LIGAND BASED PHARMACOPHORE MODELING AND VIRTUAL SCREENING STUDIES FOR DEVELOPMENT OF NOVEL GLUCAGON INHIBITORS BY USING FLAVONOIDS

A
Thesis
Submitted to the Orissa University of Agriculture & Technology, Bhubaneswar for the partial fulfillment of the requirement for the award of degree of

MASTER OF SCIENCE IN BIOINFORMATICS

BY
SUBHASHREE SUBHASMITA RAUT
ADM NO:15BI/15

DEPARTMENT OF BIOINFORMATICS
CENTRE FOR POST GRADUATE STUDIES
ORISSA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY
BHUBANESWAR-751003

Name of the Advisor **Mr. Surya Narayan Rath**

DEDICATED

TO

MY beloved Mother

सी.एस.आई.आर–भारतीय रासायनिक जीवविज्ञान संस्थान
वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद की एक इकाई
विज्ञान एवं प्रौद्योगिकी मंत्रालय के अधीन, एक स्वायत्त निकाय, भारत सरकार
4, राजा एस. सी. मल्लिक रोड, यादवपुर, कोलकाता - 700 032

CSIR - INDIAN INSTITUTE OF CHEMICAL BIOLOGY

A Unit of Council of Scientific & Industrial Research
An Autonomous Body, under Ministry of Science & Technology, Government of India
4, Raja S. C. Mullick Road, Jadavpur, Kolkata-700 032

BONAFIDE CERTIFICATE

This is to certify that the thesis entitled "LIGAND BASED PHARMACOPHORE MODELING AND VIRTUAL SCREENING STUDIES FOR DEVELOPMENT OF NOVEL GLUCAGON INHIBITORS BY USING FLAVONOIDS" is the bonafied work of SUBHASHREE SUBHASMITA RAUT from Orissa University of Agriculture & Technology, Bhubaneswar who carried out the work under my supervision at CSIR-IICB. It is further certified that to the best of my knowledge the work reported here in does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

Dr. ARINDAM TALUKDAR
Senior Scientist, (Assistant professor, AcSIR)
Organic and Medicinal chemistry division
CSIR-Indian Institute of Chemical Biology
(Council of Scientific and Industrial Research)

डॉ.अरिन्दम तालुकदार /Dr. Arindam Talukdar
वरिष्ठ वैज्ञानिक/Senior Scientist
सी.एस.आई.आर.-भारतीय रासायनिक जीवविज्ञान संस्थान
CSIR-Indian Institute of Chemical Biology
भारत सरकार/Government of India
४ राजा एस. सी मल्लिक रोड/4 Raja S.C. Mullick Rd.
यादवपुर, कोलकाता/Jadavpur, Kolkata-700 032

CERTIFICATE –II

This is to certify that the dissertation entitled **"Ligand based pharmacophore modeling and virtual screening studies for development of novel Glucagon inhibitors by using flavonoids"** submitted by **SUBHASHREE SUBHASMITA RAUT,** to the Orissa University Of Agriculture & Technology, Bhubaneswar in the partial fulfillment of the requirements for the award of the degree of **Master of Science** in **Bioinformatics** has been approved by the students advisory committee after an oral examination of the same in collaboration with external examiner.

ADVISORY COMMITTEE

1. Mr.Surya Narayan Rath — Member ……..
   Assoc. Professor
   Department of Bioinformatics

2. Mr. Sukanta kumar Pradhan
   HOD, Bioinformatics — Member ……..

3. Mr. Abhimanyu Dash
   HOD, CSA — Member ……..

External Examiner ……

ACKNOWLEDGEMENT

I express my profound gratitude to Dr. Arindam Talukdar (Senior Scientist), Organic and Medicinal chemistry, CSIR-Indian Institute of Chemical Biology, Kolkata me to carry out the dissertation work under his guidance.

I convey my deepest regards and heartiest thanks to Mr. Sourav Pal (project assistant) in CSIR-IICB Kolkata, Mr.Vinay Kumar M.S (pharm) in NIPER, CSIR-IICB, Kolkata for their guidance and support I benefited lot from their experience in the computer field of drug discovery during my project work.

It is my privilege to share my deep sense of gratitude to my advisor Mr. Surya Narayan Rath, Asst. Professor, Department of Bioinformatics, Orissa University of Agriculture and Technology, for his constant guidance, support and help without which the project would not have been successfully completed.

. I am thankful to the member of the advisory committee Mr Sukanta kumar Pradhan, HOD, Dept of Bioinformatics, OUAT for his support and encouragement to carry out this work successfully.

I convey my hearty thanks to all my faculty members Miss. Sushma Rani Martha, Miss. Sucharita Balabantaray and Mr. Sujit Kumar Dash for their suggestion and supervision that has made all difficulties very simple during my course of two years.

Also I would like to express my heartiest and cordial regards to my beloved Bapa, Bou who are real inspiration for me in every step of my life giving me unbound emotional support. I feel great pleasure to express my love to all my sweetest friends for their strong mental support throughout the project period and in these two years.

I feel honored to be a part of this auspicious university for providing me a healthy atmosphere in these two years .Last but not the least I express my gratitude to god for invaluable inspiration for accomplishment of such a splendid work.

Subhashree subhasmita Raut

Name of the Student	:	**Subhashree subhasmita Raut**
Admission No	:	**15BI/15**
Title of thesis	:	**Ligand based pharmacophore modeling and virtual screening studies for development of novel Glucagon inhibitors by using flavonoids**
Degree for which thesis submitted	:	**Master of Science in Bioinformatics**
Name of the Dept, & University	:	**Department of Bioinformatics, Centre for Post Graduate Studies, Orissa University of Agriculture &Technology, Bhubaneswar, Odisha, 751003**
Year of submission	:	**2017**
Name of the advisor	:	**Mr.Surya Narayan Rath**

ABSTRACT

The chronic hyperglycemia is a severe condition associated with Type-2 diabetics patient when blood glucose level is very high .Glucagon is a peptide hormone produced by alpha cell of pancreas which works to raise the concentration of glucose in the blood stream. Glucagon effect just opposite to insulin which decreases extra cellular glucose. Therefore inhibiting the effect of Glucagon receptor is essential to treat hyper glycemia.In the concentration flavonoids are certain secondary metabolites plants of rue family have been reported is good inhibitors against glucagon receptors. In this study has been made to identify potential inhibitors using pharmacophore model of flavonoids compound.3D QSAR pharmacophore generation virtual screening and molecular docking studies were conducted using Discovery Studio 4.1 software. The best fit pharmacophore model was proposed with high binding affinity best on good pharmacophore score. This compound may further be procedure for other experimental validation .This work would help to design a potent lead molecule against hyper glycemia.

Keywords: Chronic hyperglycemia, Type-2 diabetes, glucagon, flavonoids, 3D QSAR, docking.

CONTENTS

List of Figures

ABBREVIATIONS

CMT	California Mastitis Test
DNA	Deoxy Ribonucleic Acid.
PCR	Polymerase Chain Reaction
EMB	Eosin Methylene Blue
BLAST	Basic Local Alignment Search Tool.
NCBI	National Centre For Biotechnology Information.
MEGA	Molecular Evolutionary Genetics Analysis.
SDS	Sodium Dodecyl Sulphate
FASTA	Fast Alignment
MSA	Multiple Sequence Allignment
WMT	Wisconsin Mastitis Test

1. INTRODUCTION

Mastitis refers to an inflammatory reaction of udder tissue to bacterial, chemical, thermal or mechanical injury and is caused by microbial organisms or non infectious resulting from physical injury to the gland .Mastitis among bovines is a common disease, inflicting enormous financial bet backs on farmers. Mastitis remains the most common and the most expensive disease of dairy Indus try throughout the world. Among major diseases of during animals, Mastitis ranked first and is of major concern in Pakistan. This disease not only affects the physical, chemical, bacteriological, technological and organoleptic properties of milk but also affects quality of milk.Bovine mastitis continues to be the most costly disease to the dairy farmers. It nominates in Tamilnadu as one of the most prevalent diseases in dairy cattle among the dairy farms. (1)(2)(3).

TYPES OF MASTITIS

There are several ways of classifying mastitis. A simple classification recognizes mastitis as two major groups.

A. Contagious mastitis

Caused by bacterial live on the skin of the teat and inside the udder. Contagious mastitis can be transmitted from one cow to another during milking. Contagious mastitis is sometimes referred to as cow-to-cow mastitis, because it is generally spread from cow-to-cow. The primary habitat of bacteria causing contagious mastitis is on the udder and in teat lesions. These bacteria have poor survival in the environment when not associated with the skin or in the gland. The infection is transmitted by milk-contaminated fomites at milking, by a sponge used to wash the cow's machine. The major organisms causing contagious mastitis are streep to coccus agalactiae, staphylococcus aureaus & mycoplasm Contagious mastitis can be divided into 3 groups.such as Clinical mastitis,

Subclinical mastitis,chronic mastitis. Almost all these three sub clinical mastitis is the most important are with no visible signs of the disease. somatic cell count (SCC) of the milk will be elevated.bacteriological culturing of milk will detect bacteria in the milk,cows the greates financial loss to dairy farmers through lowered milk production.for every clinical case of

mastitis there will be 15 to 40 subclinical asses.In scc the no. of leukocytes or white blood cells per milliter of milk.normal milk will have less than 2,00,000 cells per milliliter. An elevated scc is an indication of inflammation in the udder bulk tank scc gives an indication of the level of sub clinical mastitis and the loss of the milk production in a herd due to mastitis. The udder. Bulk tank SCC gives an indication of the level of sub-clinical mastitis and the loss of the milk production in a herd due to mastitis.(4)

MASTITIS CAUSING BACTERIA

Research finding have proved that buffalo is as susceptible to mastitis as cow(11). The causative organisms of mastitis in buffaloes have been.

Bacteria that are known to cause mastitis include –

- *Pseudomas aeruginose*
- *Staphylococcus aureaus*
- *Staphylococcus agalactiae*
- *Staphyulococcus uberis*
- *Brucell a melitensis*
- *Cryhlbacterium bovil*
- *My coplasm (various species)*
- *Escherichia coli*
- *Klebsiella pneumonia*
- *Klebsiella oxytoca*
- *Enterbacter acrogens*
- *Pasteurella SPP*
- *Prulperella pyogenes*
- *Proteus SPP*
- *Protogheca Zopfii*
- *Prototheca wickerhamni(5)*

PATHOGENESIS

Mastitis in dairy animals occurs when the udder becomes inflamed and bacteria invade the teat canal and mammary glands. These bacteria multiply and produce toxins that cause injury

to the milk secreting tissue, besides, physical trauma and chemical irritants. These cause increase in the no. of leukocytes or somatic cells in the milk reducing its quality and adversely affecting the quality of milk and milk by products. The teat end serves as the first line of defense against infection from outside sphincter of smooth muscles surrounds the teat canal which functions to keep the teat canal closed [6]. It also prevents milk from escaping and bacteria from entering into the teat.

From inside, the teat canal is lined with keratin derived from stratified squamous epithelium. Damage to keratin has been reported to cause increased susceptibility of teat canal to bacterial invasion and colonization (7). The Keratin is a waxy amterial composed of fatty acid and fibrous proteins in the teat. The fat acids are both esterified and nonesterified representing myristic acid, palmitoleic acid and linolivic acid which are bacteriustatic (8).

The fibrous proteins of keratin in the teat canal bind electrostatically to mastitis pathogens. Which alter the bacterial cell wall rendering it more susceptaible to osmotic pressure? In ability to maintain osmotic pressure causes lysis and death of invading pathogen (9). The Keratin st. thus enables trapping of invading bacteria and prevents their migration into the gland cistrn (10).

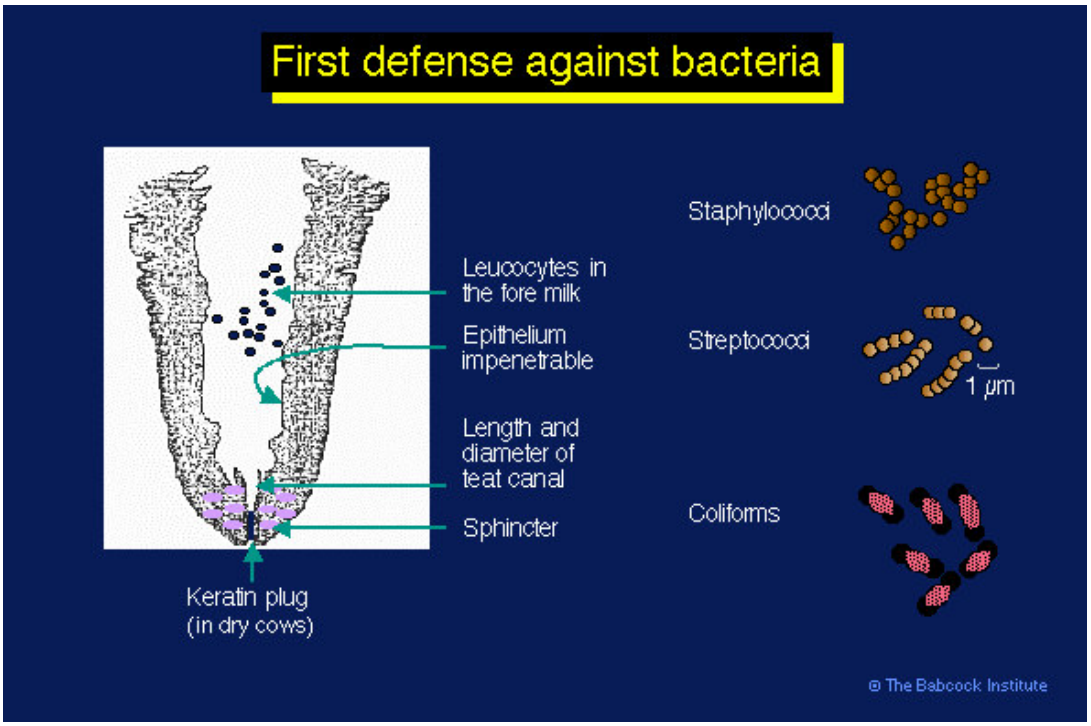


FIG-1:- First defence against bacteria

During milking bacteria present near opening of the teat find opportunity to enter the teat canal causing trauma and damage to the Keratin or mucocus membranes lining the teat sinus (11). The canal of the teat may remain partially open for 1-2 hours after milking and during this period the pathogens may freely enter into the teat canal (12).

Bacterial pathogens which are able to traverse the opening of teat end by escaping antibacterial activities establish the disease process in the mammary gland which is the second line of defense of the host. In dairy animals the mammary gland has a simple system consisting of teats and udder. Where the bacteria multiply and produce toxins, enzymes and cell-wall component which stimulate the production of inflammatory mediator's altracting phagocytes. The severity of inflammatory response however is dependent upon both the host and pathogen factors.(13)

The pathogen factors include the species, virulence, strain and the size of bacteria. Where as the host factors include parity. The stage of lactage and immune status of the animal as well as somatic cell count.(14)

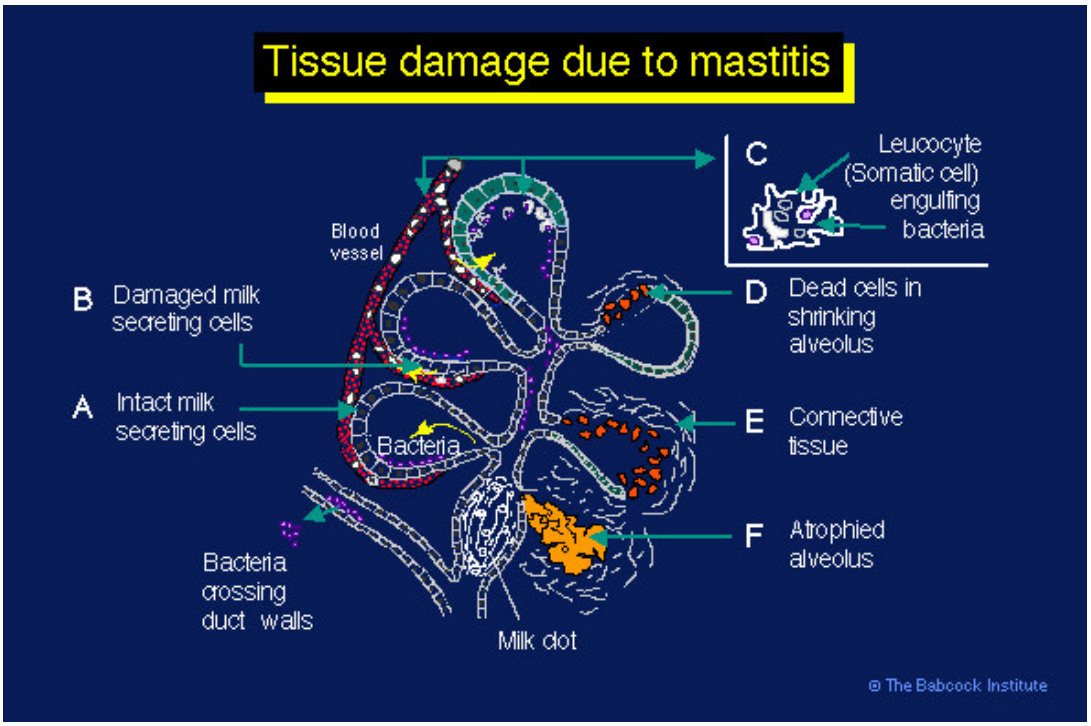


FIG-2:-Tissue damage due to mastitis

DIAGNOSIS OF MASTITIS:-

The first step in treating mastitis is to identify the causative agent. The presence of a pathogen and the inflammatory response of the udder signify an infection. The inflammatory response. Which results in abnormal milk is usually detected by the dairyman. Because mastitis is frequently subclinical "hidden" a number of tests have been developed for detecting mastitis most tests estimate the somatic cell counts (scc) of a milk sample .there is no one somatic cell count at which a cow is free from mastitis .A level of 50,000 call me of milk is usually used as a beginning point for classer observation.(15)

A variety of tests are available to determine the presence or absence of clinical and subclinical mastitis. The majority of these tests primarily indicate inflammation in the udder .They do not measure infection or bacterial presence.(16)

The stripcup test:-

The stripcup or strip plate is an important the milking parlor for determining the presence of clinical mastitis. The test is rapid and can easily be adopted as a part of milking routing few

streams of the fore milk are squirted on the strip cup and are visually scored depending on the amount of gel formed.

Table:-

Somatic cell counts for California Mastitis test score

TABLE:-1

Score	Somatic cell range
Negative	0 - 2,00,000
Trace	2,00,000 – 4,00,000
1	4,00,000 – 1,200,000
2	1,200,000 – 5,000,000
3	➢ 5,000,000

The CMT is a simple inexpensive rapid screening test and is useful inexpensive, rapid screening test and is useful in determining which quarter of the cow is most affected. However, the CMT cannot be used as a means of picking out cows for treatments as it all defects is a cellular response to inflammation. The test useful in indicated sub-clinical and clronic mastitis.(17)

FIG-3:-The Stripcup test

Wisconsin Mastitis test:-

This is a laboratory test that is generally conducted on bulk tank samples. The test uses the same reagent as the CMT. However, the reason is not estimated but measured and thus the WMT is more precise than the CMT. The WMT results are generally reported in mm. However, the scores can be used to predict the average number of somatic cells. The test is usually used as a screening test on producer's milk due to its simplicity and objectivity.
Samples score higher than 19 mm are subjected to further tests and scruting. The test also provides a convenient method of monitoring udder health on a herd basis by examining the cows, milking equipment, milking procedure and outside environment when WMT values starts to increase.

Based on the WMT, the goal for the herd should be to maintain a bulk than reading of less than 8 mm (3, 00,000 cells /ml). Milk with a WMT score of 19 mm (8, 55,000) or greater is generally considered unsuitable for human consumption.(18)

FACTORS INFLUENCING SUSCEPTIBILITY TO MASTITIS:-

Some bacteria are more virulent than others in causing mastitis.Although infection can occurs at any time, most of the new infection take place during the first three weeks of the dry period and during the 1st month after parturition, suggesting that level of milk production is not directly related to mastitis. It is likely that intra mammary pressure is a predisposing factor for mastitis during these periods.

