binding sites, Conserved Domain 4zw9; Predicted Ligand binding sites, Ligand binding sites, Conserved Domain 4zw9; Predicted Ligand binding sites, Ligand binding sites, Conserved Domain		
Norepinephrine	PHE <sup>36</sup> VAL <sup>39</sup> GLN <sup>158</sup> ASN <sup>162</sup> ARG <sup>212</sup>	4zw9; Predicted Ligand binding and functional sites, Ligand binding sites 4zw9; Predicted Ligand binding sites, Ligand binding sites, Conserved Domain 4zw9; Predicted Ligand binding sites, Ligand binding sites, Conserved Domain		



**Figure 2.** The representation of primary and secondary structure of 4zw9, 5c65, and 4zwc, and OCT1-3 proteins. The visualization of sequence features and the Conserved Domain of the protein residues are colored through an analysis performed by the Jalview 2.11.0. The probability of conserved regions decreases through the dark red to the pink. After the multiple alignments, the OCT-1 protein was set as the reference for the sequence numbering. As a result of multiple sequencing, the overlapping regions of the proteins were solely exhibited.

In OCT-3, we aligned a range 60–353 aa of OCT-3 with of 86–513 aa of 4zw9 and 4zwc by 24.53% of identity, whereas we aligned a range of 84–353 aa of OCT-3 with of 64–473 aa of 5c65 by 24.53% of identity.

## 3. 2. Analysis of Conserved Domain and Sequence Features of OCT1-3

After subjecting OCT1-3 and template proteins sequences to multiple alignments, we detected the conserved

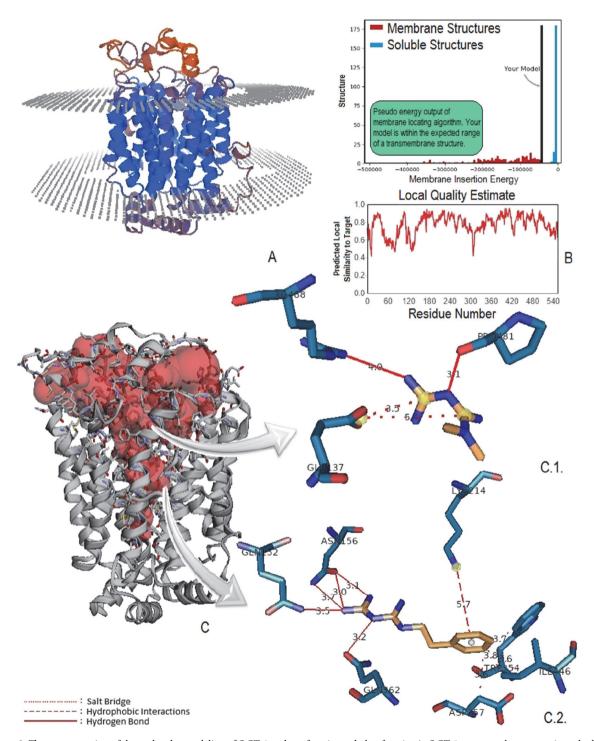


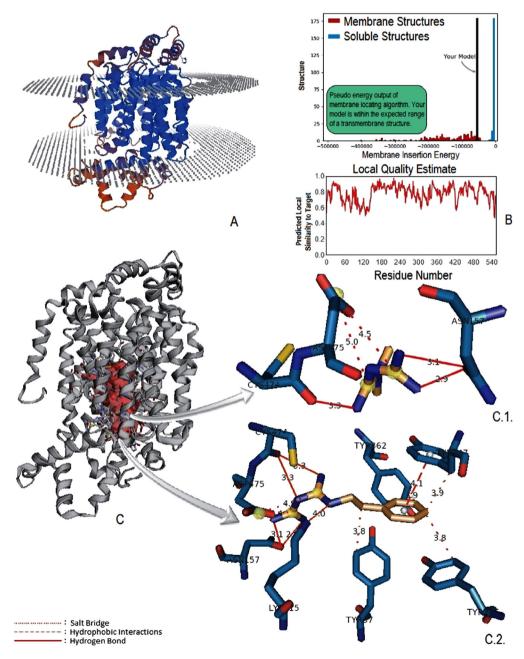
Figure 3. The representation of the molecular modeling of OCT-1 and metformin, and phenformin. A. OCT-1 transmembrane protein embedded in the plasma membrane model was predicted by QMEANBrane. B. Structure validation of modeled OCT-1 concerning membrane insertion energy and the local quality estimate of the residues of the model OCT-1. C. The surface of the binding pocket of the model OCT-1 as computed using CASTp 3.0. Molecular simulation of the best pose of the interaction of OCT1 and metformin(C1), phenformin(C2) with the highest docking scores.

