A Review on Bio Informatics for Diabetic Mellitus

Uma Makheswari M and Sudarsanam D*

Department of advanced Zoology & Bio Technology, Medical Lab Technology & Bio-medical Instrumentation Science, Loyola College – 600 034, India.

*dsloy2003@gmail.com

Abstract

Diabetes is a metabolic disorder that occurs when the pancreas does not produce enough insulin or when the body cannot effectively use the insulin it produces. It is classified in to two basic forms Type I and Type II diabetes. Computer-assisted drug design approach has contributed to the successful discovery of several novel anti diabetic agents. Molecular Docking continues to be a great promise in the field of computer based drug design. Several simulation models have been proposed to study the physiology and pathophysiology of diabetes. Biological databases and Atlas plays an important role in getting up-to-date global report on diabetes. Like so many other areas of medicine, Bio informatics has had a profound impact on diabetes research.

Keywords: Molecular Docking, Diabetes Mellitus, Simulation models, Biological databases, Atlas.

Introduction

Bioinformatics is the application of computer science and information technology to the field of biology and medicine. The ultimate goal of the field is to enable the discovery of new biological insights. It deals with algorithms, databases, artificial intelligence and soft computing, information and computation theory, data mining, image processing, modeling and simulation etc. Biological databases are an important tool in assisting scientists to understand biological phenomena from the structure of biomolecules and their interaction. This knowledge also helps to fight against diseases. Modeling and Simulation helps in developing a level of understanding of the interaction of the parts of a system, and of the system as a whole. Thus Computer modeling and simulation techniques playing an increasingly central role in medicinal field. Major research efforts in Bioinformatics include sequence alignment, gene finding, drug design, drugs discovery, protein structure alignment etc.

Diabetes Mellitus

Diabetes is a metabolic disorder that occurs when the pancreas does not produce enough insulin or when the body cannot effectively use the insulin it produces. Classic signs and symptoms of diabetes include polyuria, polydipsia, polyphagia, weight loss, headache, palpitations, and blurred vision. It is classified in to two basic forms Type I and Type II diabetes. The main risk factors for diabetes are age, family history, obesity, physical inactivity and sedentary living, insulin resistance, stress etc. Complications of Diabetes mellitus may develop after many years (10-20). The major long-term complications include damage to blood vessels. Nearly 45% of all diabetics have peripheral vascular disease. Macrovascular complications involve damage to the large blood vessels of the brain, heart, and extremities. Microvascular complications of diabetes include retinopathy and nephropathy.

Type I Diabetes Mellitus

Type 1 diabetes mellitus is caused by immunological destruction of pancreatic â cells leading to insulin deficiency (Notkins, 2002). The classical symptoms are polyuria (frequent urination), polydipsia (increased thirst), polyphagia (increased hunger), and weight loss. Complications of poorly-managed type 1 diabetes mellitus may include cardiovascular disease, diabetic neuropathy and diabetic retinopathy. Type 1 diabetes can occur at any age. However, it is most often diagnosed in children or young adults. Type 1 can be distinguished from type 2 diabetes via a C-peptide assay, which measures endogenous insulin production. Rate of β-cell destruction is quite variable in Type I Diabetes, being rapid in some individuals (mainly infants and children) and slow in others (mainly adults).

Type 1 is treated with insulin replacement therapy either via subcutaneous injection or insulin pump. Pancreatic transplants have also been used. Chatenoud et al., in 1994 have found that treatment of mice with a modified monoclonal antibody against CD3 prevents or reverses diabetes in nonobese diabetic mice and other
mouse models of type 1 diabetes mellitus. Nearly 20 different proteins have been identified as targets for T cells in the NOD mouse, and 12 of these are autoantigens in humans (Babad et al., 2010).

**Type II Diabetes Mellitus**

Type 2 diabetes is characterized with insulin resistance. Type 2 diabetes is common in individual over 40 years of age. It is characterized by hyperglycemia and associated with microvascular (ie, retinal, renal, possibly neuropathic), macrovascular (ie, coronary, peripheral vascular), and neuropathic (ie, autonomic, peripheral) complications. Type 2 diabetes can be managed by exercise and dietary modification.

