

Investigating the Interactions of Human Organic Cation Transporter 1 and Enteropeptidase with Metformin and Its Derivative to Alleviate Side Effects in Diabetes Treatments

Phoebe Tszki Zhang

*Basis International School of Guangzhou, Guangzhou, China
phoebezhangzzz1023@gmail.com*

Abstract. Diabetes is a chronic disease that is characterized by elevated blood sugar levels. Once diagnosed, patients will live with various types of treatments for the rest of their lives. The disease can burden families and result in damage to other organ systems or death. The most common treatment for type 2 diabetes is an oral medicine called metformin, coupled with life style changes and other therapies. The drug acts by binding to the protein called human organic cation transporter 1 (hOCT1). Although the medicine is low-cost, effective, and was recognized several health organizations, it still leads to side effects. The interaction of the drug with enteropeptidase is known to cause impaired digestion, which results in diarrhea, nausea, and other gastrointestinal side effects. In attempt to enhance the therapeutic effectiveness of the drug while reducing the side effects, a new derivative of metformin that has greater binding affinity to hOCT1 and lower binding affinity to enteropeptidase. Key interactions and residues were identified, and by adding an aromatic ring with substituent of hydroxyl and amine groups to the drug, H bonding and Pi-Pi interactions with hOCT1 were enhanced, increasing the binding efficacy. On the other hand, addition of bulkier substituent disrupted the H bond between the drug and a key residue HIS 57.B, allowing for the binding to be less effective in turns to minimize the side effects.

Keywords: metformin, type 2 diabetes, gastrointestinal side effects, human organic cation transporter 1, enteropeptidase

1. Introduction

Diabetes, known as a chronic disease, results from congenital insufficient insulin production (type 1) or bodily changes that counteracts insulin's effects or prevents its production (type 2). A third type, gestational diabetes, is prevalent during pregnancy and potentially affects the well being of the baby. In simpler terms, a diabetic patient will have elevated blood sugar levels that might affect other organ systems negatively. More severe and harmful consequences of diabetes are often associated with complications of other conditions of the body. If diabetes remains untreated for a prolonged period of time, it can affect lives of patients as it stimulates blindness, cardiovascular problems, kidney failure, and other health issues [1]. Rising diabetes prevalence has been an increasingly severe problem. There are various treatments needed in order to prevent complications of the

disease, and these costly therapies are likely burdensome for a lot of families. The major costs can include payments for medicine and care, as well as accumulation of other expenses. If diabetes lead to other complications, such as kidney failure or liver problems, the costs of treatment can exceed \$80,000 per person per year. As a result of the expensive treatments that continues to remain inaccessible and unaffordable to populations of lower socioeconomic status, the number of diabetic patients and patients that are left untreated has been rising.

As newly collected data states, an estimated number of 589 million adults are diagnosed with diabetes, with the value likely to reach 853 million in 2050. The disease led to approximately 3.4 million deaths in 2024 and increased spending on health for people. Statistics analyzed from detailed inspections claim that around 90% of diabetic patients are diagnosed with type 2 diabetes. Even though the disease used to be more prevalent in middle-aged populations, recently, it was found that teens that are inactive and obese have increased chances of becoming diabetic. Presently, diabetes costs cover around 12% of global health expenditures, with the proportion expected to continually increase [2]. Thus, socioeconomic status, which accounts for factors such as income and level of education became important for access to diabetes treatment. In summary, there are significant differences in diabetes prevalence rates among high (~6.6%), middle (~11.1%), and low (~12.3%) income countries, suggesting strong association between socioeconomic status and diabetes prevalence for a region [3]. This is also a phenomenon observed in the United States, where the burden of diabetes is distributed unevenly between various racial groups of differing economic conditions, impacted by awareness of the disease (education received), surrounding community, and other factors [4]. The environment around them shapes their understanding and knowledge regarding the disease, including the necessity of keeping healthy daily routines, doing regular physical health checks, and to ask for assistance when needed [5]. Racial minorities of lower economic status may not have opportunities to acquire these important knowledge, highlighting the correlation between the two factors and diabetes prevalence.

