RNA-Based Aptamer Targeting Clusterin Protein as a Diabetes Biomarker

IRENE BATTA[1,2] AND GAURAV SHARMA[2]

[1]Bothell High School, WA, Bothell
[2]Eigen Sciences, NC, Apex

Compiled November 18, 2024

Clusterin protein is primarily used as a diabetes biomarker and can be used to detect cell density in diabetes. A lower β cells are an early sign of type 1 diabetes. Aptamers are single-stranded nucleotides that bind the target protein and are prevalent in disease diagnosis and targeted therapy. Recently, aptamer targeting clusterin protein has been developed as a biomarker of the β cells. We hypothesize that the rigid clusterin protein should have specific amino acids interacting with the particular aptamers. In this research work, we have performed in silico simulations to find the optimal aptamer capable of binding to the clusterin protein can be used as a biomarker in diabetes. Specifically, we have performed molecular docking simulations to find the clusterin-aptamer binding complexes. These aptamers were selected based on their inter-actions with the clusterin protein. Phylogenetic analysis shows that the aptamers class is bifurcated into two sections. All the results were validated using the GrASP web server, which shows the aptamer binding region on the protein's surface. The finding indicates that the aptamer binds to clusterin and could potentially be used as a diabetes biomarker.

1. INTRODUCTION

Diabetes is a long-term metabolic disease marked by high blood glucose levels (1). It is sometimes referred to as diabetes mellitus (1). Over time, diabetes can cause severe damage to the kidneys, nerves, eyes, heart, and blood vessels (2). The body's inability to produce enough insulin results in type 1 diabetes, but Type 2 diabetes is caused by inefficient insulin utilization in the body (3). Moreover, gestational diabetes raises the mother's and the child's future chance of developing Type 2 diabetes during pregnancy (4). Diabetes is becoming increasingly common worldwide and is becoming more severe in middle- and low-income nations (1). This presents a serious public health concern and emphasizes the need for quick decision-making and coordinated efforts to prevent and treat this crippling condition (5).

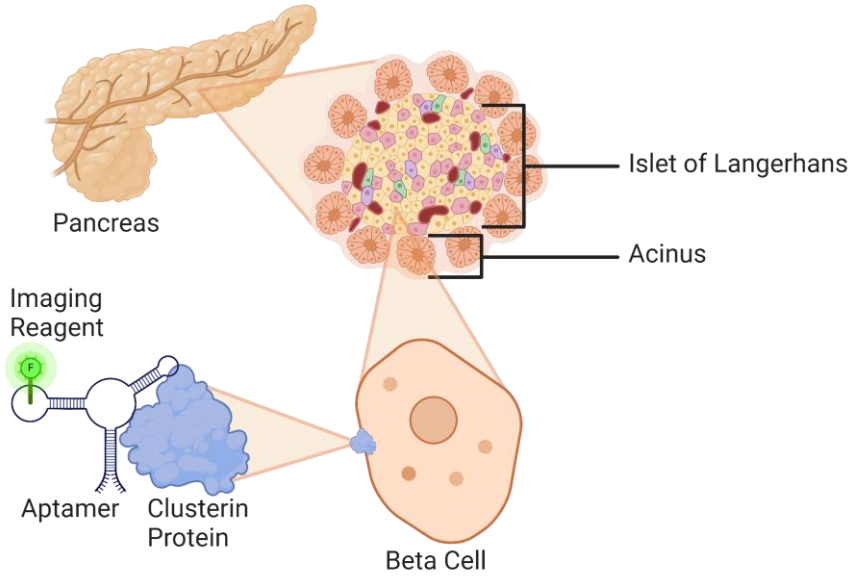

Fig. 1. Scheme of the current research. Beta cells are present in the Islet of Langerhans in the pancreas. In diabetes, the clusterin level increases in serum and tissue and, therefore, can be used as a diabetes biomarker. The aptamers used in this research specifically bind to the clusterin protein. These aptamers can be conjugated with imaging reagents and used in beta cell mass quantification (Figure created with BioRender.com).

