IN SILICO EVALUATION OF VIRAL INSULIN-LIKE PEPTIDES AS POTENTIAL THERAPIES FOR DIABETES

Massimo D. Bezoari†, Alexis Balbuena*

Northwestern State University, Louisiana Scholars’ College, Natchitoches, LA 71497

Abstract

Diabetes is a disease in which the body cannot respond to or produce insulin. People who are affected with this disease have to take synthetic insulin for the rest of their lives. This research seeks to explain insulin-like properties of selected viral insulin-like peptides (VILPs) for diabetes therapies using protein-protein docking. Schrödinger’s Biologics software was used to dock VILPs to crystallography-based structures of the human insulin receptor (6PXV). The results demonstrate that viral insulin-like compounds bind to the insulin receptor of human cells, at the same sites as with insulin as well as other sites. This work corroborates reported laboratory research showing that VILPs are effective in performing the same functions as insulin. The current work promotes the development of less expensive and potentially better synthetic insulin mimics for the treatment of type I diabetes.

†Corresponding author: bezoarim@nsula.edu.
Keywords: Diabetes, Viral Insulin-Like Peptides, Schrödinger Protein-Protein Docking Insulin Receptor.

Introduction

In 2018, 34 million people in the U.S. had diabetes, 210,000 of whom were children and young adults under the age of 20. This number is expected to triple by the year 2050. The number of children diagnosed with type II diabetes is increasing due to childhood obesity which has more than tripled since the 1970s (1, 2). Synthetic insulin is used to treat diabetes but is very expensive and only slightly similar to the structure of insulin that the body naturally produces (3). Reported research shows that viral insulin-like peptides (VILPs) have similar effects as insulin (4). One of the goals of the present work was to determine whether VILPs activate the insulin receptor (IR) in the same way as insulin and might therefore be suitable substitutes for insulin. This research promotes the development of new synthetic alternatives for the treatment of diabetes.

Diabetes is a disease in which the body cannot respond to or produce insulin. Type I diabetes, also known as “juvenile onset diabetes,” occurs when the pancreas does not produce any insulin at all. The immune system produces proteins called autoantibodies that attack the beta cells in the pancreas causing them to stop producing insulin. The cause of this autoimmune reaction is unknown. The symptoms of this disease are polydipsia (excess thirst), polyuria (excess urination), and weight loss. Type II diabetes, also known as “adult onset diabetes,” occurs when the body’s insulin receptors cannot respond to the insulin that is secreted from the body. This disease is strongly correlated with obesity. Both types of diabetes result in the body not being able to regulate blood sugar levels resulting in hyperglycemia and ketoacidosis, which can be deadly. Chronic health complications that are closely associated with diabetes include heart failure, kidney failure, heart attack, stroke, high blood pressure, and permanent nerve damage. Research relating the causal relationship between insulin deficiency and diabetes occurred throughout the 1900s. In 1978, the first prescription of human insulin was made. Currently, type I diabetes is treated with synthetic insulin, diet, and exercise (5).

Structure and Function of Insulin

There are two different forms of insulin that occur naturally in the body, an open “active” form, and a closed “storage” form. The storage form is called proinsulin. This structure consists of A, B, and C chains linked by disulfide bonds. Six cysteine residues in the structure of insulin form the interchain and intrachain disulfide bonds (6).

As insulin converts to its active form, the C chain is cleaved causing the remaining structure to be “open.” This protective hinge process enables binding to the receptor. Other factors that could influence the conformation of insulin include the presence of glucose or glycogen (7).

When insulin binds to the insulin receptor (IR) of a cell, it causes the cell to import glucose for energy to be used for growth, repair, and other cellular functions (8). If there is too much glucose in the body, insulin signals the liver, muscle, and adipose cells to convert glucose to glycogen. If there is not enough glucose in the blood, insulin signals these same cells to convert glycogen to glucose and release it back into the blood. Without insulin, blood sugar levels may rise to dangerously high levels. With no energy source, the body is forced to catabolize fatty acids that are stored in adipose tissue. Effects of these reactions include production of ketones, lowering of pH and denaturation of enzymes. This condition is called ketoacidosis (8).

Insulin Receptor Structure

The IR is formed by two monomers each containing an α chain and a β chain (Figure 1).