Examining the expenditures to treat diabetes in more detail, the most common form of treatment for type 2 diabetes is the ingestion of the oral medicine metformin, which is what is investigated in this study. As a reliable source of treatment affirmed by American Diabetes Association and the American Academy of Clinical Endocrinologists, patients with type 2 diabetes are often prescribed with metformin if the use of the medicine does not negatively impact their condition. Metformin is affordable and can reduce the probability of cardiovascular complications for diabetic patients, and the doses taken aligns with life style changes and the injection of insulin as well in order to maximize its effect in lowering blood sugar and storing energy. Although greatly effective, metformin treatment also stimulate side effects that includes vitamin B12 deficiency, hypoglycemia, liver failure, and gastrointestinal side effects. This study takes into account the side effects and aims to reduce gastrointestinal side effects of metformin while enhancing its efficacy. In order to reduce the side effects of metformin, a detailed examination of the structure and mechanism of the drug is required [6]. The chemical structure of metformin is characterized by bonds between carbon and nitrogen atoms, as depicted in Figure 1A and Figure 1B.

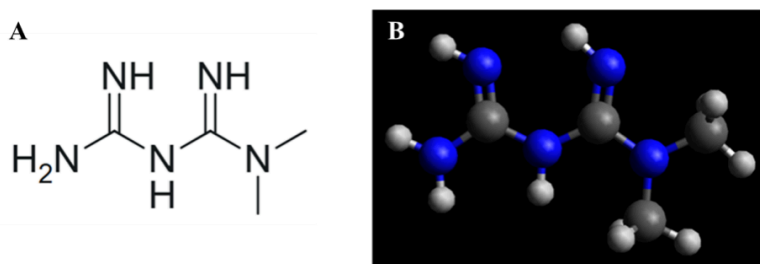


Figure 1. The structure of metformin is shown. (A) 2D structure of metformin drawn using software Chemaxon is illustrated. (B) 3D structure of metformin modeled using Avogadro software is shown (picture credit: original)

The drug performs its function through using the AMPK-dependent and AMPK-independent mechanism, in which AMPK abbreviates for AMP-activated kinase. The chemical processes occur in the liver, and was found to restrain mitochondrial glycerophosphate dehydrogenase or mitochondrial respiration, suggesting that lysozyme is involved in the mechanisms [7]. Metformin assists the regulation of blood sugar levels in the AMPK-independent pathway by affecting the liver; it can increase responsiveness to insulin and lessen the production of gluconeogenic enzymes in the AMPK-dependent pathways [7]. Metformin acts by interacting with multiple proteins. An example of a protein involved in its function will be human organic cation transporter 1 (hOCT1), a main focus in this study. Previous research has confirmed hOCT1 as an essential protein that aids the absorption of metformin in liver cells, which is a significant factor for metformin to regulate blood glucose levels in the body [8]. In particular, Hydrogen bonding and pi interactions between amino acids and aromatic structures are determining features that affect the binding affinity of metformin with the protein [9]. As discussed previously, there are also side effects that couple the ingestion of metformin. Enteropeptidase, for instance, is one of the enzymes that are crucial for digestion and works in the gastrointestinal tract in animals and humans. It performs its function by stimulating the conversion of trypsinogen, another protein, into trypsin. Yet past literature has identified gastrointestinal side effects to be associated with the interaction between enteropeptidase and metformin [10].

Metformin binds in the inhibitor site to enteropeptidase, which prevents the conversion of trypsinogen. Deficient amounts of trypsin will hinder protein digestion, leading to nausea and vomiting. In regards to the significance enteropeptidase has in causing gastrointestinal side effects, metformin is hypothesized to bind to an indicated inhibitor site in enteropeptidase [11]. Past studies conducted on metformin had objectives to propose derivatives improve the efficacy of the drug by making it effective for prolonged periods of time and increasing its solubility while reducing acid acidosis side effects [12]. However, there are certainly other areas of focus that will impact the overall effectiveness of the drug. My study proposes a new derivative of metformin to improve the therapeutic effects and to minimize gastrointestinal side effects by examining past literature and bridging gaps between known knowledge and novel research. In this paper, after careful consideration of chemical properties that will impact interactions between ligands and receptors, a new derivative of metformin will be tested. The binding affinities between metformin and the two proteins, hOCT1 and enteropeptidase, will be compared with the binding affinities between the derivative of metformin and the proteins. Ideally, the derivative should have more negative binding scores to hOCT1 and more positive binding scores with enteropeptidase, indicating that the software suggests stronger interactions between hOCT1 and the derivative and less favorable interactions

between enteropeptidase and the derivative, indicating success in fulfilling the goals of improving the efficacy of the drug while reducing gastrointestinal side effects.