Large clinical trials have emphasized that blood pressure control can help in reducing the risk of blindness in patients with diabetic retinopathy (Sivaprasad and Jackson, 2007). Bioinformatic analysis of diabetic neuropathy using functional protein sequence was reported. Sulfonylurea drugs such as glibenclamide have long been used clinically for type 2 diabetes (Malaisse, 2003). More recently, nonsulfonylurea drugs, including glinides such as nateglinide and mitiglinide, have become available for clinical use (Perfetti et al., 1998). A number of recent Internet-based interventions have been reported on for use with patients with diabetes (Bellazzi et al., 2002).

**Drug Discovery Methods**

Drug discovery is an expensive process due to the high research costs and extensive clinical testing. Two major phases are involved in creating a new drug. They are Discovery phase and clinical Testing Phase. Then the step of obtaining regulatory approval to market the drug was performed.

Before computational drug discovery was introduced, drugs were discovered by chance in a trial and error manner. High-Throughput Screening method help to quickly conduct millions of chemical, genetic or pharmacological tests in order to identify active compounds, antibodies or genes which modulate a particular biomolecular pathway. But the use of HTS is very expensive. The HTS field continues to be dynamic and very competitive. The methods of HTS are applied to the screening of combinatorial chemistry, genomics, protein, and peptide libraries. A Novel High-Throughput Screening Assay for Putative Antidiabetic Agents through PPARα Interactions was reported (Hostetler et al., 2008).

**Animal Studies**

Alloxan is a toxic glucose analogue, which selectively destroys insulin-producing cells in the pancreas (that is beta cells) when administered to rodents and many other animal species. Alloxan (2,4,5,6-tetraoxypyrimidine; 2,4,5,6-pyrimidinetetrone) is an oxygenated pyrimidine derivative. It causes insulin-dependent diabetes mellitus (called "Alloxan Diabetes"). Hence alloxan is used to induce diabetes in laboratory animals. Alloxan can be prepared by oxidation of uric acid by nitric acid. Alloxan monohydrate is used to induce Type II Diabetes Mellitus in experimental animals. Alloxan monohydrate is prepared by oxidation of barbituric acid by chromium trioxide. Alloxan monohydrate 150mg/kg body weight was dissolved in normal saline and injected intraperitoneally after 18 hours fasting to induce hyperglycemia in experimental rats (Yanarday and Colak, 1998).

Streptozotocin (STZ) is the most commonly used drug for induction of diabetes in rats (Balamurugan et al., 2003). This drugs exert their diabetogenic action when they are administered parenterally (intravenously, intraperitoneally or subcutaneously). The dose of these agents required for inducing diabetes depends on the animal species, route of administration and nutritional status (Federiuk et al., 2004).

**Computational Methods**

Computational methods can be used to predict or simulate how a particular compound interacts with a given protein target. Virtual Screening is a general term for computational methods that use computers for quick search of large libraries of chemical structures in order to identify those structures which are most likely to bind to a drug target. Virtual screening is defined in terms of finding new scaffolds rather than many of these hits.

There are two broad categories of screening techniques: ligand-based and structure-based (McInnes, 2007). In Ligand based method, Ligand is compared with pharmacophore model to determine that the ligand will bind with it (Sun, 2008). Structure-based virtual screening involves docking of candidate ligands into a protein target followed by applying a scoring function to estimate the likelihood that the ligand will bind to the
protein with high affinity (Cavasotto and Orry, 2007). The main advantages of computational methods compared to laboratory experiments are i) Low costs ii) Huge chemical search space iii) Can investigate compounds that is not yet synthesized.

Docking Studies

Molecular docking is used to predict how a drug candidate binds to a protein target without performing a laboratory experiment. Docking is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex (Lengauer and Rarey, 1996). Docking plays an important role in the rational design of drugs (Kitchen et al., 2004). There are two approaches within the molecular docking method. One approach uses a matching technique that describes the protein and the ligand as complementary surfaces (Morris et al., 1998) The second approach simulates the actual docking process in which the ligand-protein pair wise interaction energies are calculated (Feig et al., 2004).

Molecular docking software consists of two core components i) Search algorithm ii) Score function.