Because it is involved in several physiological processes frequently dysregulated in diabetic settings, clusterin, also known as apolipoprotein J, has potential as a biomarker in diabetes (6). Clusterin is a secreted glycoprotein made inside cells and transported outside the cell (7). It is formed in organs such as the liver, kidneys, brain, and adipose tissue (8). When cellular stress, inflammation, or tissue damage—all prevalent in diabetes—clusterin production and release can be increased, which helps shield cells from harm, lessen inflammation, and encourage tissue repair (1). Diabetes is characterized by metabolic dysregulation and chronic inflammation, which have an impact on clusterin expression and release (9). Inflammatory cytokines, oxidative stress, and elevated glucose levels all promote the synthesis of clusterin (10). Measuring blood levels of clusterin may help in early diagnosis and monitoring of diabetes complications, especially diabetic nephropathy, and cardiovascular illnesses, which are linked to elevated blood levels of this protein (11). Research has demonstrated a correlation between the severity of a disease and clusterin levels; higher levels may signify a more advanced disease or more significant tissue damage (12). Monitoring clusterin levels may reveal information about the efficacy of treatment measures and act as a predictive biomarker for the onset of problems in diabetes patients (13). To increase diagnostic accuracy and prediction value for diabetes complications, measurement methodologies must be standardized, large-scale validation studies must be conducted, and clustering must be integrated with other biomarkers for clinical use (14). Figure 1 shows the location of clusterin protein and also the scheme of the current

research.

A short, single-stranded DNA or RNA molecule known as an aptamer has a high affinity and selectivity for binding to particular target molecules (15,16). The systematic procedure known as SELEX is used to select these compounds. Because of their comparable binding properties to antibodies, aptamers are frequently used interchangeably, but they also have unique benefits such as simpler manufacturing, less immunogenicity, and increased thermal stability. Aptamers are used in many contexts, such as biosensors, medicinal agents, and diagnostic instruments. They can be designed to target unhealthy cells, such as cancer cells, or to obstruct particular biological pathways in therapies. Aptamers can be diagnostic tools to find proteins, small compounds, or other disease-related biomarkers. They are instrumental in research and medical contexts due to their excellent specificity in binding to various targets. Molecular docking is a computational method for predicting how a molecule will bind to another to form a stable compound (17-21). It plays a crucial role in drug development by simulating interactions between small compounds, like potential therapeutic candidates, and their biological targets, typically proteins. By identifying these molecules' binding affinity and specificity, molecular docking helps design more effective and targeted medications. This method relies on algorithms that explore various binding poses and rank them based on their stability, aiding in understanding molecular interactions and creating new drugs.

Since clusterin protein is expressed in beta cells, its expression level could be correlated with the beta cell biomass. Increased clusterin expression in beta cells is related to beta cell oxidative stress and could eventually lead to Type 1 diabetes. Since the clusterin protein is released into the bloodstream, it is easy to measure its plasma level, which could indicate the beta cell mass or cell stress level. Based on this theory, Simaeys' et al. recently designed an RNA-based aptamer targeting clusterin protein as a pre-diabetic biomarker (22). We hypothesize that the binding of aptamers should be due to their unique shape and sequence. Therefore, we have performed computational simulations to understand the aptamer-biomarker protein interactions. Specifically, we designed aptamers' 3D structures and later docked them on clusterin protein.

2. METHOD

AlphaFold 3 predicted the protein structure of clusterin protein (23). The models were acquired, and their structural accuracy was confirmed. The Vfold2D and Vfold3D servers received the aptamer sequences and produced 2D and 3D structural predictions, respectively (24,25). The stability and possible binding conformations of the secondary and tertiary structures were examined. Using molecular docking software, the 3D structures of the aptamers were positioned onto the anticipated protein structures. The binding locations and interaction strengths were determined by analyzing the docking results. Chimera software was used to visualize the individual structures and the docked complexes (26) (27). For additional investigation, figures showing the structural alignments and interactions were created. An aptamer binding site on the proteins was predicted using a graph neural network (GNN). Known aptamer-protein interactions were used to train and validate the GNN model. The protein sequences were subjected to phylogenetic analysis using MEGA software (28). Phylogenetic trees were generated to determine links and evaluate the conservation of

the aptamer binding sites.