2. Materials and methods

2.1. Software and database

Software tools utilized to assist this study include Discovery Studio, Chemaxon, Avogadro, UCSF Chimera, and AutoDock Vina [13-15]. Database used to retrieve proteins was the RCSB Protein Data bank [16].

2.2. Procedure and preparation

For the purpose of this research, a derivative of metformin was drawn using Chemaxon. Then, the drug was modeled on Avogadro and the universal force field approximation tool was used to alter the drug to the most ideal molecular geometry. The derivative was saved and fetched later for docking.

hOCT1 was fetched from PDB Databank (PDB ID 8JTS) using Chimera. The protein was dock prepped and the box was constructed around residue MF8 to indicate where the original metformin drug and the derivative will bind to. MF8 was removed, metformin was introduced, and the binding score between the unmodified drug and hOCT1 was recorded. On the other hand, enteropeptidase was fetched from PDB Databank (PDB ID 1EKB) using Chimera. The protein was dock prepped and the box was constructed around a proposed inhibition site to indicate where the original metformin drug and the derivative will bind to. Water molecules were removed, metformin was introduced, and the binding score between the unmodified drug and 1EKB was recorded.

2.3. Docking process

Data presented in this study was generated using the AutoDock Vina and Chimera. Visualizations displayed was generated using Discovery Studio. During the docking process for hOCT1, the box that was constructed around MF8 had 3D dimensions with a center at 107, 107, and 111 with a size shown as 6, 8, and 10. After that, the residue MF8 was removed and replaced with the unmodified metformin constructed in Avogadro. Default algorithms were set for the docking process, and the optimal orientation was determined with regards to that of MF8; the binding score in the optimal orientation was recorded. In enteropeptidase, the box that was constructed around the inhibitor binding site that was identified to be near HIS57 had 3D dimensions with a center at 9, 11, and 14 with a size shown as 15, 14, and 23. After that, the unmodified metformin constructed in Avogadro was introduced and docked. Default algorithms were set for the docking process, and the binding score was recorded. Qualitative representations and quantitative data were collected to be presented and discussed.

3. Results and discussion

This study aims to identify a novel derivative of metformin that enhances therapeutic effects and reduces gastrointestinal side effects that are caused by the ingestion of the oral medicine. To attain this goal, the derivative needs to bind with hOCT1 more strongly and its interaction with 1EKB must be weaker.

The structure of metformin is shown in Figure 2A and Figure 2B. The structure of metformin derivative is shown in Figure 2C and Figure 2D. Modifications made to the original metformin include the removal of two methyl groups attached to a nitrogen atom and the replacement of the groups with an aromatic ring that has substituent of a hydroxyl and amine group shown in Figure 2C and Figure 2D. In comparison to the original drug, the derivative is consisted of more polar and bulkier groups, making it an ideal analog since the bulky groups are expected to disrupt the interaction between the drug and enteropeptidase while the polar groups can contribute to the formation of H-bonds with hOCT1. Therefore, the binding affinity between metformin and hOCT1 will increase and the binding affinity between metformin and enteropeptidase will decrease.