Search algorithm (Optimization algorithm): It is used to find the best conformations of the ligand and protein system.

Score function: This is a function providing a measure of how strongly a given ligand will interact with a particular protein.

Hex is a molecular graphics program for calculating and displaying feasible docking modes of pairs of protein and DNA molecules. Hex will run on most Windows-XP, Linux and Mac OS X PCs. it is one of the few docking programs which has built-in graphics to view the results (Ritchie, 2003). Types of docking program used for determining protein-ligand interaction were listed in Table I.

<table>
<thead>
<tr>
<th>Program</th>
<th>Algorithm</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D-Dock</td>
<td>Global: FFT; rescoring: residue potentials; refinement: mean-field sidechain multicopy</td>
<td><a href="http://www.bmm.icnet.uk/docking/">www.bmm.icnet.uk/docking/</a></td>
</tr>
<tr>
<td>HEX</td>
<td>Global: Fourier correlation of spherical harmonics</td>
<td><a href="http://www.biochem.abdn.ac.uk/hex/">www.biochem.abdn.ac.uk/hex/</a></td>
</tr>
<tr>
<td>GRAMM</td>
<td>Global: FFT clustering and rescoring decoys also available</td>
<td>reco3.ams.sunysb.edu/gramm/</td>
</tr>
<tr>
<td>PPD</td>
<td>Global: geometric hashing; rescoring: multiple filters</td>
<td>ftp://flash62.bioc.columbia.edu/pub/other</td>
</tr>
<tr>
<td>DOT</td>
<td>Global: FFT for shape complementarity and approximate Poisson–Boltzmann electrostatics</td>
<td><a href="http://www.sdsc.edu/CCMS/DOT">www.sdsc.edu/CCMS/DOT</a></td>
</tr>
<tr>
<td>BIGGER (Chemera)</td>
<td>Global: bit mapping; rescoring: multiple filters</td>
<td><a href="http://www.dq.fct.unl.pt/bioin/chemera/">www.dq.fct.unl.pt/bioin/chemera/</a></td>
</tr>
<tr>
<td>DOCK</td>
<td>Global: grid-based energy function; flexible docking: random search plus incremental construction</td>
<td><a href="http://www.cmpharm.ucsf.edu/kuntz/dock.html">www.cmpharm.ucsf.edu/kuntz/dock.html</a></td>
</tr>
<tr>
<td>AutoDock</td>
<td>Grid-based empirical potential flexible docking via Monte Carlo search and incremental construction</td>
<td><a href="http://www.scripps.edu/pub/olson-web/download.html">www.scripps.edu/pub/olson-web/download.html</a></td>
</tr>
</tbody>
</table>

(Source: Smith and Sternberg, 2002).
Many researchers are trying to find out correct target for the treatment of diabetes. One of the ways is by molecular docking studies. The targets which was used by many Scientist were Glycogen phosphorylase, Protein Tyrosine Phosphatase 1-Beta(PTP-1B), Dipeptidyl peptidase IV (DPP IV), Glucokinase, Peroxisome Proliferator-activated Receptor (PPAR)-γ etc. Protein – Ligand docking studies has been done by many researchers for structural based drug designing for Diabetes mellitus (Guttula et al., 2011). Docking studies of green tea flavonoids were carried out using Auto Dock 4.0 and Argus lab 4.0.1. Analysis of the results shows that epicatechin can act as a potent insulin receptor activator (Ganugapati et al., 2011).

Foot ulcers are a very common complication of type I and type II diabetes. Individuals with diabetes have at least a 10-fold greater risk of being hospitalised for soft tissue and bone infections of the foot than the individuals without diabetes (Boykoe et al., 1995). A clinico-bioinformative study was conducted for infected foot ulcers in male and female diabetic patients. From the study it was concluded that Male diabetic patients with MDRGNB-infected foot ulcers have poor glycemic control and hence they might have higher mortality rates compared to their female counterparts. Further Amino acid residues Asn132, Glu166, Pro167, Val172, Lys234 and Thr235 of CTX-M-15 (enzyme) make important contacts with cefotaxime (drug) in the 'enzyme-drug complex', researchers are expected to duly utilize this information for designing more potent and versatile CTX-M-inhibitors (Shakil and Khan, 2010).