The incidence of mastitis increases with age. Nevertheless, it is possible for the udder of the first-calf heifer to be infected at parturition. Not directly related to incidence of mastitis. However, other factors, which affect milk, yield such as milking rates; pendulous udders may be related to mastitis incidence.Length of the leg in proportion to the udder size and relative strength of the udder attachment are examples .large, pendulous udders tend to emceed the capacity of the supporting ligaments, with a con sequent of breakdown of the udder .this will subject the udder to more physical injuries and this increases the incidence of mastitis .Improper use of milking machine (irregular fluctuation of vacuum level over milking incomplete milking) is related to tissue irritations and incidence of mastitis.Mastitis often increase when cows are turned on to pastures .chilling of the udder in cold ground in the spring or fall housing as it relates to the degree of udder and teal injury.(19)(20)

Economics loss:-

Economics loss to mastitis in the United States is estimated to be approximately & 185/cow annually. It we assume the same milk price and this value is multiplied by the total no of milking cows (9.5, million head) the total annual cost of mastitis is about & 1.

8 billion .this is approximately 10% of the total value of farm milk sales, and about two thirds of this loss is due to reduced milk production in sub clinically infected cows .the average production loss per lactation for one infected quarter is about 1,600 pounds .others losses are due to also curded abnormal milk and milk withheld from cows treated with antibiotic ,costs of early replacement of affected cows ,reduced sale value of culled cows ,costs of drugs and veterinary services ,and increased labor costs.(21)(22)

Estimated annual losses due to mastitis:-

TABLE :-2

Source of loss	Loss per cow	Percent of total

Reduced of loss	121.00	66.0
Discarded Milk	10.45	5.7
Replacement	41.73	22.6
Entra Labor	1.14	.1
Treatment	7.36	4.1
Veterinary Service	2.72	1.5
Total	**184.40**	**100.0%**

These estimates do not include additional costs arising from mastitis associated problem & related to antibiotic residues in human foods, milk quality control, dairy manufacturing, and nutritional quality of milk, degrading of milk supplies due to high bacteria or somatic cell counts, and interference with genetic improvement of dairy animals. Mastitis causes May or economical losses in dairy animals and the estimated economic loss to the us was found to be more than $ 2 billion per annum (I as per et al, 1982) the uk dairy industry a loss of 93 million per year. While Australian dairy industry faced more than $ 100 million annually (however .2005) in a recent study (itujps et al, 2008), average economics losses assessed by the Dutch dairy farmers were E78/Cow per year with a range of 17-198E/ Cow per year according to mayer (1990).the world wide cost of this disease was us $ 35 billion annually .(23)

Studies Comparing isolates Within a Geographical Region:-

A few studies have determined the proportion of resistant isolates over time performed at the same laboratory and geographic region antimicrobial susceptibility of saureus over a tour year period was determined for both clinical and subclinical isolates from British dairy co genes trends over time in the proportion lines and tart strains for clinical or subclinical humans' isolates were reported. A survey of the prevalence of mastitis and antimicrobial

resistant of in tramammary pathogens was carried out in 1988 (17,111quarten milk sample) and again in 1995(10,410 quarter samples) in Finland.(24)

E. coli Mastitis –The Past, the Present and the Future:-

Escherichia coli (E.coli) is a gram-ve rad shaped bacterium that is commonly found in the lower intestine of warm –bolded organism & (endotherems) Fig E.coli review-2
The challenge of Escherichia coli udder infection will remain so long as cows produce faces .E.coli mastitis is a disease of wet environment .infection is via the treat end, in a variety of situations the outcome of a case of E.coli udder infection is deter –mined by the in flammatery response of the cow rather than the pathogenlity of the invadinean there is a case for not treating mild cases of E.coli mastitis with antibiotics of an on –farm test for E.coli .edutoxin in mastitis milk proves fast and reliable perhaps there is a place for the farm laboratory, similar to the already accepted farm workshop, E.coli udder infection towards the end of the dry period, with infected quarters presenting as clinical cases of mastitis in the following lactation, is an important aspect of the problem. Vaccination should protect against acute clinical E.coli mastitis.(25)

HISTORY

E. coli was first recorded in bovine mastitis milk in 1896 (1). Kitt reproduced the disease by "lightly sticking (SIC) E.coli to teat orifices. There are at least 35 references to E.coli bovine mastitis for 1935.

E. coli strain has demonstrated different DNA sequence isolated from various mastitis cases and other infectious.(26)

OBJECTIVE

Since the literature on the DNA sequence of *E.coli* strain isolated from bovine mastitis is scanty, hence an attempt was made with the following objectives:

1) To isolate the causative agent *E. coli* from the milk sample of infected bovine mastitic cow
2) To isolate the genomic DNA of *E. coli*
3) To amplify the 16 S r-DNA of *E. coli* using suitable primer
4) To sequence the 16 S r-RNA of isolated *E. coli* strain
5) To perform the *in silico divergence* analysis of isolated *E. coli* strain

CHAPTER-II

REVIEW OF LITERATURE

2.1 Glucagon

Glucagon is a 29 amino acid polypeptide hormone synthesised and stored in pancreatic islet α-cells, and then secreted in response to reductions of blood glucose. Glucagon half-life in plasma is only a few minutes and its dynamic range is around 2 to 40 pmol/l. In α-cells glucagon is produced from the same proglucagon precursor as GLP-1, GLP-2 and oxyntomodulin (all biologically active) and glicentin are produced from in gut L-cells and in some brain cells. In α-cells processing is catalysed by prohormone convertase-2 (PC2) as opposed to PC1/3 in the gut cells. Processing in α-cells gives glucagon and the co-released biologically inactive glicentin-related polypeptide and major proglucagon fragment.The primary stimulus for glucagon secretion is reduction in blood glucose below the normo-glycaemic range, although how much of the effects of glucose are directly on α-cells or are indirect and involve intra-islet paracrine effects and glucose-sensing cells in the CNS is a matter of debate. β-cell products (insulin, amylin, zinc ions and gamma-amino butyric acid), somatostatin from delta cells, autonomic innervations, and glucagon itself may all be involved, and plasma concentrations of amino acids and gut hormones GLP-1, GLP-2 and GIP may also play a role, so there are potentially many different positive and negative regulators of glucagon secretion.Within α-cells low glucose allows moderate KATP channel activity which sets the membrane potential in a range where voltage dependent Na+ and Ca2+ channels open and trigger action potentials and Ca2+ influx and secretion of glucagon into the portal vein in a pulsatile manner . Elevated glucose depolarises α-cells and inactivates the voltage-gated cation channels to suppress electrical activity and secretion.

2.1.1 Glucagon action and mechanism

Glucagon is secreted into the portal vein to its main site of action the liver, the tissue with highest expression of the glucagon receptor (Gcgr), a class II seven-trans-membrane G-protein coupled receptor (GPCR) coupled to G proteins of GSα (primarily) or Gq types. Gcgr shares some homology with GLP-1 and GIP receptors.

There is some Gcgr expression in other tissues including heart, kidney, intestinal smooth muscle, brain (hypothalamus and brainstem), adipose tissue and pancreatic α-cells; its action and function in these other tissues is less well described.Glucagon binding to liver Gcgr activates GSα which activates adenylyl cyclase to increase cytosolic cyclic AMP and activate

Protein Kinase A (PKA) (Figure 1) as its main mechanism; additionally Gq activation can activate phospholipase C to increase cytosolic Ca2+ .

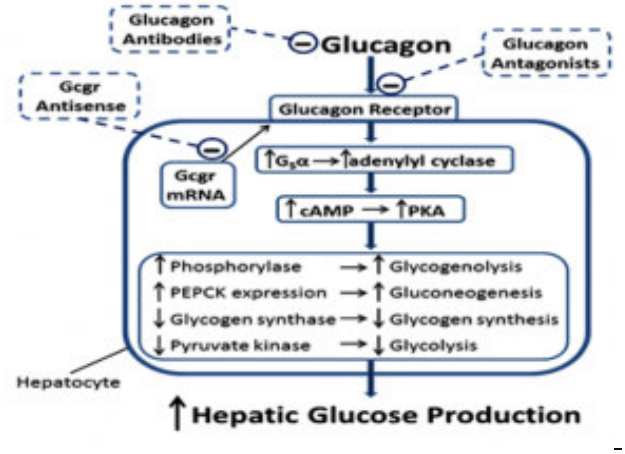


Figure 7: Schematic representation of the actions of glucagon to stimulate hepatic glucose production, and the different current approaches (dashed lines) by which these can be antagonised.

PKA activation leads to activation of glycogen phosphorylase and increased expression of phosphoenol pyruvate carboxykinase (PEPCK) which then stimulate glycogenolysis and gluconeogenesis, respectively. PKA also phosphorylates and inactivates glycogen synthase and pyruvate kinase to inhibit glucose flow to glycogen and glycolysis. These combined actions lead to increases in hepatic glucose production to ensure euglycaemia.

2.1.2 Glucagon and diabetes

In diabetes, where insulin action is defective, one of the key factors driving hyperglycaemia is that glucagon continues to act and promote hepatic glucose production even when blood glucose is high. Type 2 diabetics exhibit elevated fasting and post-prandial glucagon levels, especially when considered in terms of prevailing plasma glucose and insulin levels, and their insulin resistance means that glucagon secretion and action is not sufficiently suppressed. It has been estimated that > 50% of pathological increments in plasma glucose are due to inappropriate glucagon action. This led to the concept that antagonism of glucagon secretion or action may be a viable anti-diabetic approach.

2.1.3 GPCR (G Protein coupled receptor)

G protein–coupled receptors (GPCRs) which are also known as seven-transmembrane domain receptors, 7TM receptors, heptahelical receptors, serpentine receptor, and G protein–linked receptors (GPLR), constitute a large protein family of receptors,that detect molecules outside the cell and activate internal signal transduction pathways and ultimately modulates cellular responses. Coupling with G proteins, they are called seven-transmembrane receptors because they pass through the cell membrane seven times. G protein–coupled receptors are found only in eukaryotes, including yeast, choanoflagellates, and animals. The ligands that bind and activate these receptors include light-sensitive compounds, odors, pheromones, hormones, and neurotransmitters, and vary in size from small molecules to peptides to large proteins. G protein–coupled receptors are involved in many diseases, and are also the target of approximately 40% of all modern medicinal drugs.

There are two principal signal transduction pathways involving the G protein–coupled receptors:

- The cAMP signal pathway and
- The phosphatidylinositol signal pathway.

When a ligand binds to the GPCR it causes a conformational change in the GPCR, which allows it to act as a guanine nucleotide exchange factor (GEF). The GPCR can then activate an associated G protein by exchanging the GDP bound to the G protein for a GTP. The G protein's α subunit, together with the bound GTP, can then dissociate from the β and γ subunits to further affect intracellular signaling proteins or target functional proteins directly depending on the α subunit type ($G_{\alpha s}$, $G_{\alpha i/o}$, $G_{\alpha q/11}$, $G_{\alpha 12/13}$)

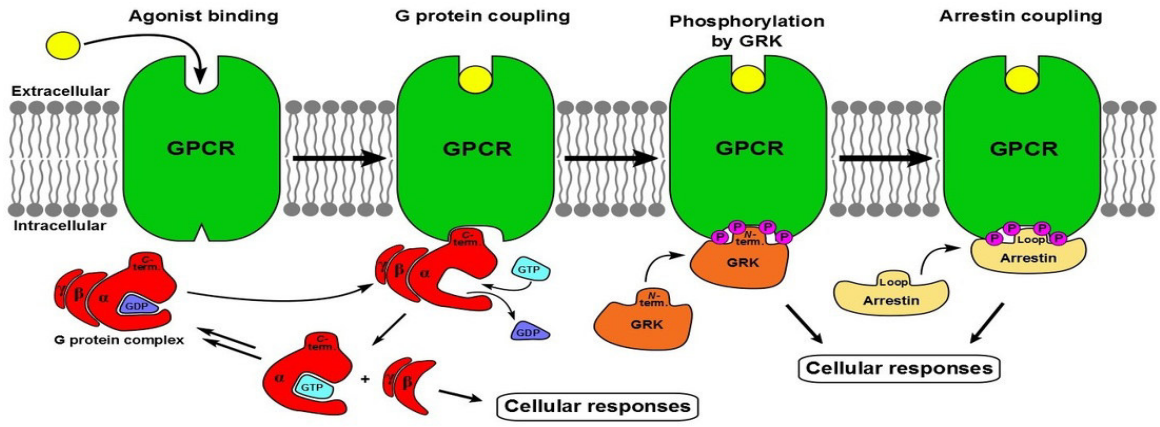


Figure8 - GPCR (G Protein coupled receptor)

2.1.4 History and significance

The 2012 Nobel Prize in Chemistry was awarded to Brian Kobilka and Robert Lefkowitz for their work that was "crucial for understanding how G protein–coupled receptors function".https://en.wikipedia.org/wiki/G_protein%E2%80%93coupled_receptor - cite_note-Nobel_committee.2C2012-8 There have been at least seven other Nobel Prizes awarded for some aspect of G protein–mediated signaling. As of 2012, two of the top ten global best-selling drugs (AdvairDiskus and Abilify) act by targeting G protein–coupled receptors.

2.1.5 Classification

The exact size of the GPCR super family is unknown, but nearly 800 different human genes (or ~ 4% of the entire protein-coding genome) have been predicted to code for them from genome sequence analysis. Although numerous classification schemes have been proposed, the super family was classically divided into three main classes (A, B and C) with no detectable shared sequence homology between classes. The largest class by far is class A, which accounts for nearly 85% of the GPCR genes. Of class A GPCRs, over half of these are predicted to encode olfactory receptors, while the remaining receptors are liganded by known endogenous compounds or are classified as orphan receptors. Despite the lack of sequence homology between classes, all GPCRs have a common structure and mechanism of signal transduction. The very large rhodopsin a group has been further subdivided into 19 subgroups (A1-A19). More recently, an alternative classification system called GRAFS (Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2, Secretin) has been proposed. By this system, GPCRs can be grouped into 6 classes based on sequence homology and functional similarity:

- Class A (or 1) (Rhodopsin-like)
- Class B (or 2) (Secretin receptor family)
- Class C (or 3) (Metabotropic glutamate/pheromone)
- Class D (or 4) (Fungal mating pheromone receptors)
- Class E (or 5) (Cyclic AMP receptors)
- Class F (or 6) (Frizzled/Smoothened)

An early study based on available DNA sequence suggested that the human genome encodes roughly 750 G protein–coupled receptors, about 350 of which detect hormones, growth

factors, and other endogenous ligands. Approximately 150 of the GPCRs found in the human genome have unknown functions.

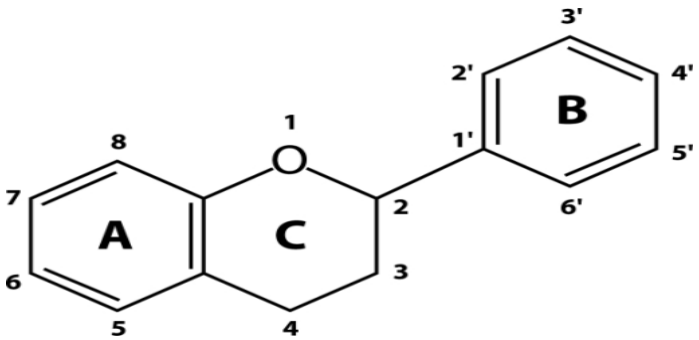
Some web-serversand bioinformatics prediction method shave been used for predicting the classification of GPCRs according to their amino acid sequence alone, by means of the pseudo amino acid composition approach.

2.1.6 Class B human glucagon G protein coupled receptor

The glucagon receptor (GCGR) is one of the 15 members of the secretin-like (class B) family of G protein-coupled receptors (GPCRs) in humans. GCGR is activated by the 29 amino acid hormonal peptide, glucagon and is a potential drug target for type 2 diabetes. During fasting, the pancreas dispatches glucagon to activate GCGR in the liver causing the release of glucose into the blood. Despite less than 15% sequence homology between class A (rhodopsin-like) and class B GPCRs, many of these receptors presumably share a 7 transmembrane (7TM) helical domain and similar signal transduction mechanisms1 . Although a structure-function understanding of the class a family of GPCRs has been greatly advanced during the last few years, a detailed understanding of class B GPCRs has lagged due to the lack of a 7TM domain structure for these receptors. All class B GPCRs contain a globular N-terminal extracellular domain (ECD) defined by three conserved disulfide bonds and a 7TM domain. They are activated by hormonal peptides that bind to both the ECD and the 7TM domain. Structural details of soluble ECDs, including the ECD of GCGR, and their role in the selective recognition of peptide hormones' C-termini have been revealed for several class B receptors.

2.1.7 Flavonoid

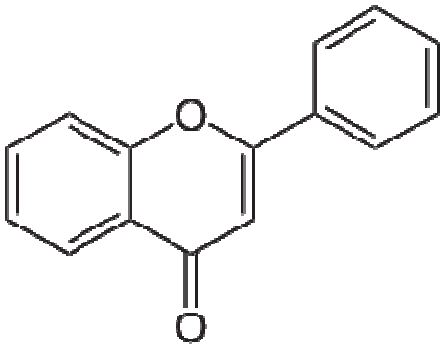
The word "Flavonoids or bioflavonoid" is isolated from the Latin word flavus meaning yellow due to their color in nature. They are a class of secondary metabolites of plant and fungus .Chemically, flavonoids have the general structure of a 15-carbon skeleton, which consists of two phenyl rings (A and B) and heterocyclic ring (C). This carbon structure can be abbreviated C6-C3-C6.



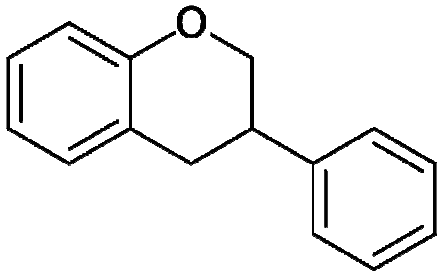
According to the IUPAC nomenclature, they can be classified into:

- Flavonoids or bioflavonoids
- Isoflavonoids, derived from 3-phenylchromen-4-one (3-phenyl-1,4-benzopyrone) structure
- Neoflavonoids, derived from 4-phenylcoumarine (4-phenyl-1,2-benzopyrone) structure

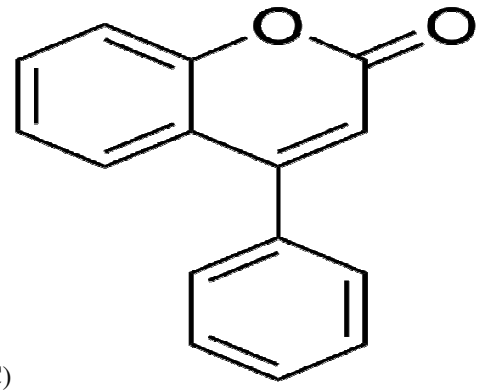
The three flavonoid classes above are all ketone-containing compounds, and as such, are anthoxanthins (flavones and flavonols). This class was the first to be termed bioflavonoids. The terms flavonoid and bioflavonoid have also been more loosely used to describe non-ketone polyhydroxy polyphenol compounds which are more specifically termed flavanoids. The three cycle or heterocycles in the flavonoid backbone are generally called ring A, B and C. Ring A usually shows a phloroglucinol substitution pattern.