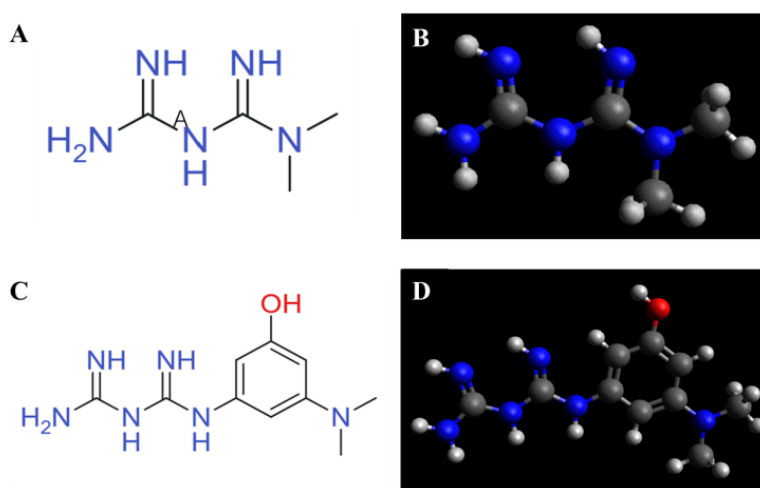


Figure 2. The structure of metformin and metformin derivative is shown. Blue spheres indicate nitrogen atoms and the red sphere indicate oxygen atom. (A) 2D structure of the metformin is shown. (B) 3D structure of metformin is shown (C) 2D structure of the metformin derivative is shown. (D) 3D structure of metformin derivative is shown (picture credit: original)

Displayed by data recorded in Table 1, an overview of the results suggests a trend that corresponds with the objective of this study. The derivative has more negative binding scores with hOCT1 compared to the original drug, showing an increase in binding affinity. On the other hand, the derivative has more positive binding scores with enteropeptidase compared to the original drug, showing a decrease in binding affinity.

Table 1. Binding scores of human organic cation transporter 1 and enteropeptidase with metformin and its derivative

Ligand	Receptor	Binding Score	RMSD l.b	RMSD u.b.
Original Metformin	hOCT1 (8JTS)	-4.195	1.435	2.37
Metformin Derivative	hOCT1 (8JTS)	-5.143	2.729	6.812
Original Metformin	Enteropeptidase (1EKB)	-4.114	1.519	4.291
Metformin Derivative	Enteropeptidase (1EKB)	-3.773	3.241	4.216

One aim of this study is to lower the binding score between hOCT1 and metformin in order to increase the binding affinity and to improve the efficacy of the drug. hOCT1 was docked prepped to record the initial binding score by constructing a box around MF8, removing the residue, then introducing metformin to the binding site in the optimal orientation. The protein with PDB ID 8JTS

was directly downloaded from RCSB PDB and loaded into Discovery Studio to generate the binding maps. The metformin corresponds to the residue MF8, which was selected as the ligand. As illustrated in the 2D binding map, there is a hydrogen bond between metformin and tyrosine, which suggests that the incorporation of the hydroxyl group in the derivative has the potential to increase the number H bonds to enhance binding affinity in Figure 3A. The 3D depiction demonstrates how metformin is spatially arranged in the binding pocket shown in Figure 3B. The closeness between nitrogen on the metformin with the colored surfaces that indicate H-bond acceptor and donors shows again the importance to consider H-bond interactions. In addition, it affirms the introduction of hydroxyl and amine groups as beneficial for the interactions.

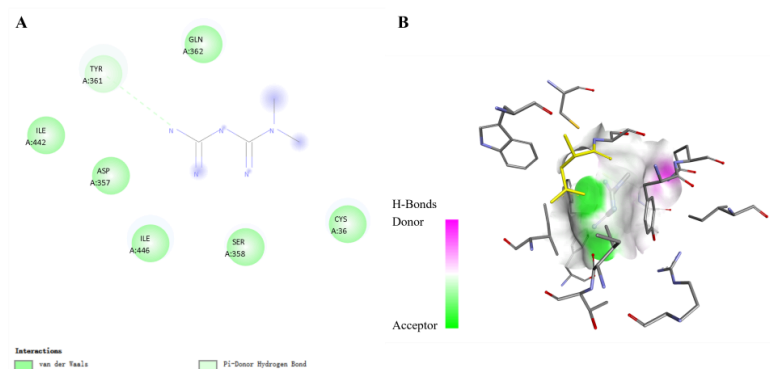


Figure 3. The binding map between original metformin and hOCT1 is shown. (A) 2D interactions between metformin and key residues in the complex of unmodified metformin and hOCT1 are shown. (B) H bond interactions in the complex of unmodified metformin and hOCT1 are shown, indicating H-bond acceptors and donors (picture credit: original)

The derivative of metformin was docked into hOCT1 in the region designated by the same box constructed previously for MF8. In the 2D binding map, there are H-bonds, Pi-Alkyl and Pi-Pi shaped interactions shown by Figure 4A. In comparison to the visualization with the original drug, the new binding map indicates more interactions that are more specific and stronger. The conventional bond that formed between the hydroxyl group and aspartate indicates electrostatic interactions that helps to stabilize the derivative at the binding site; the aromatic benzene ring that formed Pi-Pi T-shaped interactions, likewise, contributes to the stabilization of the structure in Figure 4A. The factors allow for greater binding affinity between hOCT1 and the derivative demonstrated by a decrease in the binding score. The 3D model suggests the positioning of the derivative in a cavity that allows for interactions to be maximized between H-bond acceptor and donors indicated by the colored surfaces in Figure 4B.