**Table 2.** The list of the best binding energy poses of metformin, phenformin, and norepinephrine, and hOCT1-3 proteins through the Autodock Vina

Cluster Populations The highest binding energy (kcal/mol)				
	Metformin	Phenformin		
OCT-1	-4.60	-7.00		
OCT-2	-5.20	-8.60		
OCT 2	Metformin	Norepinephrine		
OCT-3	-5.27	-5.93		

regions through the comparative analysis in the Jalview 2.11.0 (See Fig2). We summarized the results of the comparative analysis as a list in Table1.

The cellular and biological functions of a protein are highly related to its 3D structure. The pharmacodynamics of a drug on the cell decreases or has no effect if the functional parts of these proteins are mutated in the genome. On the other hand, defining protein-ligand binding sites and explaining functional parts of the protein are critical approaches for drug discovery. <sup>36</sup> Regarding the pharmaco-



**Figure 4.** Representation of molecular modeling of OCT-2 and metformin, and phenformin. A. OCT-2 transmembrane protein embedded in the plasma membrane model was predicted by QMEANBrane. B. Structure validation of modeled OCT-1 concerning membrane insertion energy and the local quality estimate of the residues of the model OCT-2. C. The surface of the binding pocket of the model OCT-2 as computed using CASTp 3.0. Molecular simulation of the best pose of the interaction of OCT-2 and metformin(C1), phenformin(C2) with the highest docking scores.

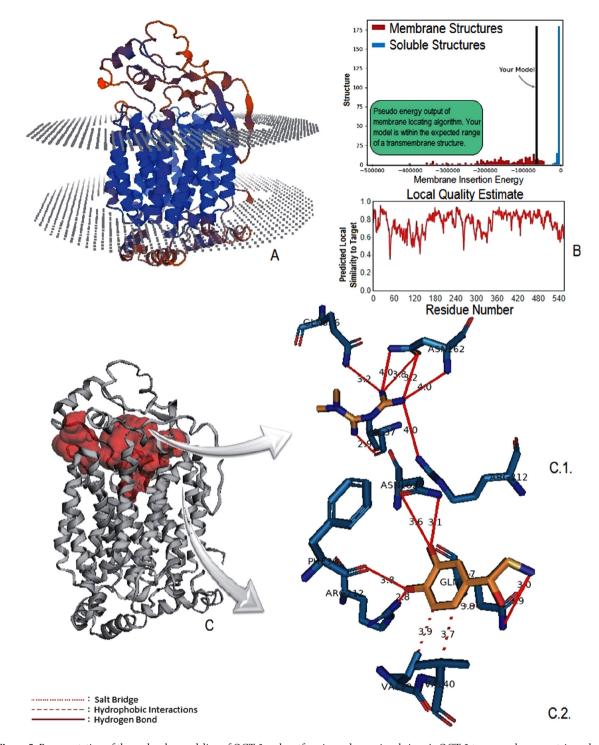


Figure 5. Representation of the molecular modeling of OCT-3 and metformin, and norepinephrine. A. OCT-3 transmembrane protein embedded in the plasma membrane model was predicted by QMEANBrane. B. Structure validation of modeled OCT-1 concerning membrane insertion energy and the local quality estimate of the residues of the model OCT-3. C. The surface of the binding pocket of the model OCT-3 is computed using CASTp 3.0. Molecular simulation of the best pose of the interaction of OCT-3 and metformin(C1), norepinephrine (C2) with the highest docking scores.

dynamics of metformin, the 3D prediction of OCT1-3 proteins and the determination of ligand binding sites in the functional sites are critical to investigate their effects on the cell.