Simulation models

The ability to simulate the glucose–insulin system in normal life conditions can be very useful in diabetes research. Several simulation models have been proposed and proven to be useful in tackling various aspects of normal physiology and pathophysiology of diabetes. The mostly used and also the simplest model is the minimal model of Bergman (Bergman et al., 1979) for type 1 diabetes patients under intensive care. The minimal model is based on an Intravenous Glucose Tolerance Test (IVGTT), where glucose and insulin concentrations in plasma are sampled after an intravenous glucose injection. Recently, a new meal simulation model has been proposed (Dalla et al., 2007). Scope of this is to develop a new simulation model of the glucose-insulin system in the normal human capable of describing the physiological events which occur during a standard mixed meal.

Sequence Alignment

Sequence alignment is a way of arranging the primary sequences of DNA, RNA, or protein to identify regions of similarity that may be a consequence of functional, structural, or evolutionary relationships between the sequences. Pairwise sequence alignment and multiple sequence alignment techniques are often used by researchers in Bio informatics. Using bioinformatics techniques and sequence analyses algorithms, a comparative study between human and rodents revealed similarity in the behavior of genes involved in the control of energy homeostasis. The Study it was inferred that brain-derived neurotrophic factor (BDNF) controls the actions of several proteins including insulin, leptin, and ghrelin and is significant to the pathobiology of type2 diabetes and obesity (Rao, Sridhar et al., 2008).

Basic Local Alignment Search Tool is the most widely used bioinformatics tool. It is optimized for speed but sacrifices only minimal sensitivity in searching databases. Relationship between altered calcium homeostasis and diabetic cardiomyopathy was found by constructing a Phylogenetic tree using Neighbour-Joining Algorithm in bioinformatics approach (Rao, Thota, et al., 2008). Role of several genes/proteins that are believed to be involved in the evolution of obesity associated diabetes by employing multiple sequence alignment using ClustalW tool was studied (by constructing a phylogram tree using functional protein sequences extracted from NCBI) (Rao, Hanuman, et al., 2008). Genomics has become a major source of drug targets, and knowledge on bioinformatics is crucial for finding and validating novel targets so as to minimize investment in laboratorial resources.

Databases for Diabetes

Biological databases are an important tool to explain a host of biological phenomena from the structure of biomolecules and their interaction, to the whole metabolism of organisms and to understanding the evolution of species. This knowledge helps facilitate the fight against diseases, and in the development of drug.

DrugBank is a freely available, fully downloadable drug interaction and drug metabolism database. It contains detailed information about 250 of the most frequently prescribed FDA-approved drugs (chemical structure, common and chemical names, 3D structure coordinates, drug/chemical class, solubility,
pharmacology, etc.), along with detailed information about their known protein, expected or measured toxicity and known metabolizing enzymes. A user-friend database of type 2 diabetes genetic association of manually curated information was constructed (Lim et al., 2010). This database can be used for research purposes, such as an association and functional study of type 2 diabetes related genes, and to construct a diabetes risk test in the preparation of personalized medicine in the future.

A database named Phyto-Mellitus with information on plants traditionally used for diabetes with their chemical constituents was present (Middha et al., 2009). The database can be accessed alphabetically using genus name for information on specific plants. From the data, 36% were whole plants, 30% leaves, 13% seeds, 11% roots, and 10% fruits. MEDLINE, the world's largest database of medical abstracts, including thousands of journal articles dealing with diabetes. MEDLINE is a service of the U.S. Government's National Library of Medicine (NLM) in Bethesda, Maryland.

T2D-Db is a database of all molecular factors reported to be involved in the pathogenesis of Type 2 diabetes in human, mouse and rat. It provides information on candidate genes, gene description, genomic loci, alleles, gene and protein sequences and the corresponding literature, gene ontology, homologous/orthologous genes, microarray expression and tissue specific expression analyses, protein-protein interaction information, SNP markers etc. It also cats information on genes candidates for the risk factors/complications reported to be associated with Type 2 diabetes (Agrawal et al., 2008).