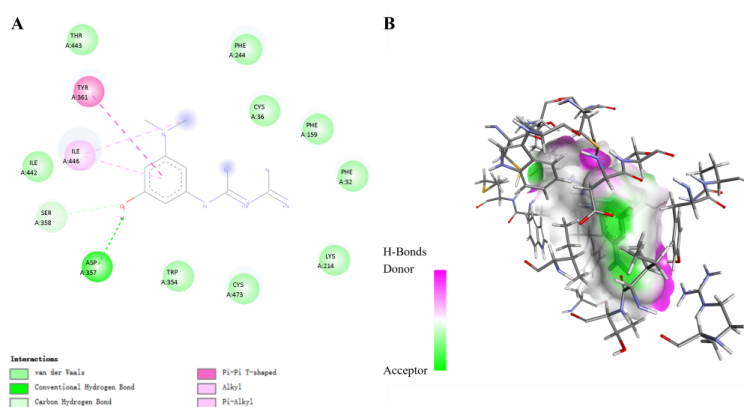


Figure 4. The binding map between metformin derivative and hOCT1 is shown. (A) 2D interactions between the derivative and significant residues in the complex of metformin derivative and hOCT1 are shown. (B) H bond interactions in the complex of metformin derivative and hOCT1 are shown, indicating H-bond acceptors and donors (picture credit: original)

The other aim of this study is to make the binding score between hOCT1 and metformin more positive in order to decrease the binding affinity and to reduce gastrointestinal side effects that occur. Unlike hOCT1, where there is a pocket that metformin is confirmed to bind to, researchers have not found a designated site where metformin will interact with enteropeptidase. But there are past literature that suggest that interactions between metformin and amino acid HIS 57.B are crucial to understand the binding of the drug to the protein based on known properties of the residue. Histidine, with a charged R group, along with aspartate and serine, is indicated as one of the catalytic triad found in enteropeptidase that catalyzes reactions [17]. Hence, it is able to form electrostatic interactions with metformin that has positively-charged nitrogen atoms. The strong interaction anchors metformin in the binding site, allowing for interactions between ring in the side chain of histidine and the amine group. Therefore, the box was constructed around HIS 57.B and the original metformin was introduced to the protein for docking. Demonstrated by the 2D binding map, there are carbon H bonds identified between metformin and HIS 57 in Figure 5A. The 3D illustration shows that metformin is in close proximity to colored portions in Figure 5B, indicating potential interactions between H-bond acceptor and donors. In order to reduce binding affinity and make the binding score more positive, steric clashes are needed to disrupt the interaction between the drug and histidine.

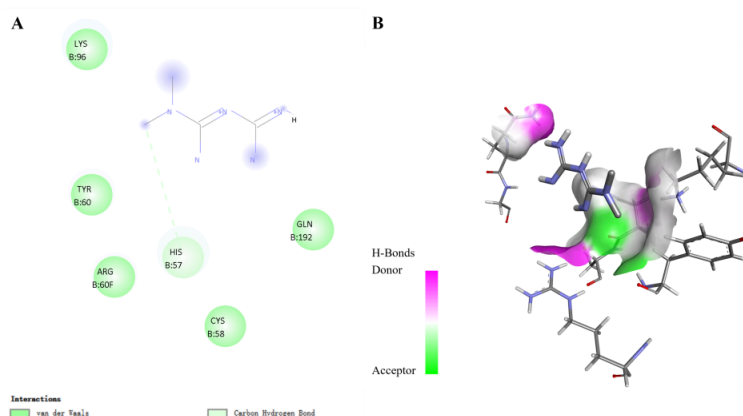


Figure 5. The binding map between original metformin and enteropeptidase is shown. (A) 2D interactions between metformin and significant residues in the complex of unmodified metformin and enteropeptidase. are shown. (B) H bond interactions in the complex of unmodified metformin and enteropeptidase are shown, indicating H-bond acceptors and donors. (picture credit: original)