Recent studies and meta-analyses have shown that patients with T2DM have a lower incidence of tumor development than healthy controls and cancer patient that use metformin has a lower risk of mortality.<sup>37</sup> Metformin

takes more attention after the discovery of its role in cancer prevention and treatment has been revealed. Improving or managing cellular uptake of therapeutic entities is mostly related to the understanding of the molecular mechanism of the interaction with the components of the cell membrane and therapeutic entities. This paper aimed to predict the 3D structure of OCT1-3 protein and identify its role in the uptake of metformin into the cells that have been studied by in vitro and in vivo studies previously.<sup>38,39</sup>

Sequence and structure analysis of proteins of unknown function with those of proteins of known function enable us to discover and deduce the function of the unknown proteins. Characterization of protein function by in vivo and in vitro studies is both time and labor-consuming. Furthermore, some proteins, especially membrane proteins are exceedingly difficult to be crystallized by experimental tools. In the modern genomic and proteomic era, a protein is mostly identified before its function is determined, therefore the role of in silico studies in structural analyses of proteins becomes more important in recent years.

The structure of OCT1-3 proteins has not been uncovered yet by any experimental tools although some of the protein's structures have already known in the same protein family. This paper is important for being the first attempt to study and predict the 3D structure of OCTs to reveal the information about how these proteins facilitate the uptake of metformin into the cells. Even though our analysis indicates no significant similarity between OCTs and the proteins of the database at a sequence level, the predicted OCTs are similar with its conservative regions to some carrier proteins that share a similar function.

It is known that 30 percent of all sequences are membrane proteins. Unlike globular proteins, a 3D model for membrane proteins can hardly be computed. Another important aspect of this paper is presenting a new pipeline to stimulate the docking of protein molecules in the absence of a similar sequence in the database. The recent algorithms in 3D structure prediction of proteins enable us to predict the structure of proteins in high accuracy even in the absence of sequence similarity. In silico analyses helped us to stimulate this biological process and propose the uptake of metformin by OCTs as it is shown in Fig 3–5.

Dakal et al. modeled the 3D structures of hOCTs by only one tool- I-TASSER in 2017. 40 In Fig 1, four key steps of this pipeline have been shown as the workflow. One of the very critical points, the prediction of the 3D structure of the protein, was performed by three different tools; Iterative Threading ASSEmbly Refinement, Phyre2 that uses protein homology, and Robetta. The output model proteins were then exposed to all proteins in the PDB by the calculation in the VAST. This approach is reflected in our results through an elevation in the accuracy in the protein structure prediction. We were eager to increase the accuracy of the prediction through the validation of these structures using the experimentally determined proteins as

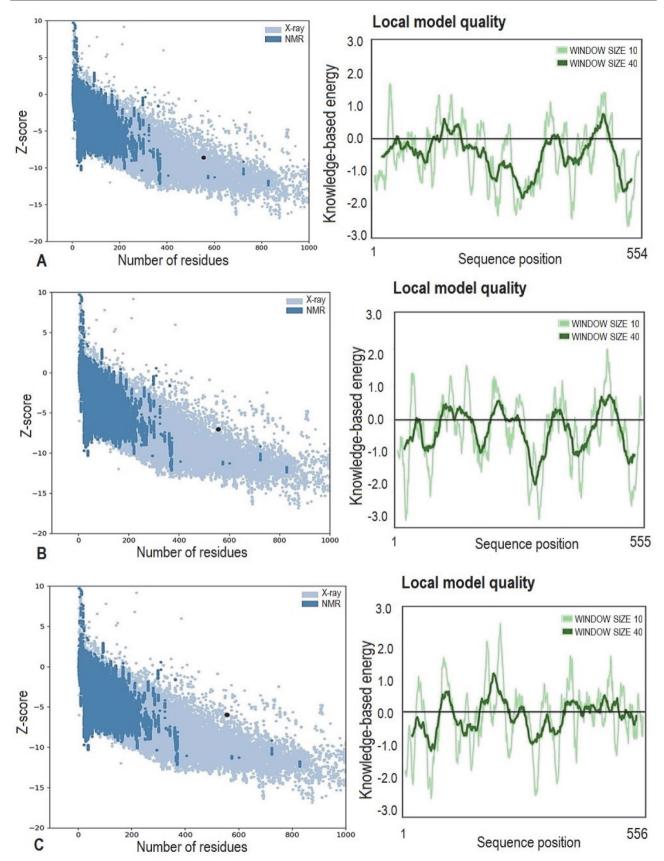
templates. After obtaining the structure of the OCTs, the orientation of these molecules in the plasma membrane was predicted using the QMEANBrane scoring function.