(A)



(B)



(C)

Figure9: Molecular structure of (A) flavone, (B) Isoflavan and (C) Neoflavonoids**.**

Flavonoids are a group of plant metabolites thought to provide health benefits through cell signaling pathways and antioxidant effects. These molecules are found in a variety of fruits and vegetables. Flavonoids are polyphenolic molecules containing 15 carbon atoms and are soluble in water. They consist of two benzene rings connected by a short three carbon chain. One of the carbons in this chain is connected to a carbon in one of the benzene rings, either through an oxygen bridge or directly, which gives a third middle ring. The flavonoids can be divided into six major subtypes, which include chalcones, flavones, isoflavonoids, flavanones, anthoxanthins and

Anthocyanins. Many of these molecules, particularly the anthoxanthins give rise to the yellow color of some petals, while anthocyanins are often responsible for the red color of buds and the purple red color of autumn leaves. Flavonoids are abundant in plants, in which they perform several functions. They are essential pigments for producing the colors needed to attract pollinating insects. In higher order plants, flavonoids are also required for UV filtration, nitrogen fixation, cell cycle inhibition, and as chemical messengers. Flavonoids secreted by a plant's roots aid the symbiotic relationship between rhizobia and certain vegetables such as peas, clover and beans. The rhizobia present in soil produce Nod factors in response to the presence of flavonoids. These Nod factors are then recognized by the plant, which induces certain responses such as ion fluxes and root nodule formation. Some flavonoids also inhibit certain spores to protect against certain plant diseases. Flavonoids are ubiquitous in plants and are the most common type of polyphenolic compound found in the human diet. The abundance of flavonoids coupled with their low toxicity relative to other plant compounds means they can be ingested in large quantities by animals, including humans. Examples of foods that are rich in flavonoids include onions, parsley, blueberries, and bananas, dark chocolate and red wine.

2.1.8 Health benefits to humans

Flavonoids are important antioxidants, and promote several health effects. Aside from antioxidant activity; these molecules provide the following beneficial effects:

Antiviral

Anticancer

Anti-inflammatory

Ant allergic

One flavonoid called quercetin can help to alleviate eczema, sinusitis, asthma, and hay fever. Some studies have shown that flavonoid intake is inversely related to heart disease, with these molecules inhibiting the oxidation of low-density lipoproteins and therefore reducing the risk of atherosclerosis developing. Flavonoids are also abundant in red wine, which some have theorized is the reason why the incidence of heart disease may be lower among the French (who have a relatively high red wine intake) compared with other Europeans,despite a higher consumption of foods rich in cholesterol (French paradox).Many studies have also shown that one to two glasses of wine a day can help protect against heart disease. Some types of tea are also rich in flavonoids and their consumption is thoughto lower levels of triglycerides and cholesterol in the blood. Soy flavonoids or isoflavones also lower cholesterol, as well as protecting against osteoporosisand alleviating the symptoms of menopause.The daily intake of dietary flavonoids typically ranges from anywhere between50 and 500 mg, meaning the contribution to antioxidant activity varies widely between individuals.

2.2 Food sources

Almost all fruits, vegetables and herbs contain a certain amount of flavonoids. They can also be found in other food sources including dry beans, grains, red wine and green and black teas. The general rule is that the more colorful food item is, the richer it will be in flavonoids. Oranges however are an exception to the rule because the flavonoids contained in this fruit are mainly found in the white and pulp interior of the skin. The best way to ensure a good intake of flavonoids is to consume plenty of fresh fruit and vegetables on a daily basis. Experts advise eating five servings of vegetables and four of fruit. Regarding the intake of red wine, men are advised not to drink more than two glasses per day and women should not drink more than one glass per day. Flavonoid supplements are also available, but people who buy these should note that experts have not confirmed what the ideal flavonoid intake is and an excess intake may even be harmful.Interactions certain drugs do interact with flavonoids. Studies have shown that the enzyme cytochrome P450, which is involved in the metabolism of drugs, is inhibited by flavonoids. An efflux transporter called P-glycoprotein, which decreases the absorption of certain drugs, is also affected. Flavonoids have also been shown to interact with certain nutrients. They can bind to nonheme iron, thereby decreasing its absorption in the intestine. Some flavonoids also inhibit cellular uptake of vitamin C and some experts advise avoiding flavonoid rich foods or drinks when taking vitamin C.

2.2.1 Binding pattern of Flavonoids with their Receptor

The glucagon receptor is available on Protein Data Bank as a PDB ID:4L6R. It is available by “Structure of the class B human glucagon G protein coupled receptor” on3.3 Åresolution. As per the reported literature available, the favorable binding sites on this protein are Tyr 138, Gln 142, Tyr 149, Val 191, Gln 232, Trp 295, Thr 296, Asn 298, Glu 362, Phe 365 and Leu 386 residues.

2.2.2 Computer Aided Drug Design

A few years ago, the National Institutes of Health (NIH) created the Biomedical Information Science and Technology Initiative (BISTI) to examine the current state of bioinformatics in the United States. BISTI’s working definition of bioinformatics included its use in biomedical research, in particular for drug discovery and development programs. Bioinformatics was seen as an emerging how drugs are found, brought to clinical trials and eventually released to the marketplace. Computer-Aided Drug Design (CADD) is a specialized discipline that uses computational methods to simulate drug-receptor interactions. CADD methods are heavily dependent on bioinformatics tools, application and on the support side of the hub, information technology, information management, software applications, databases and computational resources all provide the infrastructure for bioinformatics. On the scientific side of the hub, bioinformatics methods are used extensively in molecular biology, genomics, proteomics, other emerging areas (i.e. metabolomics, transcriptomics) and in CADD research. There are several key areas where bioinformatics supports CADD research.

2.2.3Virtual screening

Virtual screening uses computer based methods which discovers new ligand on the bases of biological structure. The basic goal of the virtual screening is the reduction of the enormous virtual chemical space of small organic molecules, to synthesize and/or screen against a specific target protein, to a manageable number of the compound that inhibit a highest chance to lead to a drug candidate. In theory, the applicability of virtual screening is limited only by what properties of a compound can be calculated computationally and the perceived relevance of those properties to the problem in hand. On a practical level, further considerations include the timescale for calculation of the properties, which may be considerable for a database of, say, one million compounds. Many drug candidates fail in the clinical trials because of the reasons unrelated to the potency against intended drug target. Pharmacokinetic & toxicity issues are blamed for more than half of the failure in the clinical

trials. Therefore first part of the visual screening evaluates the drug likeness of the same molecules most independent of their intended drug target. The term virtual screening has been used to describe a process of computationally analyzing large compound collections in order to prioritize compounds for synthesis or assay. A broad range of computational techniques can be applied to the problem. In our work we have focused on explicit receptor–ligand molecular docking as a means of yielding the most detailed model of the way in which a given ligand binds to a receptor, and hence the most informative basis on which to assess which ligands are useful candidates for synthesis or assay. Although the underlying methods of virtual screening have been in use in various guises for several years, it is worth noting the recent impact on molecular modeling made by the increased availability of high-performance computing platforms. Affordable multiprocessor workstations and PC (personal computer) clusters have enabled the modeler to employ computationally demanding algorithms on a routine basis. This change is particularly relevant in the case of virtual screening, where as in the work described in this paper, computationally intensive methods such as molecular docking must be applied to very large databases of chemical structures.

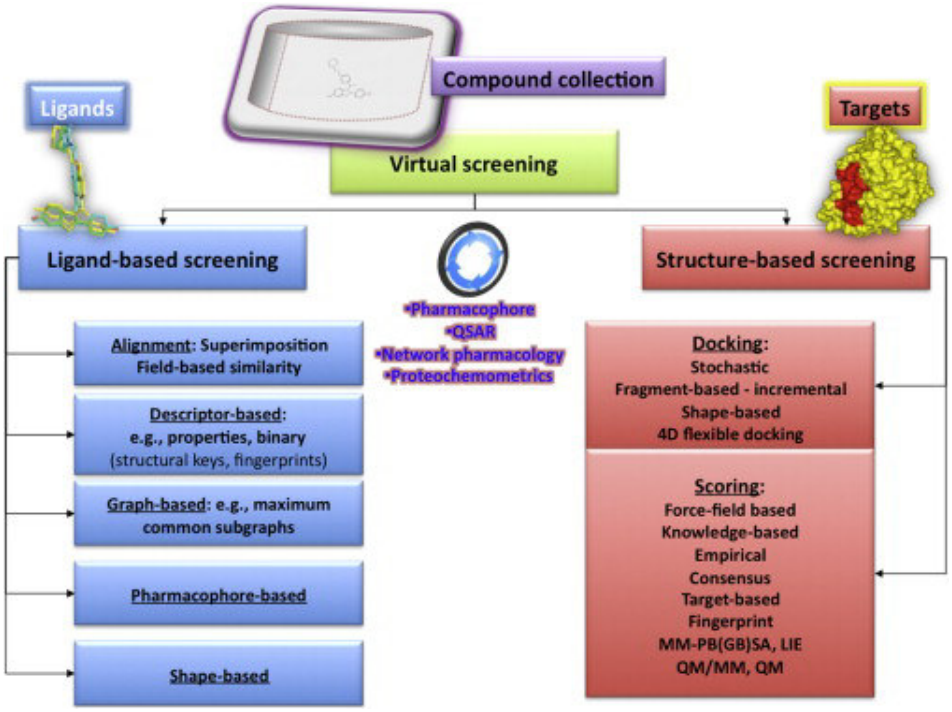


Figure10-virtual screening

2.2.4 Type of virtual screening

1:-Drug likeness screening- Many drug candidates fail in the clinical trials reasons is unrelated in the potency against the intended drug target. Pharmacokinetic and toxicity issues are blamed for more than half of all failure in the clinical trials. Therefore first part of the virtual screening evaluates druglikness of small molecules, drug like molecules exhibit favorable absorption, distribution, metabolism, excretion, toxicological (ADMET) parameters. Using following types of method currently assesses Druglikness

□ Simple counting method

□ Functional group filter

□ Topological filter

□ Pharmacophore filter

2:- Structure Based Virtual Screening (Docking) -High throughput screening docking and scoring techniques can be applied to computationally screening a database of hundreds of thousands of compound against a target of proteins. Computational methods that predict the 3D structure of a protein ligand complex are often referred to as molecular docking approaches.

2.2.5 PHARMACOPHORE

The concept of Pharmacophore was first introduced in 1909 by Ehrlich, who defined the Pharmacophore as a molecular framework that carries (Phoros) the essential features responsible for a drug's (Pharmacon) biological activity". Pharmacophore approaches have become one of the major tools in drug discovery after the past century's development. A Pharmacophore model is a geometrical description of the chemical functionalities required of a ligand to interact with the receptor. IUPAC defines a Pharmacophore to be "an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response". Typical Pharmacophore features include hydrophobic centroids, aromatic rings, hydrogen bond acceptors or donors, cations and anions.

These Pharmacophoric points may be located on the ligand itself or may be projected points presumed to be located in the receptor. Various ligand-based and structure-based methods have been developed for improve Pharmacophore modeling and have been successfully and extensively applied in virtual screening, de novo design and lead optimization.

Hydrophobicity

In chemistry, hydrophobicity is the physical property of a molecule (known as a hydrophobe) that is seemingly repelled from a mass of water. Hydrophobic molecules tend to be non-polar and, thus, prefer other neutral molecules and non-polar solvents. Hydrophobic molecules in water often cluster together, forming micelles.

Hydrogen Bond Donor

The hydrogen bond donor in a hydrogen bond is the atom to which the hydrogen atom participating in the hydrogen bond is covalently bonded, and is usually a strongly electronegative atom such as N, O, or F. the hydrogen donor is strongly electronegative, it pulls the covalently bonded electron pair closer to its nucleus, and away from the hydrogen atom.

Hydrogen Bond Acceptors

The hydrogen acceptor is the neighboring electronegative ion or molecule, and must possess a lone electron pair in order to form a hydrogen bond. This feature matches the following types of atoms or groups of atoms with surface accessibility: sp or sp^2nitrogens that have a lone pair and charge less than or equal to zero, sp^3oxygens or sulfurs that have a lone pair and charge less than or equal to zero.

Negative ionizable

This feature matches atoms or groups of atoms that are likely to be deprotonated at physiological pH, such kind of molecules are: Trifluoromethyl sulfonamide hydrogens, Sulfonic acids (centroid of the three oxygen), Phosphoric acids (centroid of the three oxygens).

Positive ionizable

This feature matches atoms or groups of atoms that are likely to be protonated at physiological pH, such ionizable molecules are such as: Basic amines, Basic secondary amidines (iminyl nitrogen), Basic primary amidines, except guanidines (centroid of the two nitrogen), Basic guanidines. Positive charges are not adjacent to a negative charge.

Importance of the Pharmacophore

A Pharmacophore model explains how structurally diverse ligands can bind to a common receptor site. The features need to match different chemical groups with similar properties, in order to identify novel ligands.

Ligand based Pharmacophore modeling

A detection of the three-dimensional binding mode between a protein and a chemical compound is valuable to optimize drug candidates for high-throughput screening. Pharmacophore models are essential functional groups of atoms in the proper three-dimensional position to interact with a given receptor, and widely used for drug design. Recent Pharmacophore models can be classified into two categories that are receptor-based Pharmacophores and ligand-based Pharmacophores. For a receptor with a known three-dimensional structure, receptor-based Pharmacophores have been studied which are based on the famous concept of a key for the lock.

On the other hand, since there are many proteins whose three-dimensional structures have not been known, ligand-based Pharmacophore models are still useful. Traditionally, ligand-based Pharmacophore models are computed by extracting common features among three-dimensional structures of compounds which are known to interact with a target protein. To do this, possible conformers of compounds should be previously enumerated. However, calculation of common features among the conformers of compounds with many rotatable bonds takes much computational time since such compounds have an extremely large number of conformers. We now focus on ligand-based Pharmacophore models and propose a new method to shorten the computational time for generating Pharmacophores of a set of chemical compounds with many rotatable bonds. Our method consists of the following processes: First, we divide bonds of each compound into two parts, where one is a set of rotatable bonds and another is a set of non-rotatable bonds. Second, we compute hard part sets which mainly include non-rotatable bonds. We then generate a few conformers for each hard part set. Finally, we superpose the compounds each other by overlapping the three-dimensional structures among the hard parts and extracting common substructures as molecular graphs among the other parts of compounds. By using our method, since we do not have to enumerate all the conformers of a compound, we can save much computational time for ligand-based Pharmacophore modeling.

2.2.6 Molecular docking

Docking and scoring is the core step of Ligand-based drug design. **Docking** is the process that places a ligand molecule into the binding site of a receptor protein and then optimizes the relative orientation and conformation of the ligand to interact with the protein. Then each pose is scored by a proper scoring function. The main aim of molecular docking is trying to predict the 3D structure of the complex between a ligand and a protein. To do this a docking program needs two components:

- **Search algorithm**
- **Scoring function**

In the molecular docking process, the search space (the set of all possible solutions to a problem) consists of all possible orientations and conformations of the protein paired with the ligand. Because it is impossible to explore the search space exhaustively, is important to use of search algorithms for sampling the search space. Then for each protein-ligand pose, the scoring function predicts the strength of the interaction between the two molecules (binding affinity). In other terms, the scoring function takes a pose as input and returns a number indicating the probability that the pose represents a favorable binding interaction. Reproducing the conformational space accessible to a macromolecule is a very difficult task and requires unavoidable approximation. Docking procedures can be classified into three categories depending on their degree of approximation:

- **Rigid body docking**: This is a highly simplistic model that regards both the ligand and the protein as two rigid solid bodies.
- **Semi-flexible docking**: This model is asymmetric; one of the molecules, usually the smaller ligand, is considered flexible, while the protein is regarded as rigid.
- **Flexible docking**: Both molecules are considered flexible, although clearly the extent of flexibility of either (or of both) is necessarily limited, or simplified.

Ligand flexibility is computationally easier to handle as ligands are much smaller than macromolecules, and thus it is standard in docking protocols. The ideal docking method would allow both ligand and receptor to explore their conformational degree of freedom. However, such calculations are computationally very demanding and most of the methods only consider the conformational space of the ligand while the receptor is assumed to be rigid. The search algorithm is used to generate ligand conformations inside the protein binding pocket. Algorithms can be grouped into:

- Deterministic algorithms which are reproducible,
- Stochastic algorithms which include a random factor and are thus not fully reproducible.

The different docking algorithms used in the molecular docking studies are:

2.2.7 Genetic Algorithm This algorithm mimics the process of evolution by manipulating a collection of data structures called chromosomes. Each of these chromosomes encodes a possible solution to the problem to be solved. In case of docking, each

chromosome encodes a possible ligand-receptor complex conformation and to each of them is assigned a fitness score on the basis of the relative quality of that solution in terms of protein-ligand interactions. The genetic algorithm is used by the software GOLD and is available as search algorithm in MOE-Dock.

2.2.7.1 Tabu search This algorithm is characterized by the imposition of restrictions to enable the search process to explore otherwise difficult regions. These restrictions take the form of a Tabu list that stores a number of previously visited solutions. By preventing the search to visit these regions, the exploration of new search space is encouraged. While the genetic algorithm usually converges quickly at the close proximity of a global minimum, it can be trapped in a local minimum. Using a Tabu list helps in avoiding this drawback.Tabu search is available as search algorithm in MOE-Dock.

2.2.7.2 Simulated Annealing It is a special molecular dynamics simulation, in which the system is cooled down at regular time intervals by decreasing the simulation temperature. The system thus gets trapped in the nearest local minimum conformation. Disadvantages of simulated annealing are that the result depends on the initial placement of the ligand and that the algorithm does not explore the solution space exhaustively. Simulated annealing is available as search algorithm in MOE-Dock.

2.2.7.3 Glide Algorithm The Glide (Grid-based Ligand Docking with Energetics) algorithm approximates a systematic search of positions, orientations and conformations of the ligand in the receptor binding site using a series of hierarchical filters. The shape and properties of the receptor are represented on a grid by several different set of fields that provide progressively more accurate scoring of the ligand pose. The fields are computed prior to docking. The binding site is defined by a rectangular box confining the translation of the center of mass of the ligand. A set of initial ligand conformations is generated through exhaustive search of the torsional minima, and the conformers are clustered in a combinatorial fashion. Each cluster, characterized by a common conformation of the core and an exhaustive set of rotamer group conformations, is docked as a single object in the first stage. The search begins with a rough positioning and scoring phase that significantly narrows the search space and reduces the number of poses to be further considered to a few hundred. In the following stage, the selected poses are minimized on pre-computed OPLS-AA Vander Waals and electrostatic grids for the receptor. In the final stage the lowest-energy poses obtained are subjected to a Monte-carlo procedure in which nearby torsional minima

are examined, and the orientation of peripheral groups of the ligand is refined. These minimized poses are finally rescored.

2.3 Plants The docking algorithm PLANTS (**Protein-Ligand ANT System**) is based on a class of stochastic optimization algorithms called **ant colony optimization** (ACO). ACO is inspired by the behavior of real ants finding a shortest path between their nest and a food source. The ants use indirect communication in the form of pheromone trails which mark paths between the nest and a food source. In the case of protein-ligand docking, an artificial ant colony is employed to find a minimum energy conformation of the ligand in the binding site. These ants are used to mimic the behavior of real ants and mark low energy ligand conformations with pheromone trails. The artificial pheromone trail information is modified in subsequent iterations to generate low energy conformations with a higher probability. Scoring function are fast approximate mathematical methods used to predict the strength of the non-covalent interaction between two molecules given the 3D structure of their complex and therefore their binding affinity. There are three general classes of scoring functions:

2.4 Force field-based scoring functions based on the non-bonded terms of a classical molecular mechanics force field. Usually a Lennard-Jones potential describes Vander Waals interactions and the Coulomb energy describes the electrostatic components of the interaction. A major disadvantage of these scoring functions lies in the fact that is unclear to what extend they can be applied to protein-ligand complexes not represented in the training set used for deriving the master equation. GoldScore and MOE Energy score are some examples.

Empirical scoring functions based on counting the number of various types of interactions between the two binding partners. They use several terms describing properties known to be important in drug binding to construct a master equation for predicting binding affinity. Multi-linear regression is used to optimize the coefficients to weight the computed terms using a training set of protein-ligand complexes for which both the binding data and the high resolution 3D crystal structures are known. ChemScore and Glidescore are some examples.

Knowledge-based scoring functions based on statistical observations of intermolecular close contacts in large 3D databases which are used to derive "potentials of mean force". This method is founded on the assumption that close intermolecular interactions between certain types of atoms or functional groups that occur more frequently than one would expect by a random distribution are likely to be energetically favorable and therefore contribute favorably to binding affinity.

2.4.1 AutoDock

In this study, we have employed Autodock docking software for dock the ligand in protein active site. AutoDock is a suite of automated docking tools. It is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure .Current distributions of AutoDock consist of two generations of software:
AutoDock 4, AutoDockVina.