There are significant differences between the binding maps showing the interactions between enteropeptidase and metformin and the binding maps showing the interactions between enteropeptidase and the derivative. As shown in the 2D binding map, the interaction between the derivative and HIS57.B is successfully disrupted with the formation of a new H-Bond between the derivative and lysine; to add on, unfavorable positive-positive interactions emerged between the nitrogen cations, arginine, and lysine, resulting in repulsion that contributes to decreasing the binding affinity, shown in Figure 6A. Even though the new analog of the drug is still positioned to allow interactions with H-bond donors and acceptor, with the elimination of the significant interaction between histidine and the derivative caused by the addition of the aromatic ring and bulkier substituent, the orientation of the derivative was deviated from the optimal arrangement, as illustrated by Figure 6B. Hence, the binding score between the receptor and the ligand became more positive, indicating lower binding affinity and reduction of side effects that occurs because of the binding of metformin onto the inhibitor site of enteropeptidase.

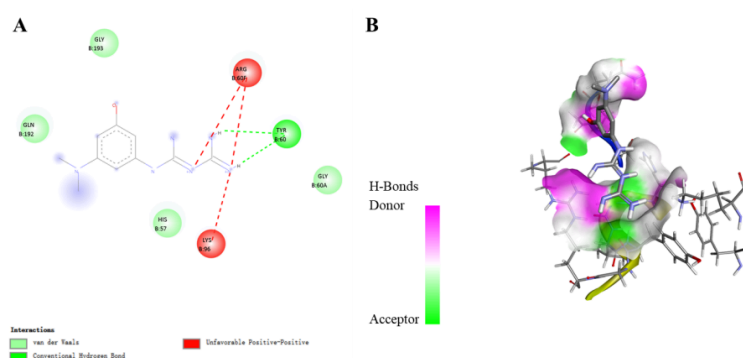


Figure 6. The binding map between the metformin derivative and enteropeptidase is shown. (A) 2D interactions between the derivative and significant residues in the complex of unmodified metformin and enteropeptidase are shown. (B) H-bond interactions in the complex of unmodified metformin and enteropeptidase are shown, indicating H-bond acceptors and donors (picture credit: original)

To conclude, the data gathered demonstrates the potential of the derivative being more effective in treatments while leading to less gastrointestinal side effects. This change is shown through the

more negative binding scores between the derivative and hOCT1 and more positive binding scores with enteropeptidase. Yet limitations of the study needs to be taken into account. First, the spatial orientation of the molecules in their hypothesized binding site needs to be confirmed for more accurate results. Although past research confirmed that metformin will bind to enteropeptidase, the binding sites used in this study were hypothesized with regards to properties of the molecule and surrounding key residues, which presents the likelihood of inaccurate estimations. Secondly, the binding sites and binding maps generated by the software lack simulation in the actual environment that the molecules will bind and exist in. The condition of the body will be influenced by various factors overlooked in the online models, including change in temperature and presence of other compounds. The sophistication to model complicated situations may result in distinctions between what is concluded in this research and under real conditions. Finally, the representations generated limits movements of the molecules, restraining contextual knowledge necessary for the binding process to be understand in more depth. The limitations set goals for studies that can be done in the future and emphasize the need for consistent effort to learn more about the drug.

4. Conclusion

Based on quantitative data collected using UCSF Chimera, adding hydroxyl and amine groups that are capable of having H-bonds and also aromatic rings that stabilize the bonding helps to enhance the therapeutic effectiveness of metformin (demonstrated by more negative binding scores to hOCT1) while minimizing gastrointestinal side effects (demonstrated by more positive binding scores to enteropeptidase). According to qualitative visualizations created by Discovery Studio, spatial orientation of molecules are affected by the addition of substituent, impacting the interactions between the key residues and the drug and subsequently altering binding affinities. Locating more specific binding pockets and identifying optimal orientations while attempting to stimulate more realistic conditions for the interactions of the drug will be foundational to continue the research process, giving more insight to how the derivative can be improved in future studies.