Transmembrane proteins play vital roles in a diverse range of essentially biological processes. Knowing about the protein position within the lipid bilayer is important and requires a computational approach, since identifying the correct orientation is possible by defining the relationship between sequence, structure, and the lipid environment. One of the commonly used tools to localize the structure of proteins within the lipid player by knowledge-based statistical potential, QMEANBrane was used and the predicted position as exhibited in Fig 3–5. As a result, all model proteins are within the expected range of transmembrane structures.

Models obtained from the other tools were determined to be inapplicable for the docking process. The Robetta is continuously evaluated with CAMEO (Continuous Automated Model EvaluatiOn), which constantly assesses the accuracy and reliability of the prediction. Among other prediction tools, CAMEO, Robetta, and QMEANBrane are the first-line with time-based statistical confidence and they show reliable performance. We also used the ProSAweb to verify the quality of the model protein structures. The Z-score designates the entire model-quality for OCT1-3, (Z-score:-8.59, -7.04, and -5.95, respectively) as shown in Figure 6.

To analyze sequence features, functional annotations of template proteins were retrieved from the PDBe-KB database. The recently released database, PDBe-KB, give us a great opportunity to analyze and visualize sequence features of the similar proteins that are used as the template to assign a novel function to our sequence of interest. Even though the sequence similarity is low, as shown in the results, there are significantly conserved regions. In this way, we assigned the predicted functional sites, predicted PTM sites, and predicted ligand binding sites, and ligand binding sites, and interaction interfaces to OCTs.

Representation of molecular modeling of OCTs and metformin was performed using the Blind Docking server. The server mainly utilizes a customized version of Autodock Vina<sup>32</sup> for the blind docking calculations. We obtained binding energy plots, and, in this way, the most energetically favorable dockings have been selected as the first best pose according to binding energy frequencies (See supplementary Fig1). Taking into account the model protein uncertainty as well as the small size of the metformin molecule, it is not surprising that many different ligands pose with similar scores. To cope with this, we used the CASTp bioinformatics tool and compared the predicted active sites of model proteins with the first best poses as the docking results. Interestingly, both output results from two servers were similar. Besides, for the results to be more meaningful, the pharmacologically important phenformin from metformin analogs was validated by the docking study of OCT-1 and OCT-2, whereas OCT-3 by



**Figure 6.** ProSA-web service analysis of human OCT1-3 proteins. The black points represent that model hOCT1-3 proteins are in the range of Z-score values of the experimental structures according to several residues. The other graph shows the local quality concerning many sequence positions (A; OCT-1, B; OCT-2, C; OCT-3). respectively (Z-score:-8.59, -7.04, -5.95).

the norepinephrine compound. We also combined these outputs with outcomes from exported functional annotations of the template proteins from PDBe-KB. We visualized the interaction of metformin, phenformin, and norepinephrine, and OCTs in PyMOL to better examine the poses and extract our images.

As listed in Table 1, OCT-1 forms hydrogen bonds with docked ligand molecules with the residue number of PRO481, ARG488, and GLN152, and ASN156, and GLN362. The other four residues in the predicted site (LYS214, TRP354, and ASP357, and ILE446) interacted by hydrophobic and salt-bridge bonds. Chen et al.<sup>41</sup> have re-

ported that OCT-1 interacts with its ligands by hydrogen binding and non-covalent interaction through the ASP357, TRP354, and ASN156, and ILE446 residues among their predicted residues. OCT-2 interacted with both metformin and phenformin through ASN157, CYS474, and ASP475 residues with noncovalent interactions such as hydrogen bond, salt bridge, and hydrophobic interaction (See Figure 4). OCT-3 protein contacts with norepinephrine and metformin in the same residue (ASN162 and ARG212) by hydrogen bonds. Given the extensive hydrogen bonding motif of metformin, water may be involved. This may significantly impact and alter the results and