AutoDock 4 Actually performs the docking of the ligand to a set of grids describing the target protein; autogrid pre-calculates these grids. In addition to using them for docking, the atomic affinity grids can be visualized. This can help, for example, to guide organic synthetic chemists design better binders. The introduction of AutoDock 4 comprises three major improvements:

1. The docking results are more accurate and reliable.
2. It can optionally model flexibility in the target macromolecule.
3. It enables AutoDock's use in evaluating protein-protein interactions.

AutoDock 4.0 not only is faster than earlier versions, it allows side chains in the macromolecule to be flexible. As before, rigid docking is blindingly fast, and high-quality flexible docking can be done in around a minute. Up to 40,000 rigid dockings can be done in a day on one CPU. AutoDock 4.0 now has a free-energy scoring function that is based on a linear regression analysis, the AMBER force field, and an even larger set of diverse protein-ligand complexes with known inhibition constants than we used in AutoDock 3.0. This is enough to discriminate between leads with milli-, micro- and nano-molar inhibition constants.

AutoDockVina does not require choosing atom types and pre-calculating grid maps for them. Instead, it calculates the grids internally, for the atom types that are needed, and it does this virtually instantly.

Materials and Methods

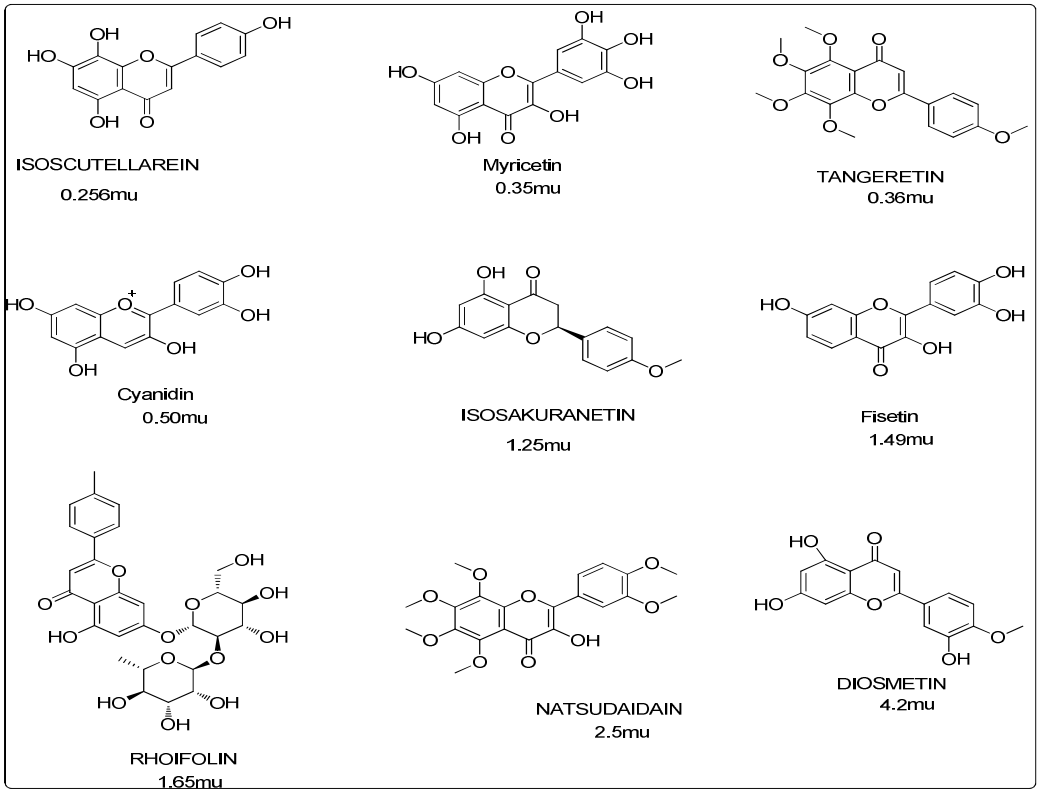
3.1 Pharmacophore generation

For the *in-silico* designing of new chemical entities there are basics two weapons. One is structure based drug design and another is ligand based drug design. Here we followed ligand based pharmacophore modelling with utilizing 34 flavonoids structure which have proper anti-diabetic activities on glucagon receptors.

3.2 Data Preparation

Based on the previous published literature I collected the 34 flavonoids structure along with their biological activities (IC50) using the same biological assay. Those are distinguished into training set and test set based on activities.

The 2D structure of the 34 molecule was sketched using ChemBiodraw program version 12.0, and those are converted into 3D structures using Discovery studio 4.1 (DS). We further used marvin-view program for reassignment of the bond order and cleaning of the steric clashes and geometric planarity of the molecules.



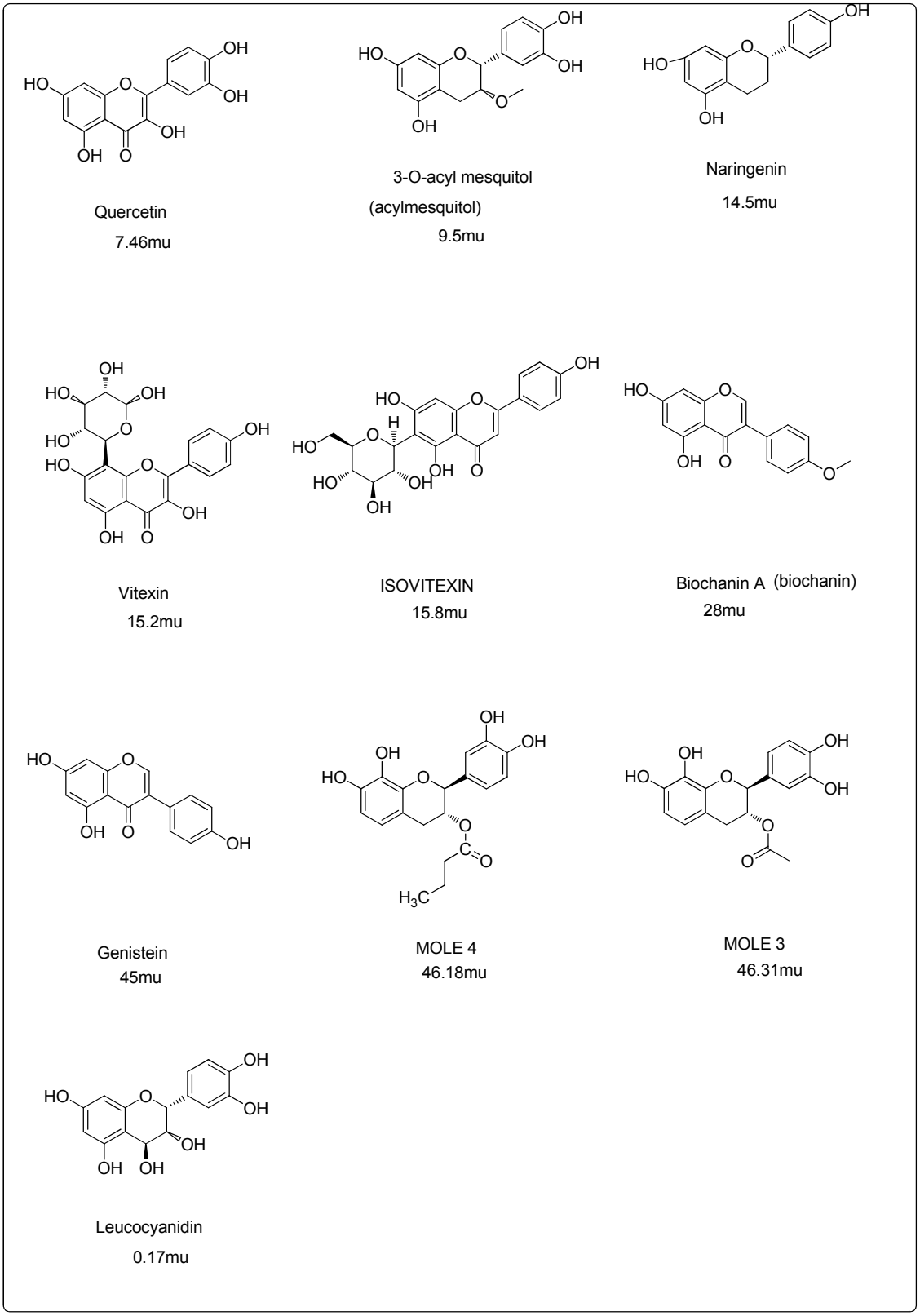
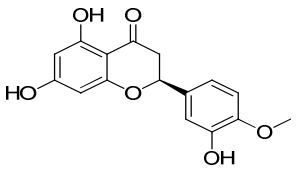
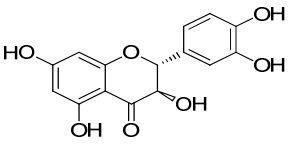


Figure11: Structure of training set compounds. The 2D chemical structures of 19 compounds of the training set together with their experimental IC_{50} values in μM.



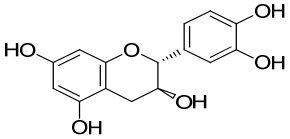
HESPERETIN

0.28mu



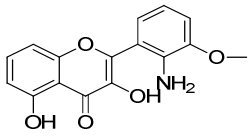
Taxifolin

0.35mu

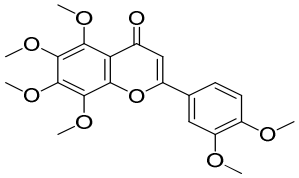


Catechin

0.37mu

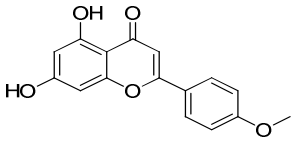


2'Amino-3'-methoxy flavone
PD 98059
(methoxy)
1mu



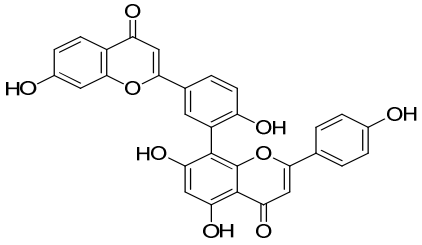
NOBILETIN

1.49mu



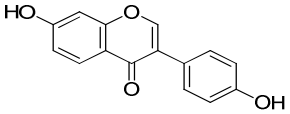
ACACETIN

2.74mu



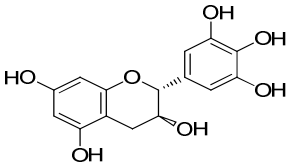
Amentoflavone

1.75mu



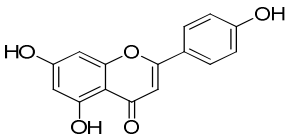
Daidzein

10mu



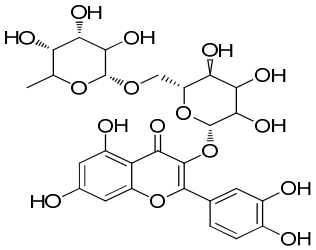
Gallocatechol

7.46mu



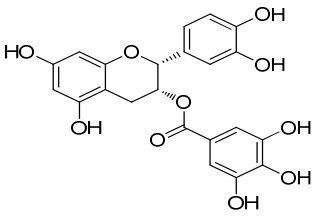
Apigenin

10.5mu



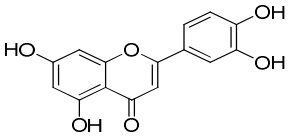
Rutin

15mu



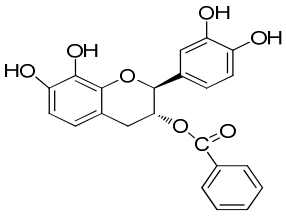
Epicatechin gallate(epigallate2)

29mu



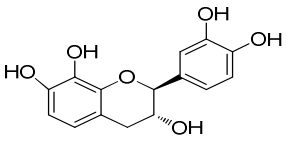
Luteolin

45.5mu



MOLE 9

77.76mu



MOLE 2

82.32mu

Figure 12: Structure of test set compounds. The 2D chemical structures of 15 compounds of the training set together with their experimental IC_{50} values in μM.

In the next step 34 Molecule divided in to training set (Fig11) and test set (Fig12).Training set contained 19molecules and Test set Consisted of 15 molecules. Then Merck Molecular Force Field (MMFF) was applied in both training set and test sets compounds. Next step was Energy minimization(Best) by performing 1000 steepest descent. These steps were necessary to produce a good set of representative conformations of different conformational space accessible to a molecule within a given energy range. A maximum of 255 conformations were generated for each molecule by keeping the RMS gradient 0.01 kcal/mol.

3.2.1 Features mapping

Feature mapping protocol was used to identify the different features of the training set molecules in Discovery studio. Feature mapping protocol only applied on training set molecule by different features like Hydrogen bond acceptor (HBA), Hydrogen Bond Donor (HBD), Positive Ionization, Negative Ionization, Hydrophobicity, Ring aromaticity with their default values.

3.2.3 Pharmacophore Model Generation

HypoGen algorithm was applied to build the pharmacophore model and in the present study four features, hydrogen bond donors(HBD), hydrogen bond acceptors (HBA), ring aromatic(RA), and hydrophobic (HY), were selected to generate the pharmacophore hypotheses . HypoGen generates pharmacophore model based on chemical features of active compounds in training set.3D QSAR module in Discovery Studio (DS) was used for developing the pharmacophore. Ligand-based pharmacophore modelling methodologies is two types; one is a Common-Feature-Pharmacophore modelling. Its utilizes common chemical features present only in the most active molecule. Another is the 3D QSAR pharmacophore methodology uses the chemical features of most active and inactive molecule with their biological activity. Here we used 3D QSAR pharmacophore methodology with utilizing training sets for generation of 10 pharmacophores.

3.2.3 Pharmacophore Validation

The pharmacophore model is validated by three steps:

1. Cost analysis
2. Fischer's randomization test
3. Test set prediction

The modelis described in terms of fixed cost, total cost, and null cost. The fixed cost represents the simplest model and it fits the data perfectly. The null cost represents no features with high cost value and it estimates the activity to be average activity of the training set molecule. The best model was selected based on the difference between the two cost values (null cost- total cost). The difference between the costs is greater than 60 means, the model has excellent true correlation. If the difference is 40–60, the model has prediction correlation of 70–90%, and if the difference is below 40, it may be difficult to predict the model. Fischer's randomization is the second approach to validate the pharmacophore model. The 90% confidence level was selected to validate the study and 9 random spread sheets were constructed. This method generates the hypotheses by randomizing the activity of the training set compounds. The correlation between the structure and biological activity was validated by this method. The Ligand pharmacophore mapping module in Discovery Studio was used to map the ligands and estimate the predicted activity of the test set molecule.

3.2.4 Database screening

Virtual screening studies were used to find novel and potential leads from virtual database forfurther development. The virtual screening studies were used to find novel leads for Diabetes. The Hypo1 model was used as a 3D query in database screening, and the ZINC database containing 17900742molecules for screening. Ligand pharmacophore mapping protocol was used with flexible search option to screen the database. Hit compounds from the

database with estimated activity less than 0.1 μM were selected for further screening using

Lipinski's rule of five; compounds have (i) molecular weight less than 500, (ii) hydrogen donors less than 5, (iii)hydrogen acceptors less than 10, and (iv) an octanol/water partition

coefficient (Log*P*) value less than 5.

3.3 Molecular Docking

MolecularDocking plays an important role in the rational drug design. Protein preparation was the main step in docking and all ligands were docked into the active site of the receptor.

Preparation of protein structure including remove water molecule and adding hydrogen atoms to target protein and applying MMFF force field. After filtration I have got the 250 molecules, These 250 zinc molecules were docked with the glucagon receptors(PDB ID-4L6R) using the CDocker protocols in Discovery Studios.Binding sites were selected depending on the 11 active residues such as TYR138,GLN142, TYR149, VAL191, GLN232, TRP295, THR296, ASN298, GLU362, PHE365, LEU386 present on the receptors. After docking we have got 20 molecules out of 250 molecules, which have least binding energies and proper binding interaction.

CHAPTER-IV

RESULTS AND DISCUSSION

A set of ten pharmacophore models was generated using a training set containing 19 compounds by selecting HBA (Hydrogen bond acceptor), HBD (Hydrogen bond donor), Hydrophobic (HY), PI (Positive ionizable) and Ring aromatic (RA) features as suggested by Feature Mapping protocol. The selected hypothesis was made of ring aromatic and hydrogen bond acceptors features. It forms a triangular structure(Figure13 A). Among the total 10 pharmacophore models (Hypotheses), we selected 1st hypothesis depending on the largest cost difference. The cost difference between the total cost and null cost must be greater and it should be smaller between total cost and fixed cost values for a significant pharmacophore model. The total cost values of ten pharmacophore models ranged from 44.046 to 20.417 bits. In our study, the pharmacophore generation runs calculated a fixed cost value of 95.9307bitsand the null cost value of 182.265 bits. Among the total cost values of generating ten pharmacophore models, first model (Hypo1) has scored the value closer to the fixed cost value when compared to other models. The cost difference between the null cost and total cost value of the first pharmacophore model is 44.046 bits (Table 3). The cost difference value greater than 70 bites imply that the pharmacophore model correlates the experimental and estimated activity values more than 90%. In this study, the cost difference value of Hypo1 signifies that it can correlate the experimental and estimated activity values of the training set compounds more than 90%. Hypo1 was made of four pharmacophoric features consisting two HBA, two Ring aromatic features (Figure 13). Further evaluation of the generated pharmacophore models was based on the correlation coefficient. The correlation values of these ten Pharmacophore models were >0.782897, and the first three pharmacophore models correlated the activity data with high correlation values, i.e., above 0.782897. Every training set molecule generated 10 hypotheses along with their different fit score and correlation coefficient and RMSD variance. For our model the correlation coefficient was 0.782897 and RMSD gradient was 1.68205 (Table 3). Each and every molecule in the training set was nicely mapped along with the resultant pharmacophore with a largest fitting score 7.55 and lowest fitting score 5.24 (Table 2).

The intermediate distance between ring aromaticity and hydrogen bond donor was calculated and the distance between ring aromaticity and hydrogen bond donor was 6.671Å and 5.030Å and the inter-features distance between hydrogen bond acceptors was 5.030Å (Figure 13).

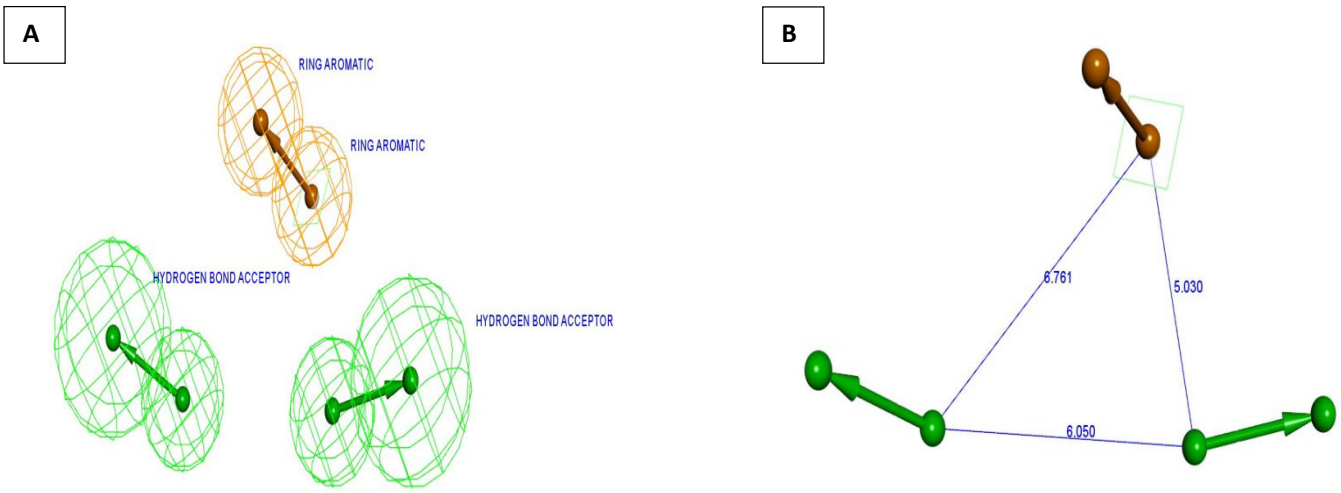


Figure1:

The best HypoGen pharmacophore model, Hypo1. (A) Chemical features present in Hypo 1 (B) 3D spatial arrangement and the distance constraints between the chemical features. Green color represents HBA, Brown color represents Ring aromatic.