3. RESULTS

We believe the aptamer should attach to a particular binding site on the clusterin protein. We have used graph neural networks using the GrASP web server to forecast this site is druggable (26). In Figure 3, the druggable site is displayed. This indicates that the druggable and binding sites are yellow and green. Understanding the surface properties of the clusterin is crucial in validating the clusterin-aptamer docking results. Vfold2D was utilized to estimate the secondary structure of aptamers (24). Aptamer 2D is vital in defining the three-dimensional aptamer shape and provides information on base pair interactions (hydrogen bonds). A popular method for representing the secondary structure of nucleic acids, including aptamers, is dot and bracket notation, Figure 2. Several symbols, such as dots and brackets, are used to denote paired and unpaired bases in this notation. In this notation, paired bases in a stem are represented by parenthesis( ) and unpaired by dots (.). A pairing begins with an opening parenthesis, "(," and ends with a closing parenthesis, ")." The dot-bracket notation is shown in Table 1, and the 2D image is shown in Figure 3.

Table 1.

Aptamer Name	Sequence dot and bracket notations
A	UAAUUCUCAGGAGGUGCGGAACGGGAUAUGGAUUGUUCGC ...((((....))))(((.((((((.....))))))))))
B	CAACAAACUAAUCAGACACGAGACAGAGAGAUAGAUCUGCCAGA .....(((((((((....))))((((((.....))))))))))

Sequence of the aptamers. The aptamer sequences and dot-bracket notations help in getting the 2D image of this DNA.

Additionally, we used CrustalW, a bioinformatics computer tool for multiple sequence alignment, to carry out the sequence alignment (29). The sequences were organized using sequence alignment to identify common portions across aptamer sequences that might suggest structural and functional similarity. A phylogenetic tree is called a branching diagram or "tree" that illustrates the connections between the various aptamer sequences. Aptamer sequences are aligned during this process to arrange similar or identical nucleotides in the same column. From SELEX, 39 aptamers were selected for further evaluation and sequence alignment was divided into a set of two. Therefore, we have chosen only two of these two sets' aptamers. A phylogenetic tree analysis was used to confirm this further, as shown in Figure 5.

Figure 6 illustrates the tertiary structure of aptamers, which results from the secondary structure's folding. The function of aptamer depends on these three-dimensional structures, which also control how well an aptamer interacts with a protein. Per our forecast, every aptamer possesses a distinct three-dimensional structure, which is evident in how they bind together in docking simulations (explained subsequently). Figure 4a and b displays the clusterin-aptamer docked structures. The figure shows that the clusterin-aptamer interaction

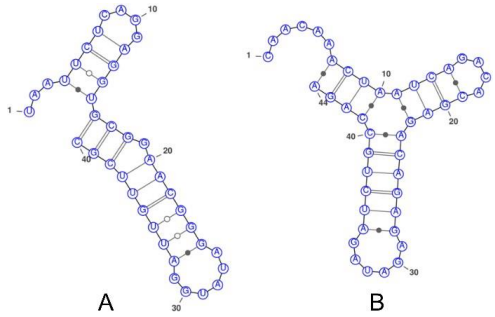

Fig. 2. Aptamer secondary structures. This shows the bounding formed between base pairs.

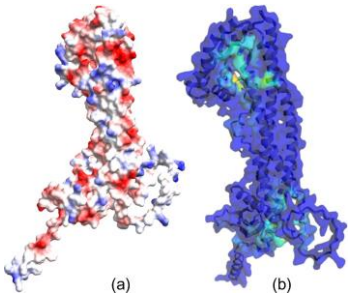

Fig. 3. The structures of clusterin protein. (a) shows the protein's electrostatic surface potential (ESP), and (b) shows the binding site in a yellow-green color.

happens to the same binding spot predicted by the graph neural network. In addition, we have also obtained the clusterin-aptamer binding interaction using AlphaFold 3, which also shows a similar binding site. After validating these results, we further calculated the binding interaction between the clusterin-aptamer complex using the PLIP web server, which shows that aptamer A formed 16 hydrogen bonds and six salt bridges. On the other hand, aptamer B formed 17 hydrogen bonds and five hydrogen bonds, Table 2.