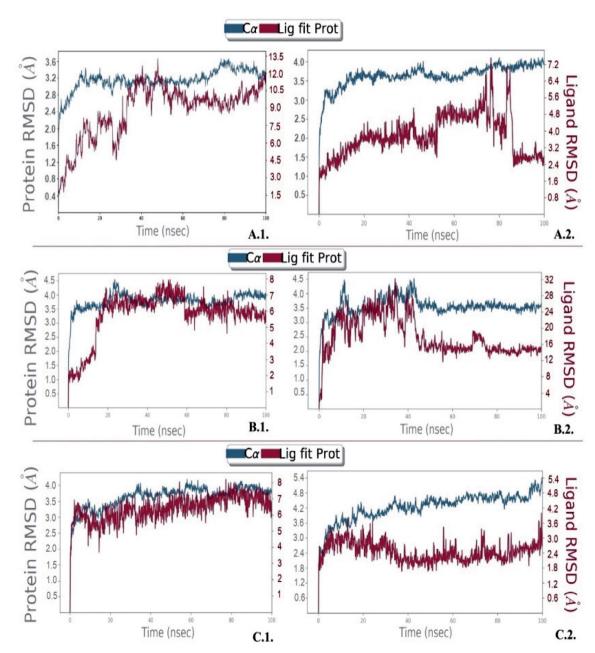


Figure 7: Desmond MD calculated Protein and Ligand RMSD: A.1.: OCT-1 and Metformin, A.2.: OCT-1 and Phenformin, B.1.: OCT-2 and Metformin, B.2.: OCT-2 and Phenformin, C.1.: OCT-3 and Metformin and C.2: OCT-3 and Norepinephrine

conclusions. Thus, we have considered performing the classic MD simulation for the docked complexes.

One of the residues that OCT-1 interacts with phenformin is GLN152 but OCT-3 interacts with norepinephrine through GLN158 as the same residue. The difference in the number of the residue is due to the setting of the sequenced reference. Our results suggest that human OCT proteins are predominantly expressed in different tissues of the human body and the active binding sites of these proteins also vary.

Although the methodology, including template definition, comparative protein modeling, and structure analysis, and molecular docking, seems pretty standard and employed in hundreds of research projects as in our workflow, there is a validation such as the quality control of the model proteins using Web services at almost every stage to increase reliability in achieving and evaluating meaningful results. Thus, the described pipeline is highly useful due to its ability to integrate the ligand-binding site and interaction interface information that is obtained from the PD-Be-KB database to the information that is derived from

similarity analysis and prediction tools. This pipeline is also promising to assign a function to predict the 3D structure even in the absence of any sequence similarity.

#### 3. 3. MD Simulations

Root mean square deviation (RMD) of protein and ligand was calculated during the MDS concerning their initial structure. RMSD of the OCTs shows its stable conformation throughout the simulation which indicates the stability of the interaction with metformin and phenformin. Besides OCT3 was stable with norepinephrine throughout the simulation.

Each OCT proteins attained equilibrium in a few nsec and remain stable throughout the simulation time up to 100 nces. Initially, the RMSD plot for metformin attained equilibrium in a few nsec as well and remain stable throughout to stimulation. Some deviations observed but no bigger changes of the order of 1–3 Å are seen in our analysis. Similar RMSD scores were recorded with the phenformin interactions as well.