Beginning at the N-terminus end of a monomer chain, there are six segments: leucine rich domain L1, comprising residues 1-157; a cysteine rich segment (CR), residues 158-310; another leucine-rich sequence, L2 (311-470); then fibronectin domains FnIII-I (471-595), FnIII-2 (596-808), and FnIII-3 (residues 809 to about 906); an insert domain (ID, 638-756) occurs within FnIII-2 about 906); an insert domain (ID, 638-756) occurs within FnIII-2 and within the αCT region (residues 704-719). The β-chain begins within the ID segment. The β chain of the receptor passes through the cell membrane after residue 906 and ends with the tyrosine kinase domain (7, 9-11).
structural analyses indicate that each primary binding site (S1) is between L1 of one monomer and the αCT and FnIII-1 domains of the other (11, 12). The secondary site, S2, has been the subject of research aiming to elucidate its structural make-up and its role in the activation of the IR by insulin. Apparently, the S2 site becomes available only after insulin has complexed at the primary site, S1. However, the role of S2 is still under investigation.

Since the receptor is composed of two monomers, there are two tyrosine kinases in the cytoplasm of the cell. Tyrosine kinase domains function in phosphorylating certain proteins to catalyze essential metabolic reactions. When insulin binds to the receptor, intracellular tyrosines in a “mobile loop” are phosphorylated, causing the loop to dissociate from the kinase active site and leaving it in an open conformation. When the tyrosine kinase domain is active, ATP binding causes phosphorylation of target substrates, including the IR substrate. This initiates many different pathways essential for growth and health. These pathways include the Ras-MAPK/Erk (MAPK/Erk) pathway, which activates T-cells, phosphorylates transcription factors in the nucleus to turn on genes, initiates the production of new epithelial cells, and promotes long term potentiation in the neurons of the hippocampus. Once these cascades are signaled, a glucose transporter, GLUT4, is translocated to the cell membrane, and molecules of glucose are imported for energy, cellular growth, storage as glycogen, or conversion to triglycerides (9, 13).

Peptide hormones similar to insulin and insulin growth factor-1, i.e., viral insulin-like peptides (VILPs), belong to a family of viruses called iridovirae that cause cancerous growths in fish (4, 14). Specific VILPs are lymphocystis disease virus-Sa (LCDV-Sa), Singapore grouper iridovirus (SGIV), lymphocystis disease virus-1 (LCDV-1), and grouper iridovirus (GIV). These viruses have been identified in human fecal and blood samples (14). They show structural similarities to insulin, including the occurrence of all six cysteine residues essential for intrachain and interchain disulfide bonds that are components of insulin, and play a role in binding to the insulin receptor. The synthesis of VILPs has been reported. They have been subjected to in vitro and in vivo experiments and found to lower blood glucose levels in human cells and mice. It was also found that all the VILPs activate insulin receptors. Additionally, although specific measurements such as IC_{50} are not available to compare the glucose uptake effects of VILPs with those of insulin, competition studies have shown that the four VILPs compete very effectively with insulin for the IR binding site (14).

Therefore, previous research reports imply that VILPs might conduct cellular signaling such as cell growth and cell proliferation in addition to the uptake of glucose. In fact, the VILPs also exhibited metabolic effects similar to insulin. The VILPs also showed effects on the Ras-MAPK pathway that were comparable to insulin.

**Computational Chemistry Strategy**

In the current work, a computational chemistry study was performed using BioLuminate 3.6 in the Schrödinger software to simulate binding of VILPs to the insulin receptor. No previous computational docking research has been reported, to our knowledge. The VILP structures were imported from the original experimental reports (14). The structure of human insulin described in the PDB file 3I3Z was selected for this experiment because of its open conformation, i.e., the conformation that insulin takes in binding to the IR. The PDB file of the dimeric insulin receptor, 6PXV, was used as the receptor. This dimeric structure includes binding of 4 insulin
molecules—2 at the equivalent primary binding sites (S1) and 2 at the equivalent secondary sites (S2).

Structures that are imported from the protein data bank often require adjustments due to crystallographic experimental procedures and difficulties. These difficulties may result in inclusion of compounds used to obtain and analyze the solid proteins, missing hydrogen atoms, and undetermined amino acid side chains. Therefore, protein structures must be prepared and optimized (15).