Table 2: Experimental and estimated IC_{50} values of the training set compounds based on best Pharmacophore Hypothesis Hypo1.

Name	Fit Score	Estimated activity (IC_{50})	Experimental Activity (IC_{50})	Error	Status
Leucocyanidin	6.75	1.9	0.17	+11	active
Isoscutellarein	6.64	2.5	0.26	+9.8	active
Hesperetin	7.54	0.31	0.28	+1.1	active
Taxifolin	7.55	0.31	0.35	-1.1	active
Tangeretin	7.00	1.1	0.36	+3	active
Catechin	6.36	4.7	0.37	+13	active
Cyanidin	6.81	1.7	0.5	+3.4	active
Methoxy	6.69	2.2	1	+2.2	moderately active
Isosakuranetin	7.14	0.8	1.2	-1.6	moderately active
Fisetin	6.73	2	1.5	+1.4	moderately active
Nobiletin	6.99	1.1	1.5	-1.3	moderately active
Rhoifolin	7.30	0.54	1.6	-3	moderately active
Amentoflavone	7.05	0.98	1.8	-1.8	moderately active
Acacetin	6.27	5.8	2.7	+2.1	moderately active
Diosmetin	6.26	5.9	4.2	+1.4	moderately active
Gallocatechol	6.35	4.9	7.5	-1.5	moderately active
Quercetin	6.73	2	7.5	-3.7	moderately active
acylmesquito	6.30	5.5	9.5	-1.7	moderately active
Apigenin	6.25	6.1	11	-1.7	moderately active
Naringenin	5.24	62	14	+4.3	moderately active
Vitexin	6.09	8.8	15	-1.7	moderately active
Isovitexin	6.18	7.1	16	-2.2	moderately active
biochanin	5.75	19	28	-1.5	moderately active

Genistein	5.74	20	45	-2.3	moderately active
Luteolin	6.28	5.7	45	-8	moderately active
Mole4	5.97	12	46	-3.9	moderately active
Mole3	5.76	19	46	-2.5	moderately active
Mole9	5.61	27	78	-2.9	moderately active

***a Positive value indicates that the estimated IC_{50} is higher than the experimental IC_{50}; negative value indicates that the estimated IC_{50} is lower than the experimental IC_{50}.[b]fit value indicates how well the features in the Pharmacophore map the chemical feature in the compound.**

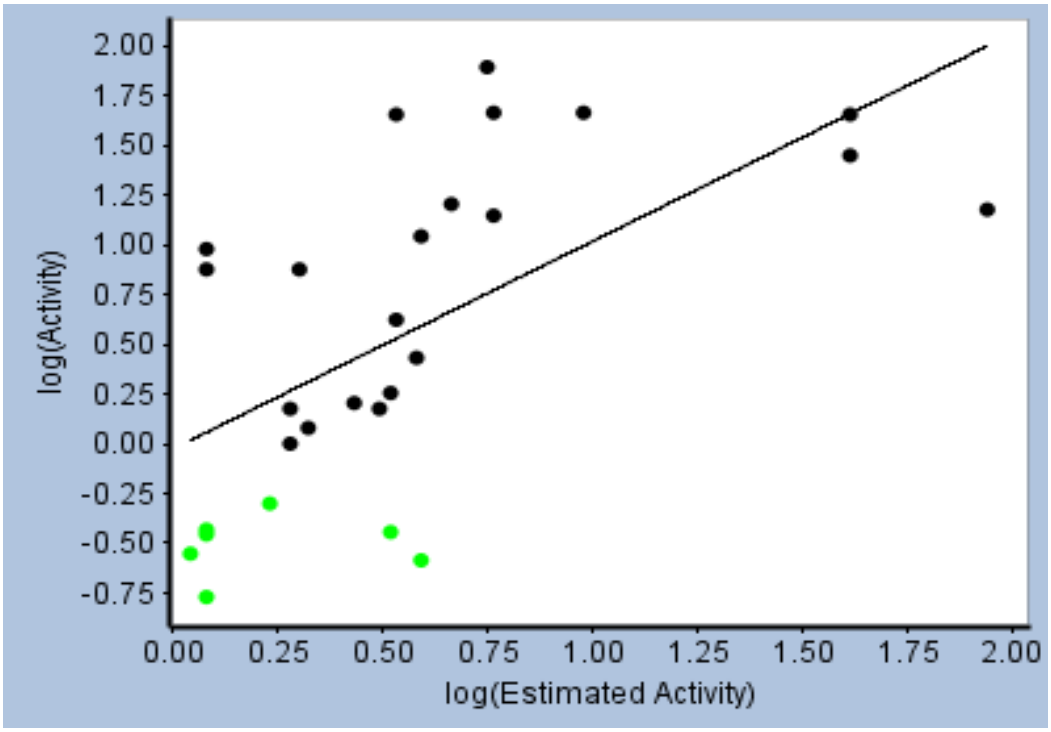


Figure 14: Correlation between experimental activity and estimated activity.

4.1 Pharmacophore validation

During the pharmacophore generation, the HypoGen algorithm in Discovery Studio produces three cost values, such as **Fixed cost**, **Total Cost** and **Null cost**. The model is validated by depending on the cost differences which were calculated by the difference between the null cost and total cost. If the model has the cost difference above 60, it has the predictability chance of greater than 90%. For our hypogen1 is the sign model which has the cost difference 44.046 bits and correlation with 0.782897 and RMSD value of 1.68205 bits (Shown in Table 3).

Table 3: Statistical results of the top 10 pharmacophore hypotheses generated by HypoGen algorithm

Hypo. NO.	Total cost	Cost Difference	RMSD	Correlation	Max. fit	Features
1	138.219	44.046	1.68205	0.782897	8.09674	HBA HBA RING_AROMATIC
2	157.357	24.908	2.09247	0.626982	5.54113	HBA HBA RING_AROMATIC
3	159.782	22.483	2.13121	0.608921	6.65645	HBA RING_AROMATIC RING_AROMATIC
4	159.872	22.393	2.08236	0.632681	4.92761	HBA HBA HBA HYDROPHOBIC
5	159.887	22.378	2.08163	0.633019	4.90049	HBA HBA HBA HYDROPHOBIC
6	160.776	21.489	2.08892	0.629741	4.68966	HBA HBA HBA HYDROPHOBIC
7	161.431	20.834	2.10206	0.62366	4.74114	HBA HBA HBD HYDROPHOBIC
8	161.833	20.432	2.12108	0.614602	5.08245	HBA HBD HYDROPHOBIC RING_AROMATIC
9	161.837	20.428	2.11673	0.61671	4.95438	HBA HBA HYDROPHOBIC RING_AROMATIC

10	161.848	20.417	2.12222	0.614048	5.10921	HBA HBA HBA HYDROPHOBIC

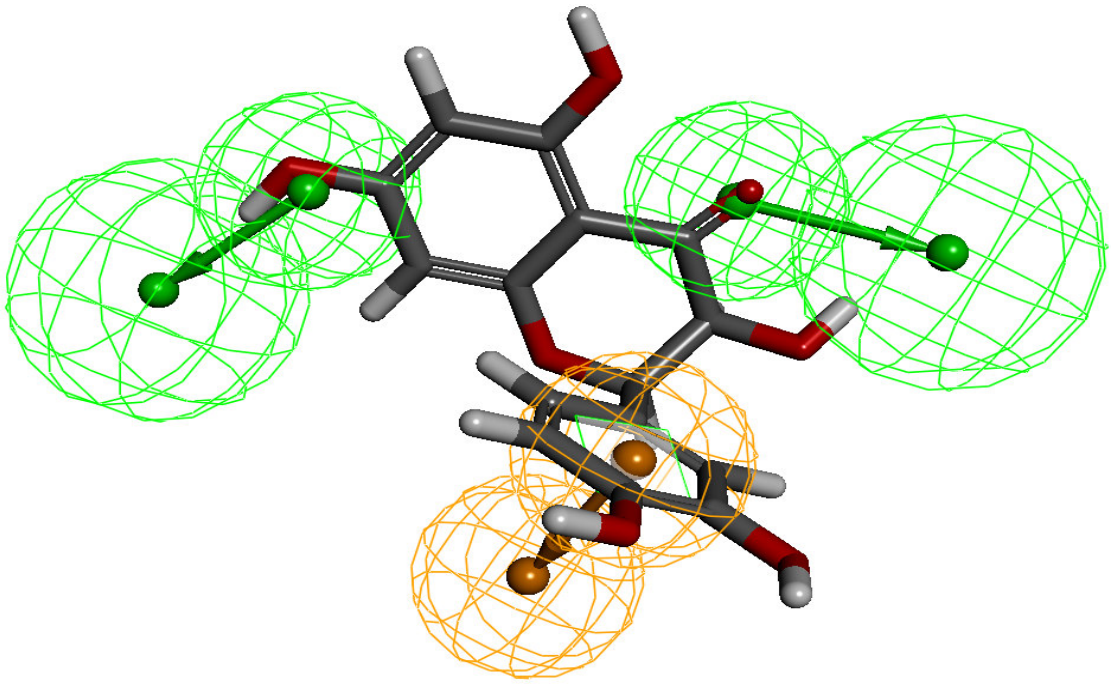
(Null cost- 182.265, *Fixed cost- 95.9307, * Best record in pass- 5, *Pharmacophore space-1.175103e+004

(RA- RING_AROMATIC, HBA-Hydrogen bond acceptor, HYD- HYDROPHOBIC)

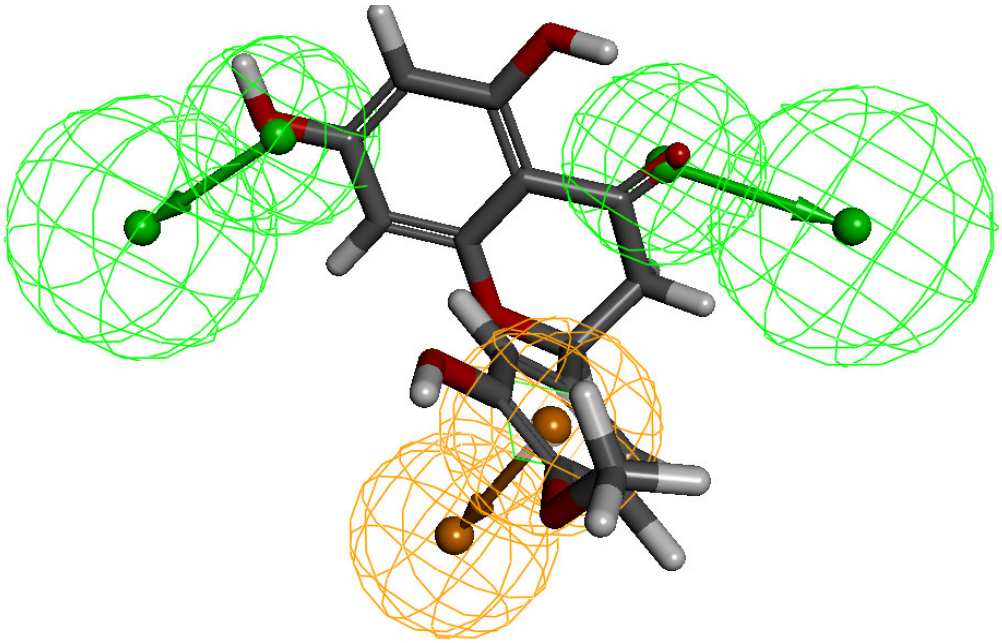
4.2 Test set analysis

The activities of the training set compounds and test set compounds were predicted by the good Pharmacophore Model. Test sets were consisted of 15 flavonoids compounds with different activities range. To check the predictability power of the pharmacophore model Ligand Pharmacophore mapping protocol with flexible search option was used to map the test set compounds. In test set analysis, for most of the compounds were beautifully mapped. 15 compounds presented on test set were beautifully mapped on the pharmacophore. The maximum fit score was 7.55319 obtained for Taxifolin.

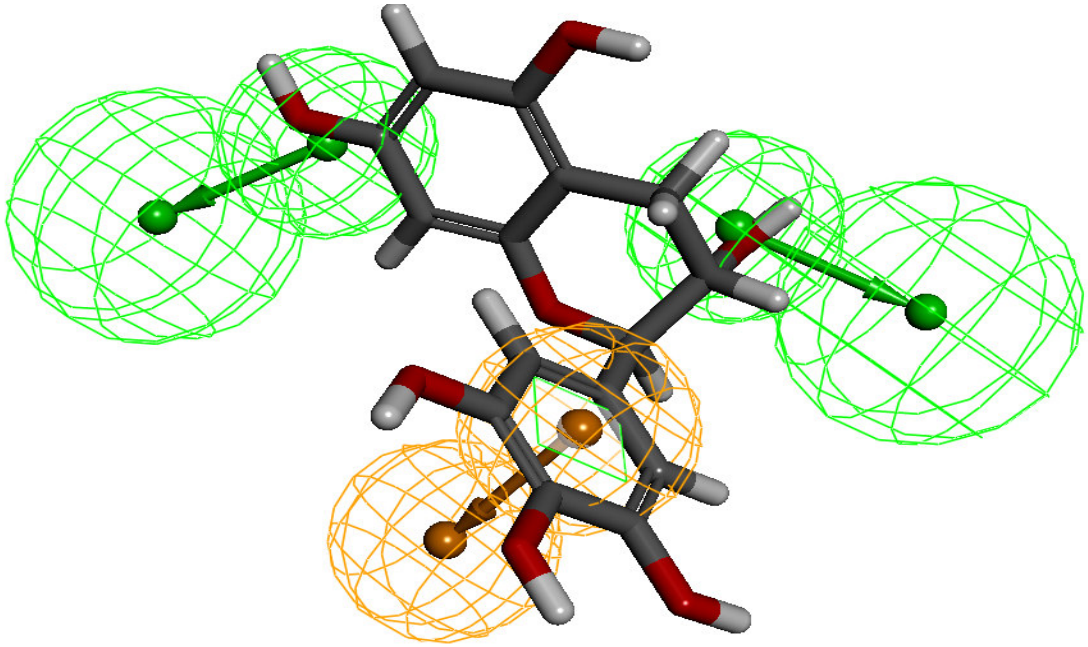
(A)



(B)



(C)



(D)

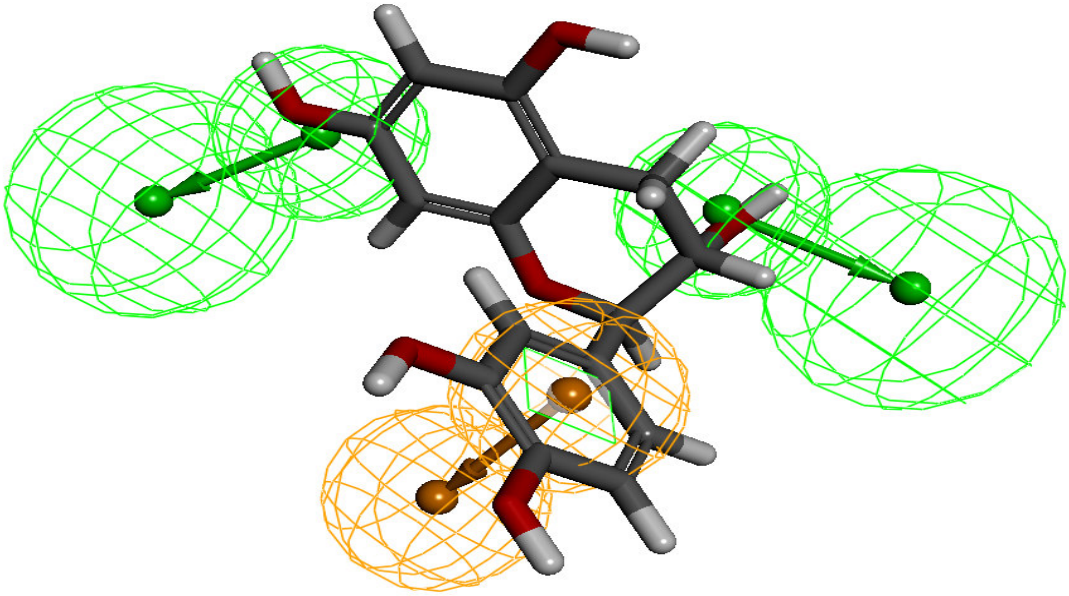


Figure 15: Pharmacophore mapping of the compounds in the Test set along with their fit score. (A) Hypo 1 mapped on to the Taxifolin and fit score is 7.55319, (B) Hypo 1 mapped on Hesperetin and fit score is 5.68148, (C) Hypo 1 mapped on to Catechin and fit score is 7.30873. (d) Hypo1 mapped on to Gallocatechol with a fit score 7.30836. The features are color coded with Green, hydrogen bond acceptor; Brown Ring aromatic features.

4.2.1 Fischer Randomization Test

The Fischer randomization test was the third approach to validate the Hypo1 using DS. In this method the experimental activity of the training set compounds was randomly scrambled and generates the random pharmacophore model using the same parameters as used in developing the Hypo1 hypothesis. A confidence level of 90% was set and it created 9 spreadsheets, all 9 random spreadsheets have high cost values than total cost, and correlation value is less than the Hypo1. It clearly shows none of the randomly generated pharmacophores has good statistical values than Hypo1. The difference in costs between the HypoGen and Fischer randomizations was shown in Figure 16.

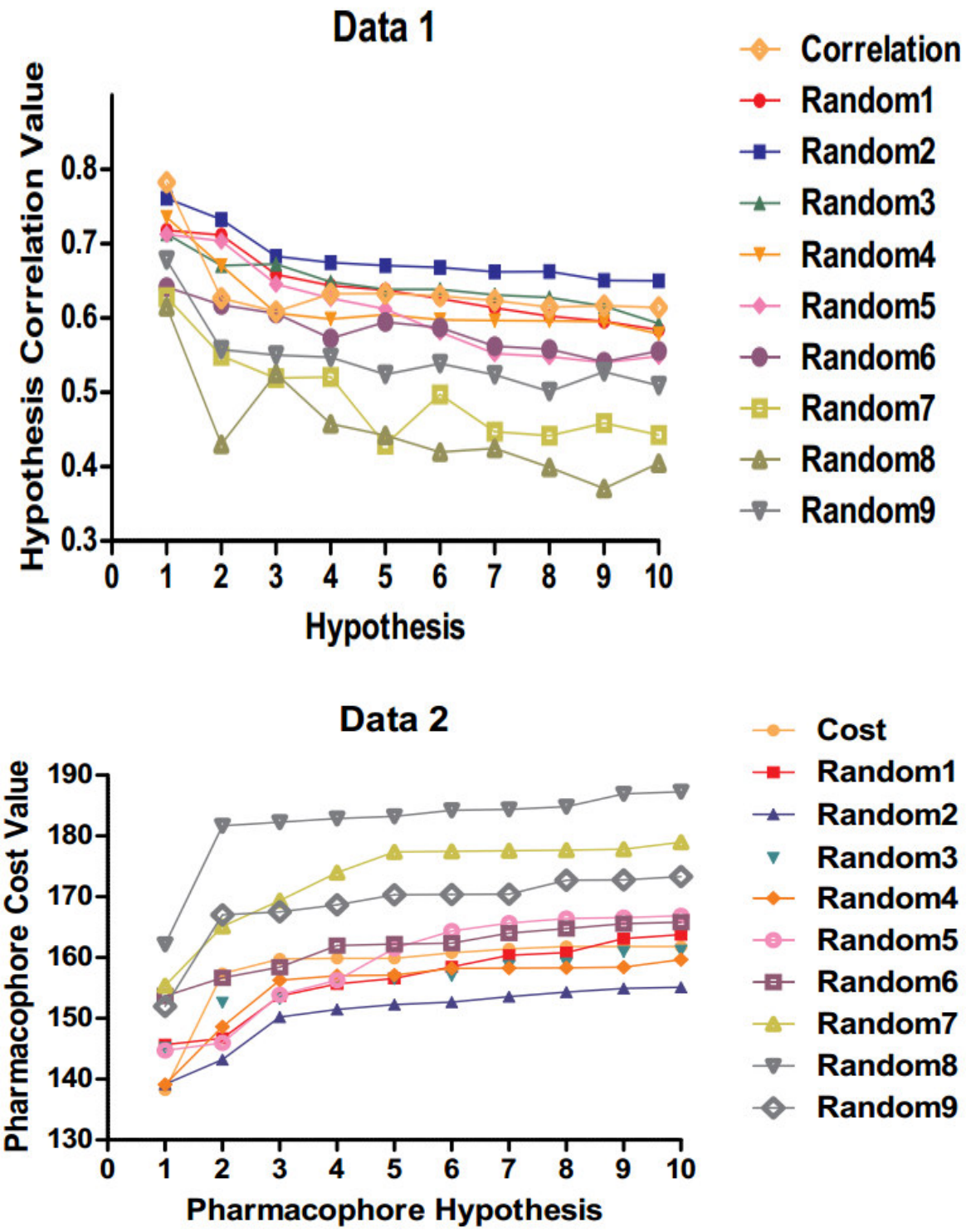


Figure: 16. Data1: The difference in correlation values of hypotheses between Hypo1 spreadsheet and 9 random spreadsheets. Data2: The difference in total cost of hypotheses between the initial hypothesis (Hypo1) spreadsheet and 9 random spreadsheets.