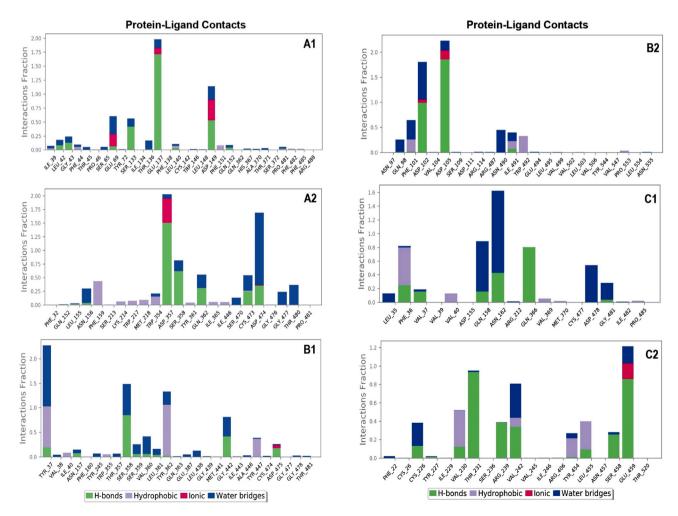


Figure 8: Protein-ligand contact interaction profile analyzed for the A1. OCT-1 and metformin, A2. OCT-1 and phenformin, B1. OCT-2 and metformin, B2. OCT-2 and phenformin, C1. OCT-3 and metformin, and C2. OCT-3 and norepinephrine

OCT1-3 interactions with the ligands were monitored throughout the simulation. The interactions are rich with H bonds these play significant roles in ligand binding.

Even though the involvement of human OCT in the uptake/transport of metformin is already mentioned in the literature, this study is demonstrating the hOCTs-metformin interaction at the atomic level for the first time and describing how and where the binding occurs. Mimicking this binding with the absence of the structural information of the protein was possible with the unique approach that was described in the pipeline in Figure 1.

#### 4. Conclusion

The three-dimensional structure of a protein is a direct association with its comprehensive cellular and biological function. To investigate the tertiary atomic structure of OCT1-3 proteins and their localization in the cell membrane, it is significant to evaluate the pharmacodynamics of metformin, frequently preferred in Type 2 Diabetes Mellitus medication, through the determination of the residues which interact with metformin in the case of cell translocation of these proteins. One of the important limitations of mimicking protein-ligand interactions is the absence of the protein 3D structure. The presented new pipeline is promising especially for the interaction simulation studies that are conducted with proteins with unknown structures. To determine the therapeutic effect of Metformin or other life-saving drugs into the cells, further studies are needed to examine genetic variants of human OCTs in specific patient populations. Analyzing insertions, deletions, and other genetic variants effects on hOCTs in structure level are important to explore the role of these proteins in metformin pharmacokinetics and response. Our study could be a front preparation and inspiration for future positively charged drug discoveries and development by examining the atomic level of OCT proteins.

#### **Conflict Of Interest**

The authors declare that they do not have any conflict of interest.

#### Acknowledgements

We are thankful to Mungo Carstairs and his coworkers from the Barton group for their support in our study in the Jalview 2.11. This work has been funded by grants from the Spanish Ministry of Economy and Competitiveness (CTQ2017-87974), and by Fundacioń Seńeca de la Regioń de Murcia under Project 20988/PI/18. This research was partially supported by the supercomputing infrastructure of Poznan Supercomputing Center, the e-infrastructure program of the Research Council of Norway via the super-

computer center of UiT—the Arctic University of Norway, and by the supercomputing infrastructure of the NLHPC (ECM-02), Powered@NLHPC. This research project has been co-financed by the European Union (European Regional Development Fund—ERDF)

#### 5. Reference

- T. M. Bakheet, A. J. Doig, *Bioinformatics*. 2009, 25, 451–457.
   DOI:10.1093/bioinformatics/btp002
- 2. L. Fagerberg, K. Jonasson, G. Von Heijne, M. Uhlén, L. Berglund, *Proteomics.* **2010**, *10*, 1141–1149.

DOI:10.1002/pmic.200900258

- 3. S. Tan, T. T. Hwee, M. C. M. Chung, *Proteomics.* **2008**, 8, 3924–3932. **DOI**:10.1002/pmic.200800597
- A. Marchler-Bauer, Y. Bo, L. Han, J. He, C. J. Lanczycki, S. Lu, F. Chitsaz, M. K. Derbyshire, R. C. Geer, N. R. Gonzales, et al., Nucleic Acids Res. 2016, 45, D200-D203.