Protein-protein docking is carried out with the PIPER program in Schrödinger’s BioLuminate software. Simply stated, a “ligand” protein binds to a “receptor” protein. This is achieved by an algorithm that places the ligand in multiple different orientations about the receptor. Each possible complex is evaluated by a repeated algorithm

\[
E = w_1 E_{\text{rep}} + w_2 E_{\text{attr}} + w_3 E_{\text{elec}} + w_4 E_{\text{DARS}}
\]

in which \(E_{\text{rep}}\) and \(E_{\text{attr}}\) determine the repulsions and attraction energies due to van der Waals interactions, \(E_{\text{elec}}\) represents the electrostatic attraction and repulsion energies, and \(w\) represents the weighting factors. The DARS (Decoys As the Reference State) potentials \(E_{\text{DARS}}\) are obtained using large sets of known complexed structures as “decoys.” Atom pair interactions in the complexes being analyzed can then be compared to distributions in the reference state decoys (16). In brief, Boltzmann statistics are used to provide potential energies and have been found to be ideal for finding near-native conformations among structures obtained by docking.

The receptor protein is treated as a rigid structure. The user can specify which ligand or receptor chains and/or residues to include or exclude in the possible complexes to be obtained. For the flexible ligand, the user also defines the number of possible ligand rotations, the default being 70,000. Using a greedy algorithm, the best 1000 docked conformations are grouped in “clusters” according to the RMS distances between analogous atoms in the structures. Thus, a “pose” is the centroid structure within a set of equivalent complexes of a cluster. The poses are ranked by the sizes of the clusters that they embody. The default number of poses returned is 30. The user can select a larger number of poses to be returned, but it is known that the 30 poses with the highest numbers of clusters include the near native structure for 93% of protein-protein docking experiments (17, 18).

In the current work, the first protein-protein docking was for insulin (3I3Z) to the insulin receptor (6PXV). This docking of insulin, after deletion of insulins from the original structure and optimization of the resulting protein, was to test the reliability of the computational approach. The 6PXV structure was then used to evaluate the dockings of VILPs to the IR.

### Materials and Methods

Schrödinger BioLuminate 3.6 was used to determine the protein-protein interactions between the VILPs and the IR. The insulin structure and the IR were obtained from Protein Data Bank files (3I3Z and 6PXV, respectively). Ligands (including insulin) were deleted from the protein structures. Resulting structures were adjusted utilizing the Protein Prep Wizard with the selections: Remove original hydrogens and Add hydrogens; Create zero-order bonds to metals and disulfides; Convert selenomethionines to methionines; Fill in missing side chains using Prime; Delete waters beyond 5.00 Å; Include aromatic hydrogens as hydrogen donors and halogens as acceptors in weak hydrogen bonding; Delete waters more than 5 Å from het groups; Generate het states using Epik with a pH of 7.0 +/- 1.0; Hydrogen bond assignments; Sample water orientations, and Use PROPK, pH = 7.0; Remove water molecules beyond hets = 3.0 Å, and with fewer than = 3 H-bonds to non-waters; Restrained minimization, converge heavy atoms to an RMSD of 0.30 Å; Restrained minimization, hydrogen only. Structures of chemically synthesized single-chain peptides were imported from reported research (14).

Two sets of protein-protein docking experiments were conducted. In one set, both monomers of the IR were selected as the receptor; residues 865-880, which border a cavity of the FnIII-2β and FnIII-3 monomers that is very close to the cellular surface, were defined as a “repulsive region” to prevent possible dockings within this cavity; in the second set, a single monomer of the insulin receptor (IR) was selected as the “receptor protein.”

### Results and Discussion

In the interpretation of the docking results, greater emphasis has been placed on docking to the dimeric IR. It is logical to assume that virtual binding to the dimeric IR is more reliable than to the

<table>
<thead>
<tr>
<th>Ligand Protein</th>
<th>Docking to IR Dimer</th>
<th>Docking to IR Monomer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pose</td>
<td>Cluster</td>
</tr>
<tr>
<td>Insulin</td>
<td>1-5, 9, 10, 12, 15, 17-18, 22, 25, 30</td>
<td>66</td>
</tr>
<tr>
<td>SGIV</td>
<td>1-3, 5, 9-12, 25, 28</td>
<td>47</td>
</tr>
<tr>
<td>LCDV-Sa</td>
<td>1-2, 8, 12, 14, 16-17, 21-25, 27, 30</td>
<td>44</td>
</tr>
<tr>
<td>LCDV-1</td>
<td>1, 2, 14, 23</td>
<td>30</td>
</tr>
<tr>
<td>GIV</td>
<td>6-7, 14, 17-18, 24, 28-29</td>
<td>18</td>
</tr>
</tbody>
</table>

1. Complexation at equivalent primary active sites.
2. Maximum number of clusters possible = 1000, but some dockings returned fewer than 1000.
3. Different locations.
monomer, as pairwise potential and shape complementarity will better represent the native receptor when the dimer is used.