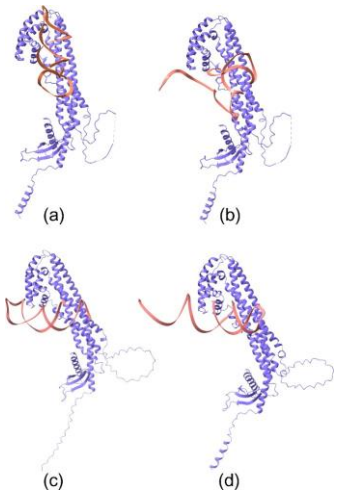

Fig. 4. Clustrin-aptamer binding. (a) and (b) are the docking results of aptamer A and B, respectively, and (c) and (d) are the structures predicted by AlphaFold 3 of aptamer A and B, correspondingly.

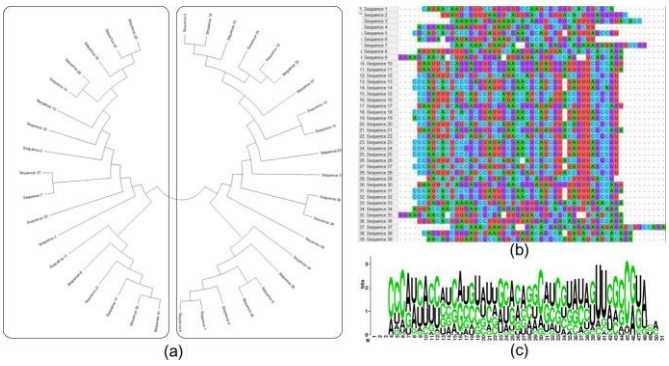

Fig. 5. Aptamer analysis. The phylogenetic tree of aptamer used in the SE-LEX shows that aptamers were divided into two sections.

4. DISCUSSION

Aptamers have been used in diabetes detection before. For example, Lee et al. developed ssDNA-based aptamers that bind to the retinol-binding protein for the early detection of diabetes (30). These aptamers were attached to a gold chip, and SPR was used to detect the aptamer-protein specificity. They have also stated that these aptamers are more sensitive than ELISA assay (31). In a different study, Apiwat et al. have developed graphene-based aptasensor diagnosing diabetes mellitus (32). Moreover, aptamer has been used to identify SecinH3 in human liver cells (33). Apart from diabetes mellitus, aptamer has been used in various metabolic diseases such as aptamer targeting adipose

tissues could have an impact on obesity and metabolic syndrome (34). Till now, aptamer targeting is as follows: HbA1c, GHSA, RBP4, adipocyte cell line, and visceral adipose tissue-derived serpin (35). Noninvasive aptamer-mediated therapy targeting pancreatic B cells will be a valuable tool and pave the way for treating diabetes and its complications. We hope future research will focus on the aptamer-mediated inhibitory effect of T cells and macrophages in diabetes complications (36).

First, the quality of the models and algorithms employed significantly impacts the overall accuracy of the molecular docking, and poor models can and often do produce false positives. In the future, we hope to incorporate the behavior of the protein in the presence of water surrounding its ligand through molecular dynamics simulations. MD simulations will also reveal the strength of the interaction and whether the ligands remain with the protein during the simulation. Moreover, ITC or SPR should be used to validate the computational research after further analysis further. These assays will prove that the suggested ligands bind to the protein in the anticipated manner.

In this work, we investigated the potential of aptamers as diabetes diagnostic agents by studying their interactions with the biomarker clusterin protein, which is implicated in cellular stress and inflammation associated with diabetes. This aptamer can bind to the β cells and could be used in the β cell mass for the early detection of Type 1 diabetes. We found particular binding sites on clusterin where aptamers produced stable complexes characterized by numerous hydrogen bonds and salt bridges. Our results show that aptamers can specifically target and bind to clusterin, indicating its potential use in the early detection and tracking of problems associated with diabetes. The accuracy and effectiveness of aptamer-based diagnostics were demonstrated by applying graph neural networks (GNNs), which improved the identification of druggable locations on the protein. The research highlights the potential of aptamers in developing tailored treatments and diagnostic instruments for diabetes, opening the door to more precise and efficient treatment. To combine the clusterin protein with other biomarkers for better diagnostic and predictive potential, future studies should concentrate on standardizing measurement techniques and carrying out extensive validation studies.