- All the three validation methods demonstrated that Hypo1 hypothesis has good predictability and can be chosen as the best model.

4.2.2 Database screening

The best pharmacophore model, Hypo1, was used as a 3D query to search ZINC chemical database containing totally 17900242 compounds. Search 3D Database protocol with the Best Search option as available in DS was employed to search these databases.The hit compounds were primarily filtered based on the estimated activity less than 0.1 value followed by drug-likeness prediction. Smart/ADMET and Lipinski's rule of five properties was used to select the compounds with favoring drug-like properties. A compound has to obey the following parameters to be predicted as Lipinski-positive:

(i) Molecular weight less than 500

(ii) Number of hydrogen bond donors and acceptors less than 5 and 10, respectively

(iii)Octanol / water partition coefficient value less than 5.

Along with the training set compounds 250 drug-like compound were subjected to molecular docking study. Database screening and drug-likeness prediction steps are shown in detail in Figure 17.

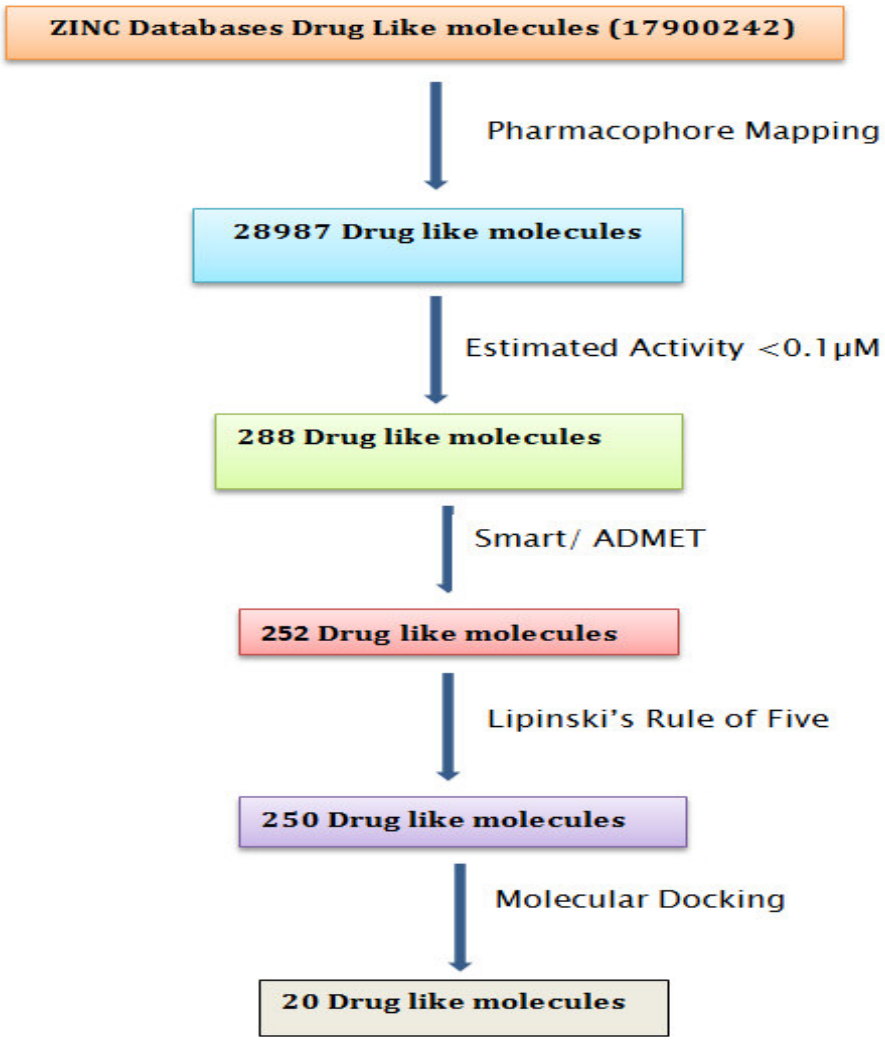


Figure 17: Schematic representation of the virtual screening process implemented in the identification of Glucagon inhibitors

4.2.3 - Molecular Docking

The drug-like hit compounds were docked into the active site of Glucagon receptors (PDB ID 4L6R). The active site was defined based on the bound inhibitor, PEG1200 and previous published journal (Raj et al., EnzEng 2015, 4.2). Arg346, Leu399, Tyr 138, Gln 142, Tyr 149, Val 191, Gln 232, Trp 295, Thr 296, Asn 298, Glu 362, Phe 365 and Leu 386 residues are interacted with the active sites of the glucagon proteins. These residues are very important for binding in protein active site. The docking was performed using CDocker protocol in Discovery Studio 4.0. Molecules were ranked according to their binding energy.The binding modes, molecular interactions with the active site components and binding energy scores

were considered as important components in selecting the best poses of the docked compounds.

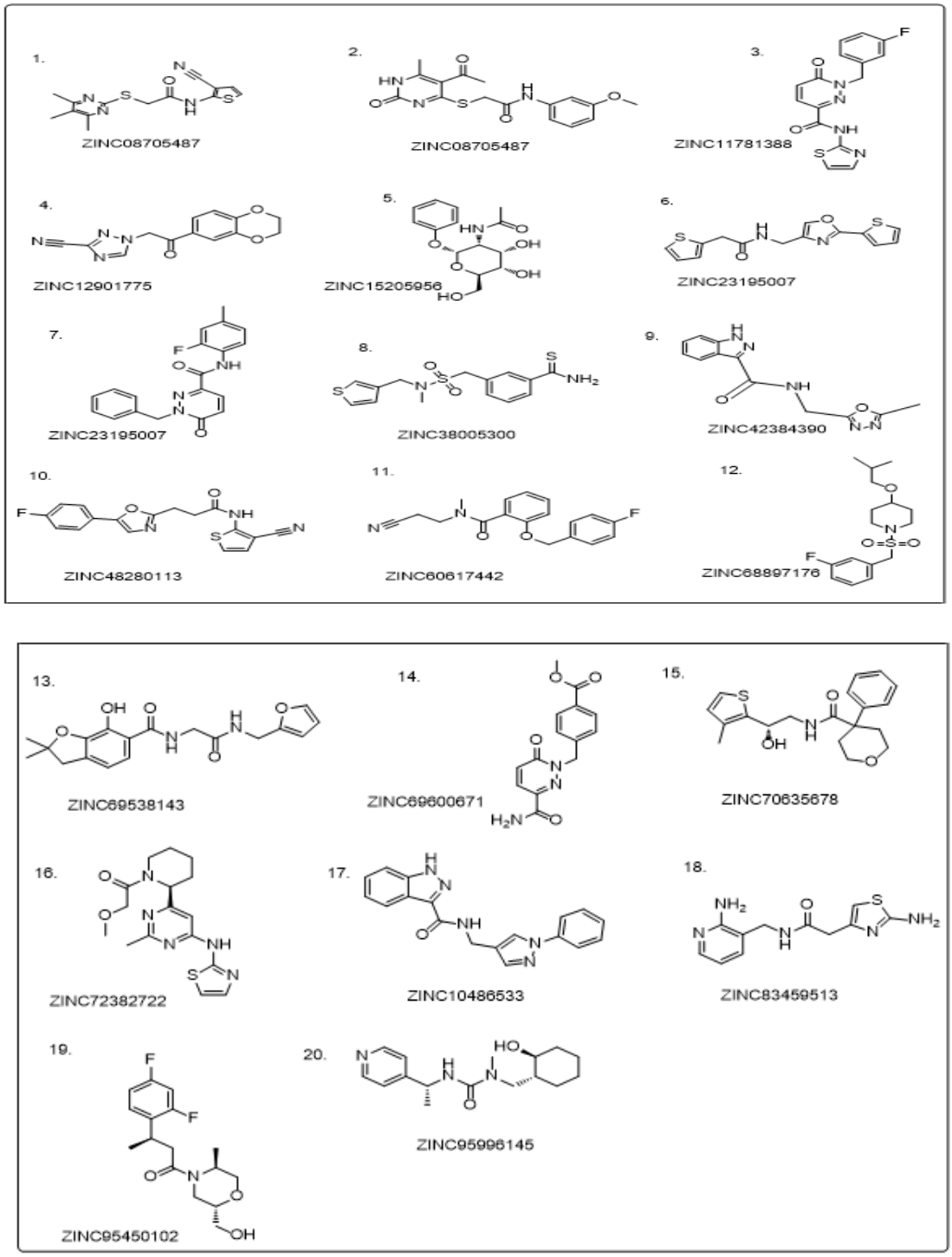
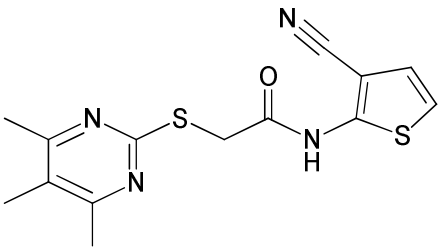
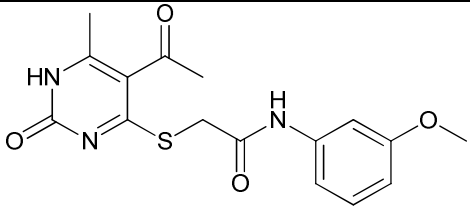
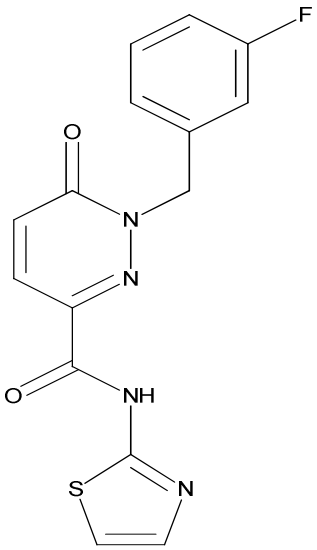
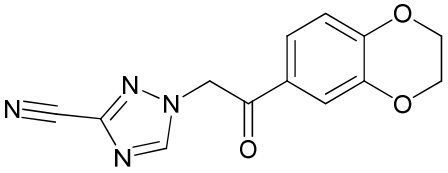
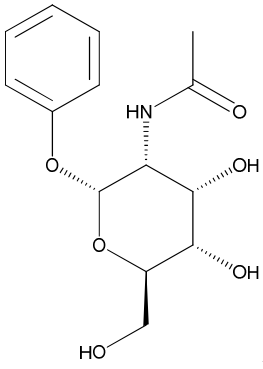
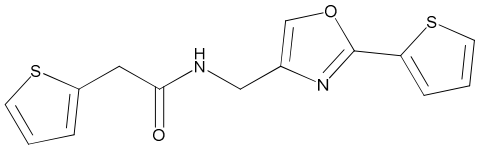
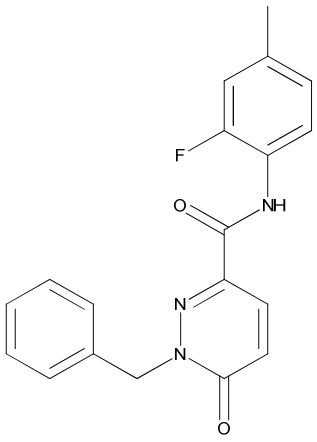


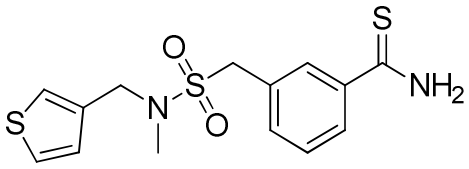
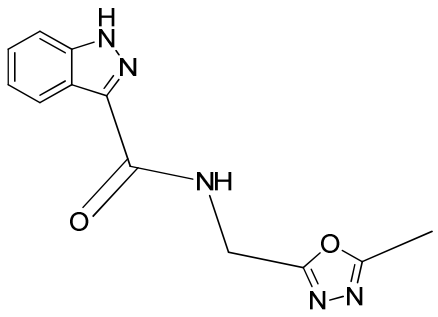
Figure 18: Identified lead compounds through NCI and Asinex database search**.**

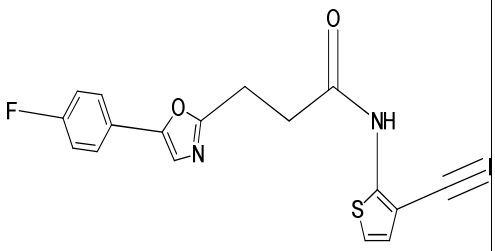
Table 4: The binding analysis of Ligand on binding with active site cavity of Glucagon receptor

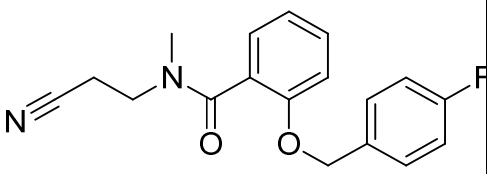
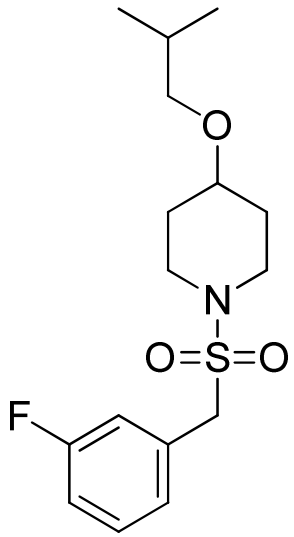
SL.NO.	STRUCTURE	AMINO ACID	HYDROGEN BOND DISTANCE (Å)	CDOCKER ENERGY
1	 ZINC08705487	ZINC08705487:N13 - ZINC08705487:N4 ZINC08705487:S9 - A:PHE365 ZINC08705487 - A:PHE303 ZINC08705487:C8 - A:VAL191 ZINC08705487:C8 - A:LEU386 A:TRP295 - ZINC08705487:C21 A:TRP295 - ZINC08705487:C21 ZINC08705487 - A:LEU386	3.32477 4.70843 5.05593 3.60804 4.75405 4.51199 4.52091 5.26091	-26.0358
2	 ZINC10913777	A:ASN298:N - ZINC10913777:O24 A:LEU386:CD2 - ZINC10913777	2.92135 3.81157	-11.6718

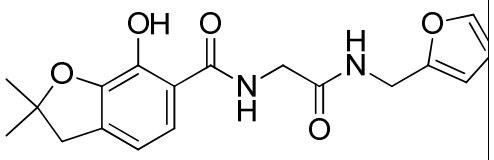
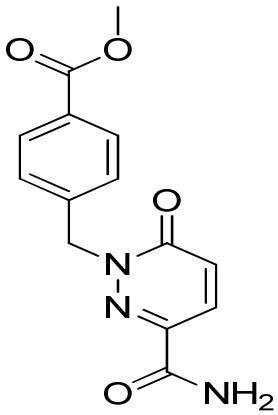
3	 ZINC11781388	A:GLN232:HE21 - ZINC11781388:O17 A:ASN298:HN - ZINC11781388:O8 A:ARG308:HH21 - ZINC11781388:F1 A:PHE365:O - ZINC11781388:F1 A:PHE365 - ZINC11781388 A:PHE303 - ZINC11781388 ZINC11781388 - A:MET231	1.99895 2.61635 2.15933 3.43798 5.13705 5.6241 5.21188	-20.3729
4	 ZINC12901775	A:ASN298:N - ZINC12901775:O12 A:ARG378:NH1 - ZINC12901775:O1 A:ARG378:CD - ZINC12901775:O1 ZINC12901775:C11 - A:THR296:O A:TRP295 - ZINC12901775	3.04113 2.76296 3.26722 3.56805 5.87666	+16.1257
5	 ZINC15205956	A:ASN298:N - ZINC15205956:O20 A:ARG378:NH1 - ZINC15205956:O10 A:ARG378:NH1 - ZINC15205956:O11 ZINC15205956 - A:LEU382	3.28225 2.86833 3.1785 5.18165	-5.99497

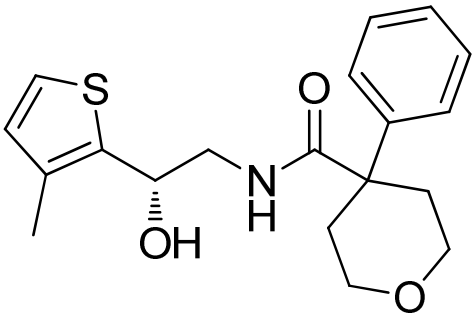
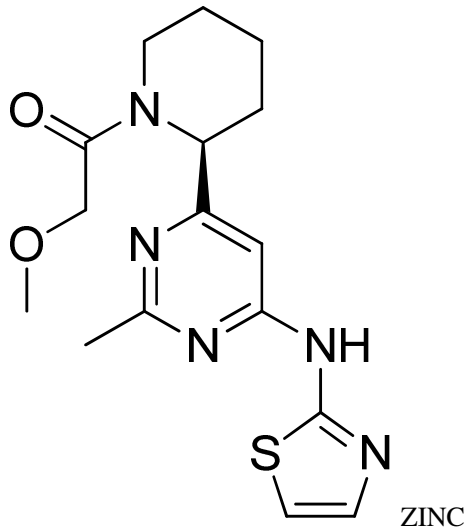
6	 ZINC23195007	A:SER389:OG - ZINC23195007:O15 A:LYS187:NZ - ZINC23195007:S16 A:SER389:OG - ZINC23195007:S16 A:GLU362:OE2 - ZINC23195007 A:ILE235:CG2 - ZINC23195007 A:ILE235:CD1 - ZINC23195007 A:TRP295 - ZINC23195007 ZINC23195007 - A:TRP295 ZINC23195007 - A:PHE365 ZINC23195007 - A:ILE235	2.85827 2.88506 3.05737 3.41606 3.94449 3.76355 5.40148 5.48057 4.58524 5.33836	-26.3469
7	 ZINC36798750	A:GLN232:NE2 - ZINC36798750:O11 A:TRP295 - ZINC36798750 A:TRP295 - ZINC36798750 A:PHE303 - ZINC36798750 ZINC36798750 - A:VAL191 ZINC36798750 - A:MET231 ZINC36798750 - A:LEU386	2.88473 5.59897 5.72587 5.17666 5.36734 4.80914 5.36265	-23.6019

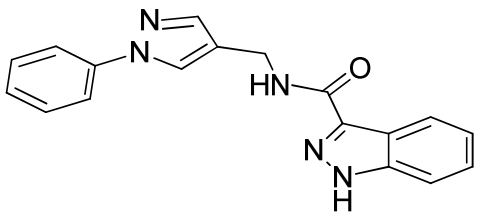
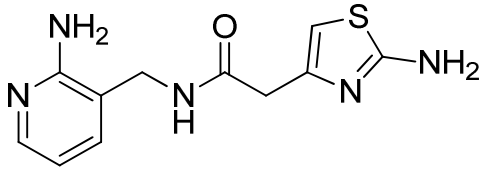
8	 ZINC38005300	A:ASP195:N - ZINC38005300:S20 A:TYR145:OH - ZINC38005300:S20 A:SER389:OG - ZINC38005300:S7 ZINC38005300:S7 - A:TYR149:OH A:ILE235:CD1 - ZINC38005300 ZINC38005300:S7 - A:TYR149 ZINC38005300:S20 - A:TYR145 ZINC38005300 - A:VAL191 ZINC38005300 - A:MET231 ZINC38005300 - A:LEU386	3.42306 3.77303 2.95383 3.04705 3.92847 5.72977 5.71527 5.33754 5.26108 5.4274	-26.055
9	 ZINC42384390	A:GLN232:NE2 - ZINC42384390:O10 ZINC42384390:N18 - A:PHE303 A:MET231:CE - ZINC42384390 ZINC42384390 - A:VAL191	2.93966 3.45227 3.79218 5.07288	+18.5768