For insulin docking to the IR, the results (Table 1) showed that 66% of insulin complexes occurred at the primary active site (S1) when both chains of the IR were used as the receptor protein, and 98% when only one chain of the IR receptor was employed. Other binding sites were found that have not been reported previously (Figure 3), at different CR locations, at FnIII-1, and at L2 (very close to the S1 site). One of these lies within the inverted V pocket between the two monomers, at FnIII-1. Based on the 6PXV structure (Figure 2) this binding site is not within S2. However, the analysis is complicated by the many structural changes that have been shown to occur on binding of insulin to the IR (12, 19-20). Docking of insulin to the IR monomer gave equivalent docking sites, located on one chain. It is worth noting that some of the alternate binding sites—those at the top of the IR structure, were very close to, and even partially overlapped, the primary binding site. (Docked VILPs that overlapped complexed insulin of 6PXV were not considered to bind at the active site unless the VILP overlapped insulin by about 50% or more, by visual inspection.) Insulin binding to the secondary binding site (Figure 1) was not observed for the monomer receptor experiments. It is not known what effects binding at the secondary site or possible binding at other sites may have on the IR.

Of the 4 VILPs tested, SGIV gave the best results in mimicking insulin docking to the IR dimer, with 47% of the complexes occurring at the S1 primary active site. Of the remaining complexes, 36% were at sites adjacent to the primary sites, and directly within the S2 secondary active site (Figure 4).

The majority of complexation (44%) of LCDV-Sa was at the IR primary active site (Table 1). The second most populated clusters (31%) were for complexation at FnIII-1 (Figure 5) where the monomers form an inverted V gap that appears to involve secondary active site residues, as discussed above. A much more favorable value of 84% was obtained for docking to the primary site when the IR monomer structure only was used as the receptor (Table 1).

LCDV-1 showed 31% of complexes at the primary binding sites for the IR dimer, and 53% for the monomer. For the IR dimer, most of the alternate binding occurred at the inverted V-site. Although many of these complexes appear to occupy S2, 3 dimensional visualizations reveal that, at best, some may overlap residues of S2 as defined by the 6PXV structure (Figure 6).

GIV demonstrated the least propensity for complexation at the primary active site, with 18% complexation at that site for the dimer as the receptor. All the alternate binding sites were adjacent to the primary and directly within the secondary binding sites (Figure 7). The preference level for docking to the S1 primary site of the IR monomer (as the receptor) was 15%.

Conclusions

This work indicates that VILPs stimulate the IR by binding to the same primary active sites as insulin. Judging by their similarity to insulin in complexation at the primary activate site of the IR (Table 1), the ranking of the VILPs as promising insulin mimics is:
SGIV > LCDV-Sa > LCDV-1 > GIV. This ranking ignores possible effects of binding at the IR secondary active site (S2) as well as other sites. This is somewhat implied by research showing that the competition of VILPs with insulin for the IR active sites follows a different ranking (14). The effects of binding to S2 or alternate sites bears further research and investigation.

Due to similarities between the IR and insulin growth factor (IGF) receptor, docking studies might be expanded to include the IGF receptor, which relates to its role in anabolic effects in adults (14, 21). Reported research has already shown that the VILPs in the current work have shown activation of the IGF receptor. It would be of significant value if IC_{50} properties were available, rather than results assessing the competition of VILPs with insulin or insulin growth factors for the active sites of the IR or IGF receptor.

This computational docking research indicates that the known insulin-like effects of VILPs relate to their propensities to bind to the IR at sites analogous to the binding sites of insulin. Therefore, these compounds could eventually provide a more efficient and cost-effective treatment for type 1 diabetes. The research promotes advancements in the treatment and management of diabetes, either as new stand-alone therapeutic compounds or as constituents of a combined regimen. Although this research has focused on the glucose uptake effects of the IR receptor functions, the receptor has other important roles and dysfunction of the IR receptor is linked to several human diseases including Alzheimer’s disease and cancer (22). Thus, the discovery of novel IR activators is intimately related to many aspects of human health.
Acknowledgement

The authors gratefully acknowledge support of this work by the Richard Lounsbery Foundation, in the form of a research professorship to Massimo D. Bezoari.

References


The authors gratefully acknowledge support of this work by the Richard Lounsbery Foundation, in the form of a research professorship to Massimo D. Bezoari.