DOI:10.1093/nar/gkw1129

- C. Colas, P. M. U. Ung, A. Schlessinger, Medchemcomm. 2016, 7, 1069–1081. DOI:10.1039/C6MD00005C
- I. Pernicova, M. Korbonits, Nat. Rev. Endocrinol. 2014, 10, 143. DOI:10.1038/nrendo.2013.256
- L. Chen, B. Pawlikowski, A. Schlessinger, S. S. More, D. Stryke,
   J. Johns, M. A. Portman, E. Chen, T. E. Ferrin, A. Sali, et al.,
   Pharmacogenet. Genomics. 2010, 20), 687.

DOI:10.1097/FPC.0b013e32833fe789

8. A. T. Nies, H. Koepsell, K. Damme, M. Schwab, *Handb. Exp. Pharmacol.* **2011**, 105–167.

DOI:10.1007/978-3-642-14541-4 3

- R. Saitoh, K. Sugano, N. Takata, T. Tachibana, A. Higashida, Y. Nabuchi, Y. Aso, *Pharm. Res.* 2004, 21, 749–755.
  - **DOI:**10.1023/B:PHAM.0000026423.48583.e2
- D.-S. Wang, J. W. Jonker, Y. Kato, H. Kusuhara, A. H. Schinkel, Y. Sugiyama, *J. Pharmacol. Exp. Ther.* **2002**, *302*, 510–515.
   **DOI:**10.1124/jpet.102.034140
- R. J. Shaw, K. A. Lamia, D. Vasquez, S.-H. Koo, N. Bardeesy,
   R. A. DePinho, M. Montminy, L. C. Cantley, *Science* (80-.).
   2005, 310, 1642–1646. DOI:10.1126/science.1120781
- 12. H. D. McIntyre, C. A. Paterson, A. Ma, P. J. Ravenscroft, D. M. Bird, D. P. Cameron, *Aust. N. Z. J. Med.* **1991**, *21*, 714–719. **DOI:**10.1111/j.1445-5994.1991.tb01375.x
- M. R. Owen, E. Doran, A. P. Halestrap, *Biochem. J.* 2000, 348, 607–614. DOI:10.1042/bj3480607
- J. E. Gunton, P. J. D. Delhanty, S.-I. Takahashi, R. C. Baxter, J. Clin. Endocrinol. Metab. 2003, 88, 1323–1332.
   DOI:10.1210/jc.2002-021394
- A. Drozdetskiy, C. Cole, J. Procter, G. J. Barton, *Nucleic Acids Res.* 2015, 43, W389–W394. DOI:10.1093/nar/gkv332
- L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, *Nat. Protoc.* 2015, *10*, 845.
   DOI:10.1038/nprot.2015.053
- Y. Song, F. DiMaio, R. Y.-R. Wang, D. Kim, C. Miles, T. J. Brunette, J. Thompson, D. Baker, *Structure* **2013**, *21*, 1735–1742.
   **DOI:**10.1016/j.str.2013.08.005

- 18. S. Raman, R. Vernon, J. Thompson, M. Tyka, R. Sadreyev, J. Pei, D. Kim, E. Kellogg, F. DiMaio, O. Lange, et al., *Proteins Struct. Funct. Bioinforma.* 2009, 77, 89–99. DOI:10.1002/prot.22540
- A. Roy, A. Kucukural, Y. Zhang, *Nat. Protoc.* 2010, 5, 725.
   DOI:10.1038/nprot.2010.5
- 20. J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, *Nat. Methods* **2015**, *12*, 7. **DOI**:10.1038/nmeth.3213
- J. Yang, Y. Zhang, Nucleic Acids Res. 2015, 43, W174-W181.
   DOI:10.1093/nar/gkv342
- G. Studer, M. Biasini, T. Schwede, Bioinformatics 2014, 30, i505-i511. DOI:10.1093/bioinformatics/btu457
- 23. M. Wiederstein, M. J. Sippl, *Nucleic Acids Res.* **2007**, *35*, W407-W410. **DOI:**10.1093/nar/gkm290
- J.-F. Gibrat, T. Madej, S. H. Bryant, Curr. Opin. Struct. Biol. 1996, 6, 377–385. DOI:10.1016/S0959-440X(96)80058-3
- W. Tian, C. Chen, X. Lei, J. Zhao, J. Liang, Nucleic Acids Res.
   2018, 46, W363–W367. DOI:10.1093/nar/gky473
- S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* 1990, 215, 403–410.
  - **DOI:**10.1016/S0022-2836(05)80360-2
- 27. H. McWilliam, W. Li, M. Uludag, S. Squizzato, Y. M. Park, N. Buso, A. P. Cowley, R. Lopez, *Nucleic Acids Res.* 2013, 41, W597-W600. DOI:10.1093/nar/gkt376
- A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, G. J. Barton, *Bioinformatics* 2009, 25, 1189–1191.
   DOI:10.1093/bioinformatics/btp033
- Pdb.-K. consortium, *Nucleic Acids Res.* 2019. DOI:10.1093/nar/gkz853.