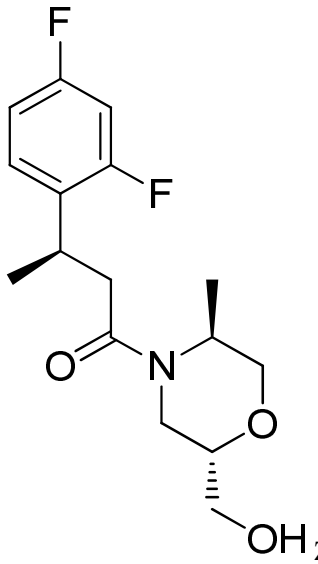
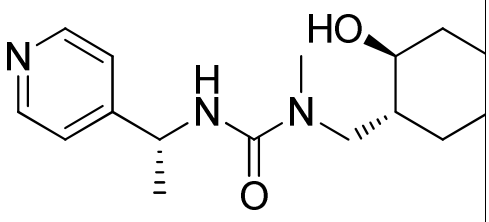
10	 ZINC48280113	A:GLN232:HE21 - ZINC48280113:O1 A:GLN232:HE21 - ZINC48280113:S22 A:TRP295 - ZINC48280113 A:TRP295 - ZINC48280113 ZINC48280113 - A:PHE303 ZINC48280113 - A:VAL191 ZINC48280113 - A:MET231	2.23666 2.64579 4.57418 4.40377 5.45359 4.88527 5.03918	-17.6489

11	 ZINC60617442	A:ARG308:NH2 - ZINC60617442:F23 A:ARG378:NH1 - ZINC60617442:O8 A:ARG378:CD - ZINC60617442:O8 A:PHE365:O - ZINC60617442:F23 A:TRP295 - ZINC60617442 ZINC60617442:C5 - A:LEU382 A:TRP295 - ZINC60617442:C5	3.3951 5.5226 4.52088 5.01044 3.08366 2.86651 2.9167	-7.94317
12	 ZINC68897176	A:ARG308:NH2 - ZINC68897176:F22 A:ARG378:NH1 - ZINC68897176:O5 A:ARG378:NH1 - ZINC68897176:O14 A:ARG378:CD - ZINC68897176:O14 A:PHE365:O - ZINC68897176:F22	3.10815 3.36849 3.09405 3.64303 3.61446	-15.8082

13	 ZINC69538143	A:GLN232:NE2 - ZINC69538143:O12 ZINC69538143:C15 - A:GLN232:OE1 A:MET231:CE - ZINC69538143 ZINC69538143:C1 - A:TRP295 A:SER297:C,O;ASN298:N - ZINC69538143 ZINC69538143:C25 - A:LEU386	3.02044 3.69613 3.95303 3.8655 4.34612 5.16184	-21.29
14	 ZINC69600671	A:LYS187:HZ3 - ZINC69600671:O4 A:GLN232:HE22 - ZINC69600671:O20 A:SER389:HG - ZINC69600671:O4 ZINC69600671:N21 - A:THR296:O A:LYS187:HE1 - ZINC69600671:O4 ZINC69600671:C1 - A:GLU362:OE2 ZINC69600671:C1 - A:PHE365 A:MET231:SD - ZINC69600671 ZINC69600671 - A:MET231	2.2948 2.73714 2.07601 2.76744 2.34426 3.25003 3.94804 5.93548 5.25812 5.25875	-27.0392

		ZINC69600671 - A:ILE235		
15	 ZINC70635678	A:GLN232:NE2 - ZINC70635678:O11 A:GLN232:NE2 - ZINC70635678:O24 ZINC70635678:C8 - A:THR296:O ZINC70635678:O24 - A:PHE303 ZINC70635678:S5 - A:PHE303 A:TRP295 - ZINC70635678 A:TRP295 - ZINC70635678 ZINC70635678 - A:PHE303 A:LEU386 - ZINC70635678	3.37699 3.22805 3.3335 3.65731 3.67375 4.43505 4.34387 5.00675 5.26375	-22.8173
16	 ZINC 72382722	A:GLN232:NE2 - ZINC72382722:O21 ZINC72382722:C22 - A:THR296:O A:GLU362:OE2 - ZINC72382722 A:ILE235:CD1 - ZINC72382722 A:LEU386:CD2 - ZINC72382722 A:MET231:SD - ZINC72382722 ZINC72382722 - A:PHE365	3.15713 3.11704 3.40616 3.71009 3.74283 5.45416 4.66075	-29.2679

17	 ZINC10486533	A:ASN298:N - ZINC10486533:N24 A:ARG378:NH1 - ZINC10486533:N15 A:ARG378:NH1 - ZINC10486533 A:TRP295 - ZINC10486533 ZINC10486533 - A:PHE303 ZINC10486533 - A:ARG378	3.2996 3.20495 3.62902 5.21726 5.64317 5.09182	+25.4604
18	 ZINC83459513	A:GLN232:NE2 - ZINC83459513:N7 ZINC83459513:N1 - ZINC83459513:O11 ZINC83459513:N18 - A:TYR145:OH ZINC83459513:C6 - A:THR296:O ZINC83459513 - ZINC83459513 A:TRP295 - ZINC83459513 ZINC83459513 - A:ILE194	3.16315 3.2365 3.0392 3.61152 5.73886 5.90416 5.44808	-23.8631

19	 ZINC95450102	A:SER389:OG - ZINC95450102:O4 A:SER389:OG - ZINC95450102:O22 ZINC95450102:O22 - A:GLU362:OE2 A:GLN232:NE2 - ZINC95450102:O9 ZINC95450102:C1 - ZINC95450102 A:VAL191 - ZINC95450102 ZINC95450102 - A:MET231	3.0008 2.94112 2.71012 2.92807 3.9372 5.44245 5.34087	-20.8769
20	 ZINC95996145	A:ARG308:NH2 - ZINC95996145:N6 A:ARG378:NH1 - ZINC95996145:O21 A:TRP295 - ZINC95996145	3.02968 2.92206 5.26878	-16.7823

After analyzing the docking results four molecules were found to be more active at binding side in Glucagon Protein. **ZINC72382722** having binding energy -29.2679 kcal\mol, Interacting residues is GLN232, THR296, GLU362, ILE235, LEU386, MET231, PHE365 which is most important for binding in glucagon protein. **ZINC69600671** having binding energy -27.0392 kcal\mol, interacting residues is LYS187, GLN232, SER389, THR296, LYS187, GLU362, PHE365, MET231, ILE235. **ZINC23195007** has binding energy -26.3469 kcal\mol, with interacting residues SER389, LYS187, SER389, GLU362, ILE235, TRP295, and PHE365. **ZINC38005300** has binding energy is -26.055 kcal\mol, Interacting residues ASP195, TYR145, SER389, TYR149, ILE235, TYR149, TYR145, VAL195, MET231, LEU386. Binding orientation of compound in 3D and 2D is given below.

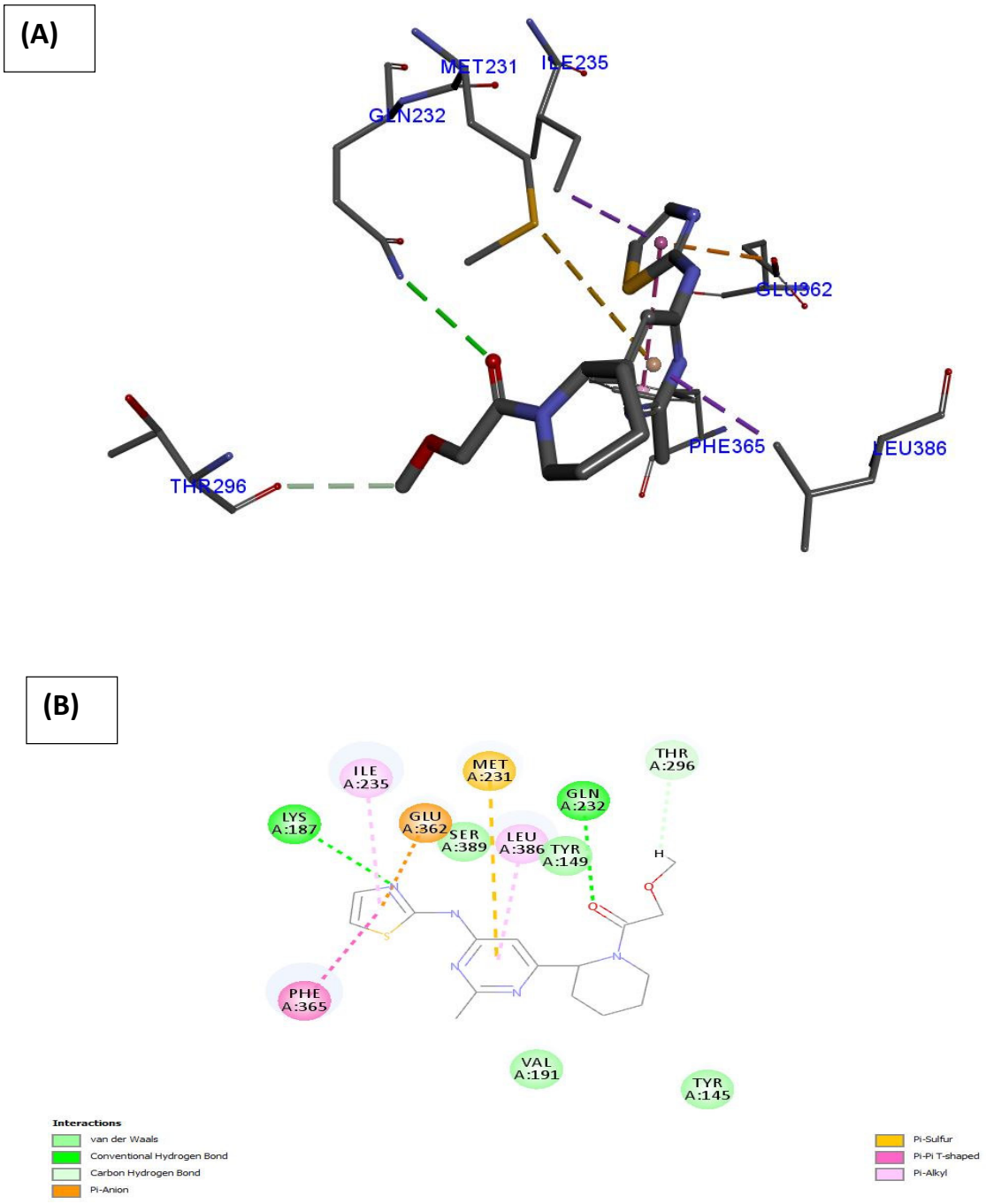


Figure 19: Binding orientation of **ZINC72382722** with glucagon protein (4L6R). **(A)** Binding orientation in 3D, Blue color residue represent the important interacting residue, Green dotted line represent the hydrogen bond distance and other dotted line represent the non-hydrogen bond distance. **(B)** Binding orientation in 2D Green color residue represent the important residues for interaction, green color dotted line re represent the hydrogen bond

distance and other dotted line represent the non-hydrogen bond distance, Binding energy -29.2679 kcal\mol.

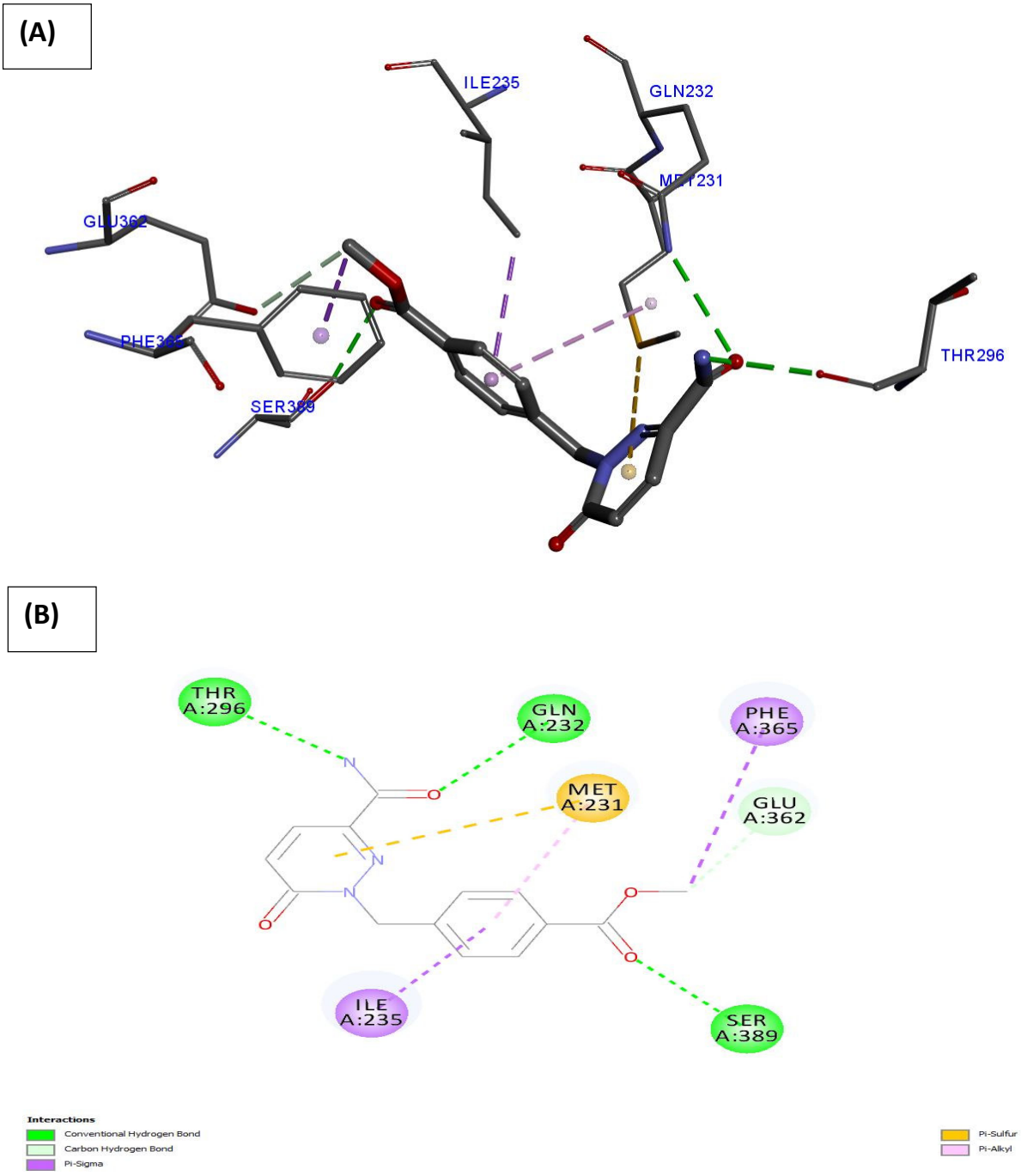


Figure 20: Binding orientation of**ZINC69600671** with Glucagon (4L6R). **(A)** Binding orientation in 3D, Blue color residue represent the important interacting residue, Green dotted line represent the hydrogen bond distance and other dotted line represent the non-hydrogen bond distance. **(B)** Binding orientation in 2D Green color residue represent the important

residues for interaction, green color dotted line re represent the hydrogen bond distance and other dotted line represent the non-hydrogen bond distance, Binding energy -27.0392 kcal\mol.

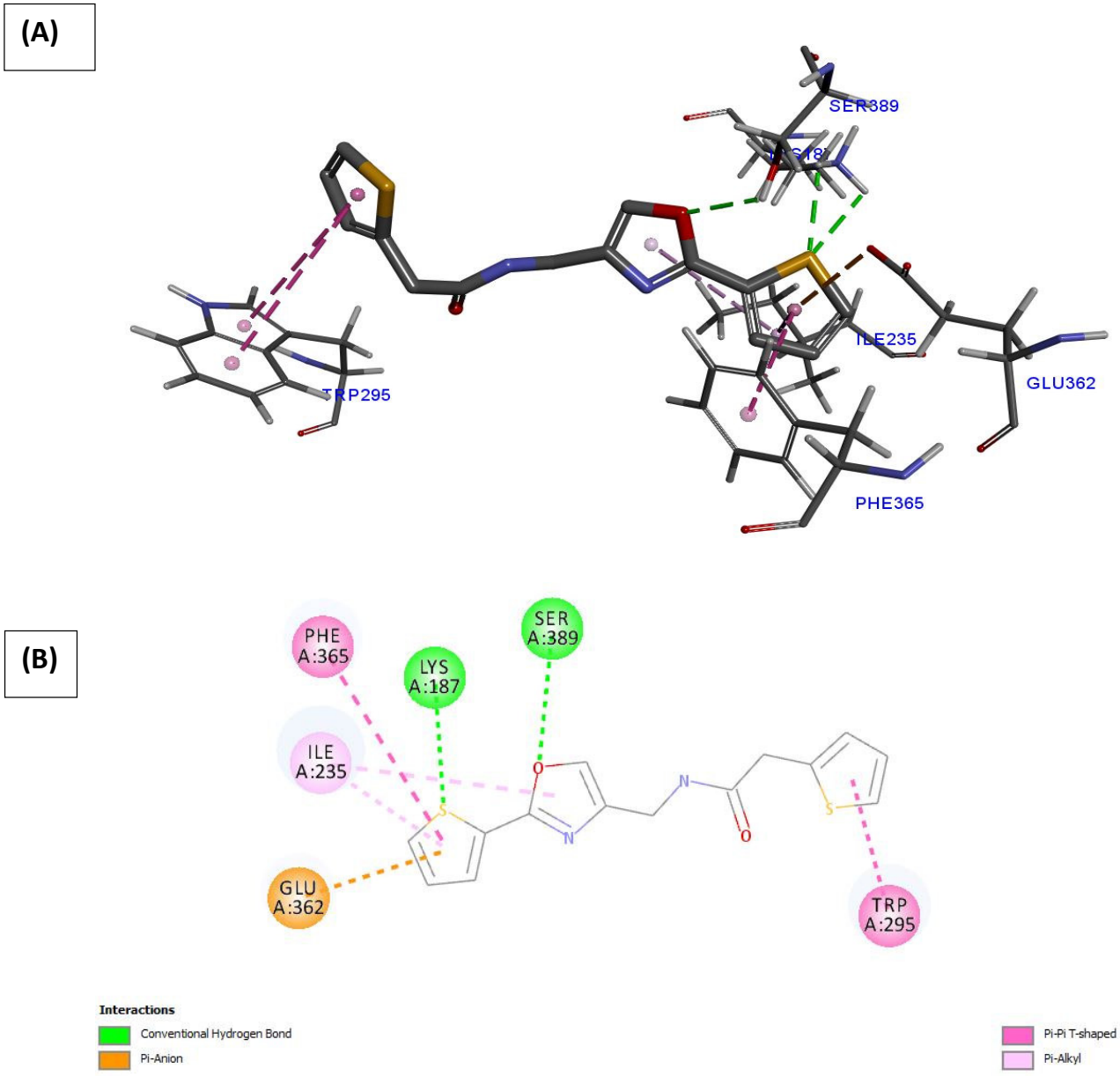


Figure 21: Binding orientation of **ZINC23195007** with Glucagon protein (4L6R). **(A)** Binding orientation in 3D, Blue color residue represent the important interacting residue represent the others interacting residues, Green dotted line represent the hydrogen bond distance and other dotted line represent the non-hydrogen bond distance. **(B)** Binding orientation in 2D Green color residue represent the important residues for interaction, green

color dotted line re represent the hydrogen bond distance and other dotted line represent the non-hydrogen bond distance, Binding energy -26.3469 kcal\mol.