**Atlas on Diabetes**

The Diabetes Atlas will be used by the World Diabetes Foundation, the International Diabetes Federation and the World Health Organization to communicate the global impact of diabetes and underline the need for immediate intervention from governments, healthcare professionals, international health organisations and other bodies.

According to the World Health Organisation estimates, India had 32 million diabetic subjects in the year 2000 and this number would increase to 80 million by the year 2030 (Wild et al., 2004). The first Edition of Diabetes Atlas was published in 2000. The 5th edition of the IDF Diabetes Atlas was officially released on World Diabetes Day 14 November 2011. The Atlas is the most up-to-date global report on diabetes, covering all aspects of the disease from epidemiology to health economics and education.

**Advances in Information Technology**

Information Technology now plays an increasingly important role in diabetes healthcare and is continuing to revolutionize day-to-day management for millions of people.

Telemedicine is the use of telecommunication and information technologies in order to provide clinical health care at a distance. The Informatics for Diabetes and Education Telemedicine (IDEATel) Project has been performed by Steven Shea in 2007. It was found that diabetes case management delivered using telemedicine improved hemoglobin, blood pressure and cholesterol levels in older patients with diabetes mellitus at one year of follow-up, compared to usual care. The IDEATel study provides evidence that medical informatics and telemedicine technology can help to translate advances in treatment of chronic diseases into effective health care.

A system that provides real-time individualized medical treatments that are easily accessible using Internet and wireless technology is called as ubiquitous healthcare (u-healthcare) system. The u-healthcare system can potentially provide disease prevention and early treatment, as well as continuous follow-up that are available whenever and wherever they are needed. Appropriate self-care, is essential in diabetes care but is difficult to monitor. Hence a new health care delivery model with the u-healthcare system has been introduced to induce effective glucose control. A glucometer with a mobile system and Zigbee communication protocol, which is a specification for a suite of high-level communication protocols using small, low-power digital radios for wireless home area networks, showed that diabetic patients could be more autonomous in controlling their glucose levels (Lee et al., 2009). To improve quality and efficiency of care for elderly patients with type II diabetes, Clinical Decision Support System (CDSS)-based ubiquitous healthcare (u-healthcare) service, which is an individualized health management system using advanced medical information technology have been introduced by many researchers.

AIDA is a freeware computer program that allows the interactive simulation of plasma insulin and blood glucose (BG) profiles for demonstration, self-learning and research purposes (Lehmann, 1999). A web-based version of the program, called AIDA online, has been available on the World Wide Web since 1997/1998.
for diabetes education. AIDA online includes a standard Web-browser interface. Input parameters include carbohydrate ingestion and insulin dosing. The output of the program is the simulated BG with supplemental information regarding the simulated plasma insulin values.

Computer-Prompted Diabetes Care (CPDC) is a multifaceted software program designed to allow the physicians of the Maccabi Health Care Services, for diabetes care through the use of real-time, on-screen, situation-specific messages (Spero et al., 1998). The CPDC software was developed relatively inexpensively by Maccabi Health Care Service. Novel approach for diagnosing diabetes illness using pervasive healthcare computing and artificial neural networks on small mobile and wireless devices have been reported on 2010 (Bayraktar et al., 2010). All over the world scientists are using nanotechnology to create new treatments for diabetes. The future possibilities on the latest technological advancement in treating Diabetes is too firm to ignore.

Conclusion

Rapid urbanization and industrialization have resulted in dramatic lifestyle changes leading to lifestyle related diseases. The design of new drug is based on the requirement of active binding site present in a protein. Bio-informatics has been applied effectively in the management of chronic diseases such as Diabetes. Databases are key to both bioinformatics and Drug Discovery.

Computer-assisted drug design approach has contributed to the successful discovery of several antidiabetic agents. There are various tools, which can be used for Computer aided drug design such as QSAR, Docking, Homology modeling etc. CADD methods are heavily dependent on bioinformatics tools, applications and databases. As such, there is considerable overlap in CADD research and Bioinformatics. Various computer aided drug design and medicinal chemistry tools were used by many researchers to design novel potential drug candidates for diabetes (Semighini et al., 2011). Bioinformatics plays a central role in diabetes research.

References