- 30. Schrödinger, LLC, The {PyMOL} Molecular Graphics System, Version~1.8, 2015.
- S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, et al., *Nucleic Acids Res.* 2016, 44, D1202–D1213. DOI:10.1093/nar/gkv951
- 32. O. Trott, A. J. Olson, J. Comput. Chem. 2010, 31, 455-461.
- E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* 2004, 25, 1605–1612. DOI:10.1002/jcc.20084
- I. Sánchez-Linares, H. Pérez-Sánchez, J. M. Cecilia, J. M. García, *BMC Bioinformatics* 2012, *13*, S13.
   DOI:10.1186/1471-2105-13-S14-S13
- S. Release, others, Maest. Interoperability Tools, Schrödinger, New York, NY 2017.
- Q. Wu, Z. Peng, Y. Zhang, J. Yang, Nucleic Acids Res. 2018, 46, W438-W442. DOI:10.1093/nar/gky439
- 37. F. Zi, H. Zi, Y. Li, J. He, Q. Shi, Z. Cai, Oncol. Lett. 2018, 15, 683-690.
- A. T. Nies, U. Hofmann, C. Resch, E. Schaeffeler, M. Rius, M. Schwab, *PLoS One* **2011**, *6*, e22163.
   DOI:10.1371/journal.pone.0022163
- M. Foretz, B. Guigas, L. Bertrand, M. Pollak, B. Viollet, *Cell Metab.* 2014, *20*, 953–966.
   DOI:10.1016/j.cmet.2014.09.018
- T. C. Dakal, R. Kumar, D. Ramotar, Comput. Biol. Chem. 2017, 68, 153–163. DOI:10.1016/j.compbiolchem.2017.03.007
- E. C. Chen, N. Khuri, X. Liang, A. Stecula, H.-C. Chien, S. W. Yee, Y. Huang, A. Sali, K. M. Giacomini, *J. Med. Chem.* 2017, 60, 2685–2696. DOI:10.1021/acs.jmedchem.6b01317

#### Povzetek

Metformin, zdravilo, ki ga bolniki s sladkorno boleznijo pogosto uporabljajo kot prvi izbor zdravljenja po vsem svetu, je pozitivno nabit in se v celice vnaša s pomočjo človeških transportnih proteinov za organske katione (hOCT 1-3). Z različnimi metodami in bioinformatičnimi orodji smo želeli posnemati celični privzem metformina s hOCT1-3. 3D struktura proteinov OCT1-3 je bila napovedana z upoštevanjem struktur in funkcij teh proteinov. Predvideli smo funkcionalne regije (aktivna in ligand-vezavna mesta) OCT1-3 in izvedli primerjalno bioinformatično analizo. Napovedana struktura hOCT1-3 je bila nato analizirana na strežniku Blind Docking in rezultati potrjeni s predvidenimi preostanki vezavnih mest in ohranjenimi regijami domen. Simulirali smo sidranje OCT1-3 in metformina ter potrdili postopek sidranja z drugimi substrati proteinov HOCT1-3. Izbrali smo najboljše konformacije simuliranja sidranja metformina glede na energijo vezave (-5,27 do -4,60 kcal/mol). Nazadnje smo statični opis interakcij protein-ligand (OCT-Metformin) potrdili s simulacijo molekularne dinamike. Pri tem smo dosegli stabilno stimulacijo interakcije OCT-metformin.



Except when otherwise noted, articles in this journal are published under the terms and conditions of the Creative Commons Attribution 4.0 International License