References

- [1] Ogbeide, S.A. (2017) Review of Integrated behavioral health in primary care: Step-by-step guidance for assessment and intervention (Second edition). Families Systems & Health, 35, 391.
- [2] Reed, J., Bain, S. and Kanamarlapudi, V. (2021) A Review of Current Trends with Type 2 Diabetes Epidemiology, Aetiology, Pathogenesis, Treatments and Future Perspectives. Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy, 14, 3567–3602.
- [3] Dagenais, G. R., Gerstein, H. C., Zhang, X., McQueen, M., Lear, S., Lopez-Jaramillo, P., Mohan, V., Mony, P., Gupta, R., Kutty, V. R., Kumar, R., Rahman, O., Yusoff, K., Zatonska, K., Oguz, A., Rosengren, A., Kelishadi, R., Yusufali, A., Diaz, R. and Avezum, A. (2016) Variations in Diabetes Prevalence in Low-, Middle-, and High-Income Countries: Results From the Prospective Urban and Rural Epidemiological Study. Diabetes Care, 39, 780–787.
- [4] Liu, C., He, L., Li, Y., Yang, A., Zhang, K. and Luo, B. (2023) Diabetes risk among US adults with different socioeconomic status and behavioral lifestyles: evidence from the National Health and Nutrition Examination Survey. Frontiers in Public Health, 11, 1197947.
- [5] Contesa, A., Malini, H. and Rahman, D. (2025) Comprehensive risk analysis of diabetic peripheral neuropathy in type II diabetes mellitus. Journal of Diabetes and Metabolic Disorders, 24, 123.
- [6] Sparkes, S. T., Patel, D. K. and Goldman-Levine, J. D. (2024) Past, Present, and Future Research Avenues for Metformin. Journal of Pharmacy Technology, 30, 227–234.
- [7] Rena, G., Hardie, D. G. and Pearson, E. R. (2017) The mechanisms of action of metformin. Diabetologia, 60, 1577–1585.
- [8] Umehara, K-I., Iwatsubo, T., Noguchi, K. and Kamimura, H. (2007) Functional involvement of organic cation transporter1 (OCT1/Oct1) in the hepatic uptake of organic cations in humans and rats. Xenobiotica; the Fate of

Foreign Compounds in Biological Systems, 37, 818–831.

- [9] Zeng, Y. C., Meghna Sobti, Quinn, A., Smith, N. J., Simon, Vandenberg, J. I., Ryan, R. M., O'Mara, M. L. and Stewart, A. G. (2023) Structural basis of promiscuous substrate transport by Organic Cation Transporter 1. *Nature Communications*, 14, 6374.
- [10] Kelly, C., Verdegaaal, A. A., Anderson, B., Shaw, W. L., Bencivenga-Barry, N. A., Folta-Stogniew, E. and Goodman, A. L. (2023) Metformin inhibits digestive proteases and impairs protein digestion in mice. *Journal of Biological Chemistry*, 105363–105363.
- [11] Lu, D., Fütterer, K., Korolev, S., Zheng, X., Tan, K., Waksman, G. and Sadler, J. E. (1999) Crystal structure of enteropeptidase light chain complexed with an analog of the trypsinogen activation peptide. *Journal of Molecular Biology*, 292, 361–373.
- [12] Hernández-Velázquez, E.D., Alba-Betancourt, C., Alonso-Castro, Á.J., Ortiz-Alvarado, R., López, J.A., Meza-Carmen, V. and Solorio-Alvarado, C.R. (2023) Metformin, a biological and synthetic overview. *Bioorganic & Medicinal Chemistry Letters*, 86, 129241.
- [13] Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E. and Hutchison, G. R. (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, 4, 17.
- [14] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25, 1605–1612.
- [15] Trott, O. and Olson, A. J. (2009) AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring function, Efficient optimization, and Multithreading. *Journal of Computational Chemistry*, 31, 455–461.
- [16] Berman, H., Henrick, K. and NakaFmura, H. (2023) Announcing the worldwide Protein Data Bank. *Nature Structural & Molecular Biology*, 10, 980–980.
- [17] Hedstrom, L. (2022) Serine Protease Mechanism and Specificity. *Chemical Reviews*, 102, 4501–4524.