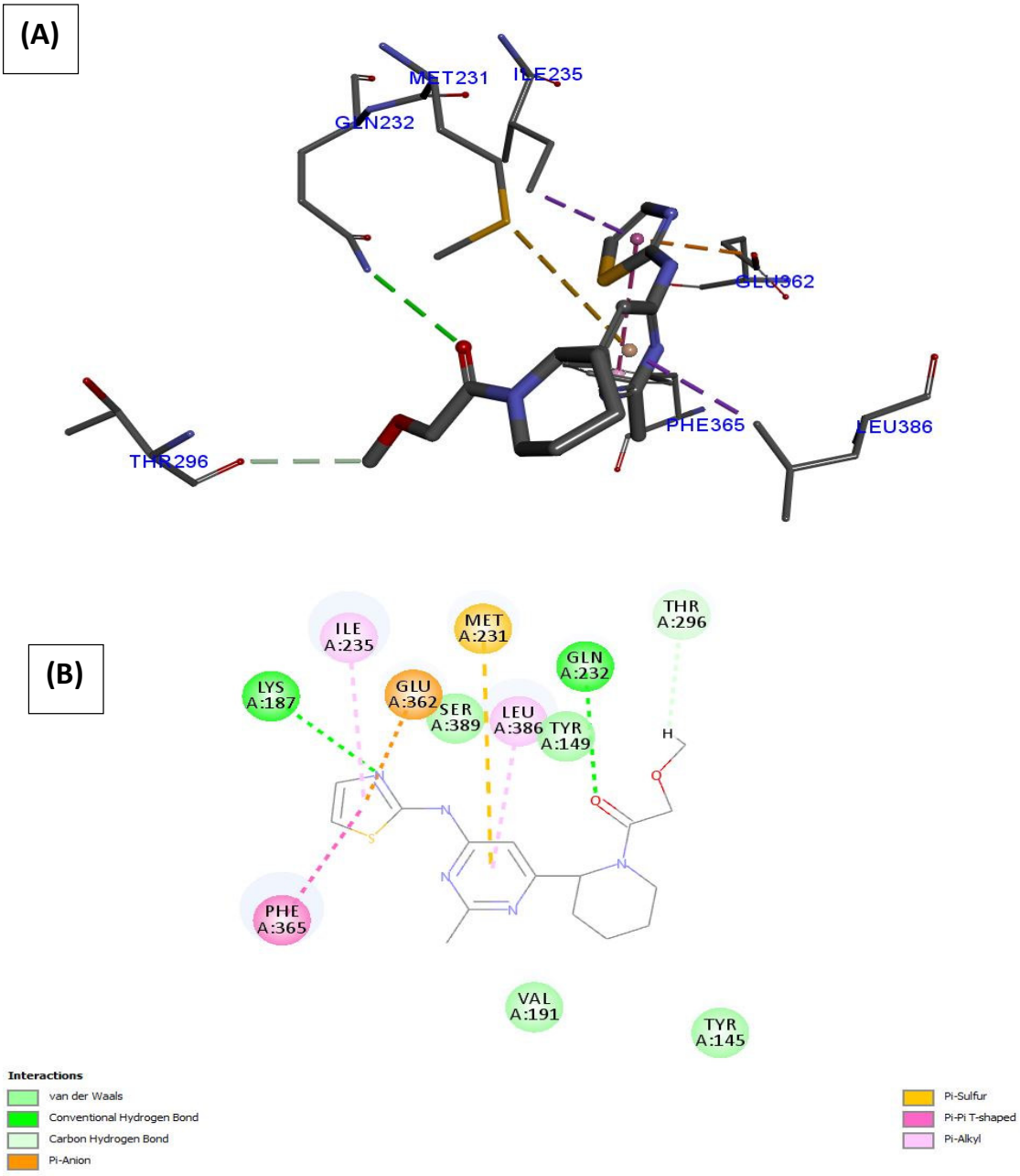


Figure 22: Binding orientation of **ZINC38005300** with Glucagon protein (4L6R). **(A)** Binding orientation in 3D, Blue color residue represent the important interacting residue represent the others interacting residues, Green dotted line represent the hydrogen bond distance and other dotted line represent the non-hydrogen bond distance. **(B)** Binding

orientation in 2D Green color residue represent the important residues for interaction, green color dotted line re represent the hydrogen bond distance and other dotted line represent the non-hydrogen bond distance, Binding energy -26.055 kcal\mol.

For further validation we have overlay the generated pharmacophore on hits compound and found that they overlay properly, each group in hits structure align properly in their respective feature. The results of overlay of pharmacophore on hits compounds are given below.

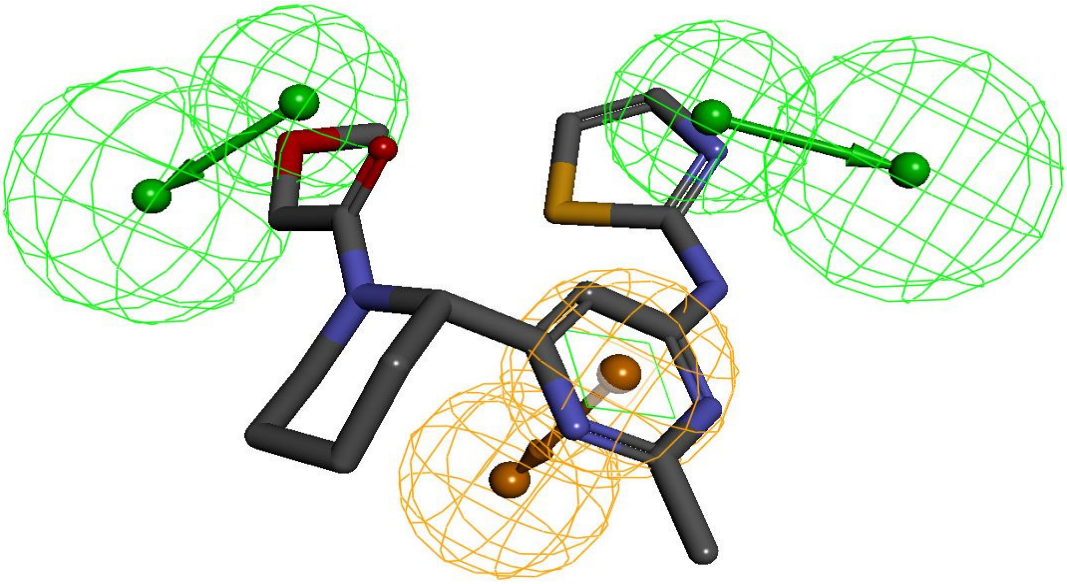


Figure-23: Compound**ZINC72382722** properly overlay on generated pharmacophore

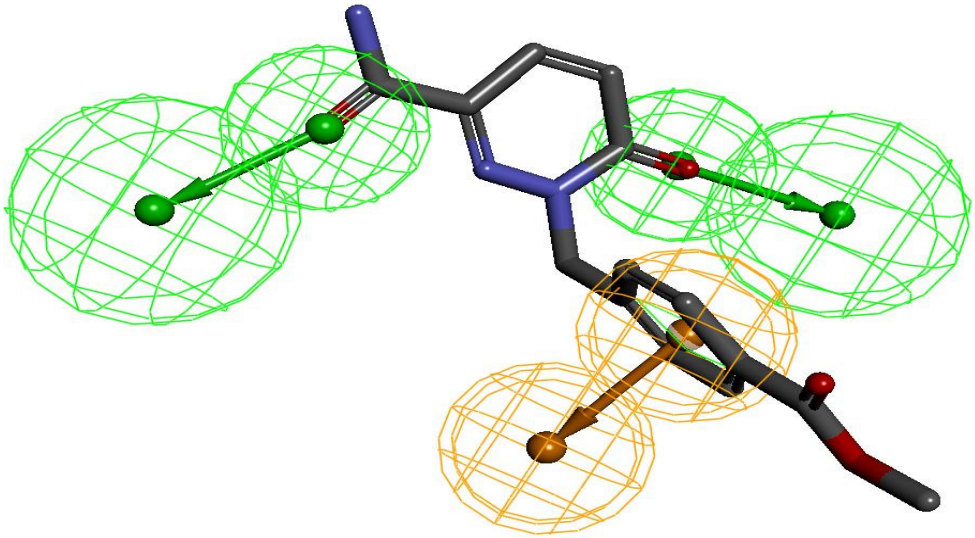


Figure-24: Compound ZINC69600671 properly overlay on generated pharmacophore

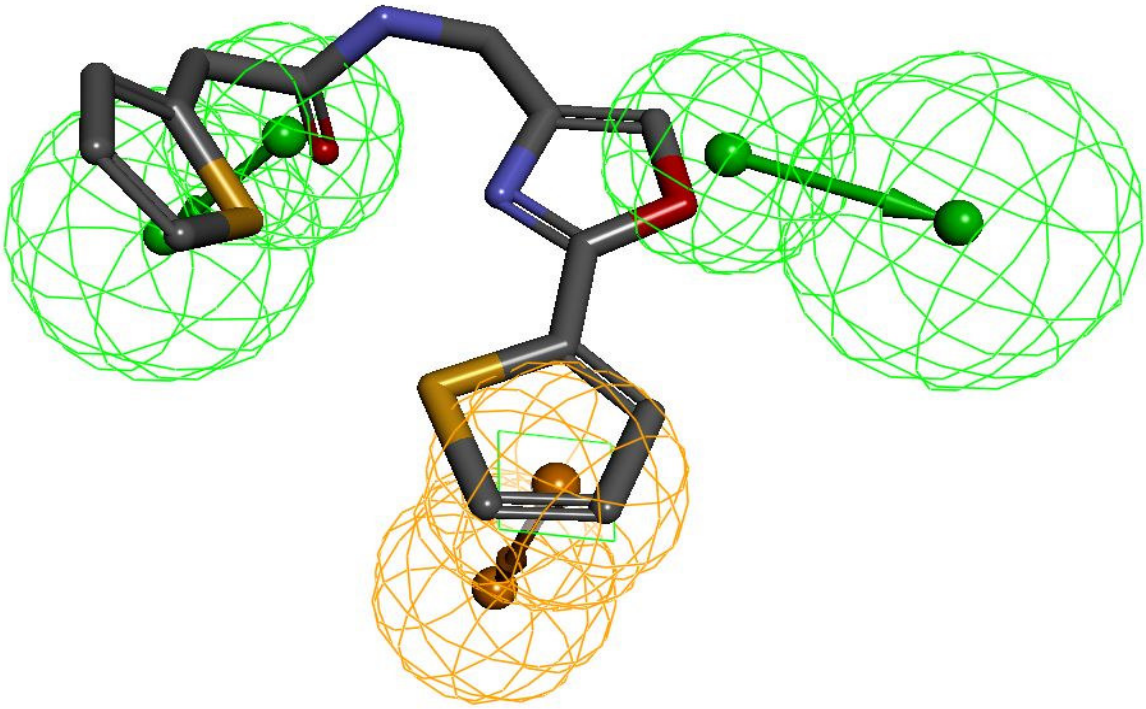


Figure-25: CompoundZINC23195007 properly overlay on generated pharmacophore

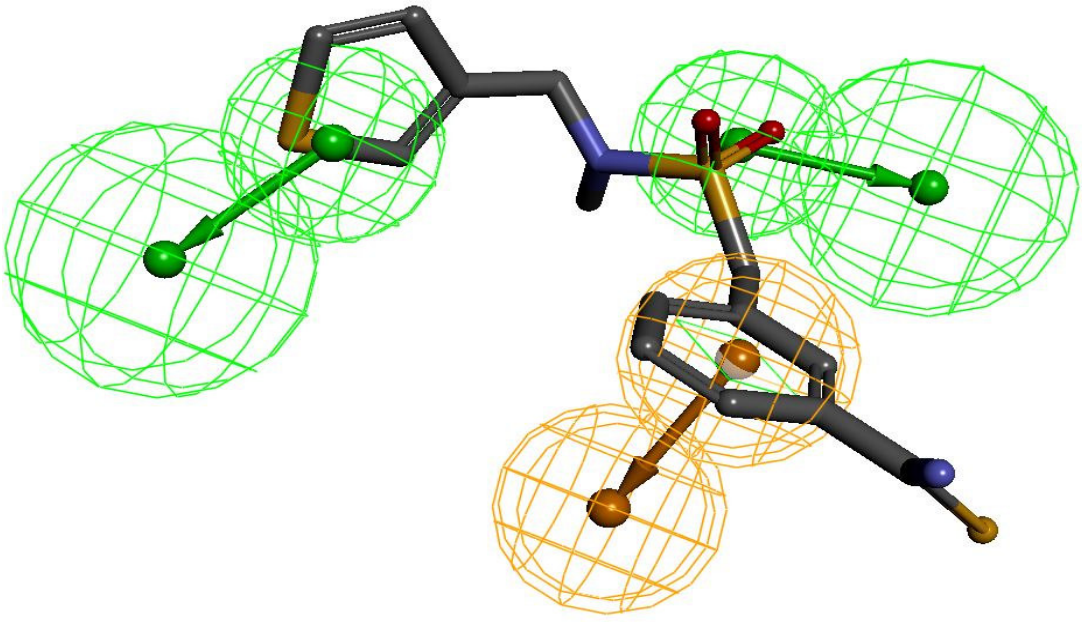


Figure-26: Compound ZINC38005300 properly overlay on generated pharmacophore

CHAPTER-VI

5.SUMMARY

Mastitis remains the most common and the most expensive disease of dairy industry throughout the world. Bovine mastitis disease not only effects the physical, chemical, Bacteriological and organoleptic properties of milk but also effects quality of milk. In winter *E. coli* effects is more. We isolated the genomic DNA of *E. coli* and amplified the 16S r-RNA through PCR. Than sequenced it through ABI instrument and submitted to Gen bank with Accession no. JX122856.We predicted the evolutionary relation through MEGA software. We concluded *E. coli o9* shares a common cluster with *E. coli* strain U5/41 and *Shigella flexneri* strain ATCC 29903.both of which have 99% identity. The other species in the same cluster are Yersinia kristensenii strain ATCC 33638, Serratia marcescens subsp. sakuensis strain KRED and Dickeya zeae strain CFBP 2052. Although the sequence identities of these species are 95%, 96% and 95% respectively they seemed to be evolutionary related with the query sequence.

APPENDIX I

TABLE 5:-MSA OF 18 GENUS

```
E.Coli_16SrRNA                  --------------------------ATGCAGTCG--------------- 9
Escherichia_coli_strain_U_5/41  ---AGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACAT 47
Shigella_flexneri_strain_ATCC_  -------------TGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACAT 37
Enterobacter_cowanii_strain_88  ------------------------------GGCGGCAGGCCTAACACAT 19
Salmonella_bongori_strain_BR_1  ------TTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACAT 44
Cronobacter_muytjensii_strain_  ----------------------------------------CTAACACAT 9
Trabulsiella_guamensis          -------------------------ACGCTGGCG--CAGCCTTACACAT 22
Citrobacter_farmeri_strain_CDC  -------------TGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACAT 37
Pantoea_dispersa_strain_LMG260  --------------------------------------------------
```

```
Pectobacterium_cypripedii_stra      --------------GGCTCAGATTGAACGCTGGCGGCAGGCCTAACACAT 36
Erwinia_papayae_strain_CFBP_11      --------------GGCTCAGATTGAACGCTGGCGGCAGGCCTAACACAT 36
Cedecea_neteri_strain_GTC1717       -AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACAT 49
Serratia_marcescens_subsp._sak      GAGAGTTTGATCCTGGCTCAGATTGAACGCTGGCGGCAGGCTTAACACAT 50
Raoultella_planticola_strain_A      ---------ATCCTGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACAT 41
Klebsiella_oxytoca_strain_ATCC      ---------ATCCTGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACAT 41
Kluyvera_cryocrescens_strain_1      ----------------------------CTGGCGGCAGGCCTAACACAT 21
Buttiauxella_izardii_strain_S3      ---------------GCTCAGATTGAACGCTGGCGGCAGGCCTAACACAT 35
Yersinia_kristensenii_strain_A      --------------------ATTGAACGCTGGCGGCAGGCCTAACACAT 29
Dickeya_zeae_strain_CFBP_2052       ------------------------GAACGCTGGCGGCAGGCCTAACACAT 26


E.Coli_16SrRNA                      ---------ACGGTAACAGGAAGCAGCTTGCTGCTTTGCTGACGAGTGGC 50
Escherichia_coli_strain_U_5/41      GCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGCTTTGCTGACGAGTGGC 97
Shigella_flexneri_strain_ATCC_      GCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGTTTCGCTGACGAGTGGC 87
Enterobacter_cowanii_strain_88      GCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGCTTG-CTGACGAGTGGC 68
Salmonella_bongori_strain_BR_1      GCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGCTTCGCTGACGAGTGGC 94
Cronobacter_muytjensii_strain_      GCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGCTTTGCTGACGAGTGGC 59
Trabulsiella_guamensis              GCAAGTCGAGCGGCAGCGGGGGAAAGCT-GCTTTCCCGCCGGCGAGCGGC 71
Citrobacter_farmeri_strain_CDC      GCAAGTCGAACGGTAACAGGAAGCAGCTTGCTSCTTCGCTGACGAGTGGC 87
Pantoea_dispersa_strain_LMG260      --------------------------------------------------
Pectobacterium_cypripedii_stra      GCAAGTCGAACGGTAGCACAGAGGAGCTTGCTCCTTGGGTGACGAGTGGC 86
Erwinia_papayae_strain_CFBP_11      GCAAGTCGAACGGTAGCACAAGGGAGCTTGCTCCCCGGGTGACGAGTGGC 86
Cedecea_neteri_strain_GTC1717       GCAAGTCGAGCGGTAGCACGGGAGAGCTTGCTCTCTGCTTGACGAGCGGC 99
Serratia_marcescens_subsp._sak      GCAAGTCGAGCGGTAGCACAGGGGAGCTTGCTCCCTGGGTGACGAGCGGC 100
Raoultella_planticola_strain_A      GCAAGTCGAGCGGTAGCACAG-AGAGCTTGCTCTCGGG-TGACGAGCGGC 89
Klebsiella_oxytoca_strain_ATCC      GCAAGTCGAACGGTAGCACAG-AGAGCTTGCTCTCGGG-TGACGAGTGGC 89
Kluyvera_cryocrescens_strain_1      GCAAGTCGAACGGTAGCACAG-AGAGCTTGCTCTTGGG-TGACGAGTGGC 69
Buttiauxella_izardii_strain_S3      GCAAGTCGAGCGGTAGCACAG-GGAGCTTGCTCCTGGG-TGACGAGCGGC 83
Yersinia_kristensenii_strain_A      GCAAGTCGAGCGGCAGCGGGAAGTAGTTTACTACTTTGCCGGCGAGCGGC 79
Dickeya_zeae_strain_CFBP_2052       GCAAGTCGGGCGGTAGCACAAGGGAGCTTGCTCCCTGGGTGACGAGCGGC 76


E.Coli_16SrRNA                      GGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC 100
Escherichia_coli_strain_U_5/41      GGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC 147
Shigella_flexneri_strain_ATCC_      GGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC 137
Enterobacter_cowanii_strain_88      GGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC 118
Salmonella_bongori_strain_BR_1      GGACGGGTGAGTAATGTCTGGGAAACTGCCTGGTGGAGGGGGATAACTAC 144
Cronobacter_muytjensii_strain_      GGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC 109
Trabulsiella_guamensis              GGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC 121
Citrobacter_farmeri_strain_CDC      GGACGGGTGAGTAATGTCTGGGAAACTGCCCGATGGAGGGGGATAACTAC 137
Pantoea_dispersa_strain_LMG260      GGACGGGTGAGTAATGTCTGGGAAACTGCCCGATGGAGGGGGATAACTAC 50
Pectobacterium_cypripedii_stra      GGACGGGTGAGTAATGTCTGGGGATCTGCCCGATGGAGGGGGATAACTAC 136
Erwinia_papayae_strain_CFBP_11      GGACGGGTGAGTAATGTCTGGGAAACTGCCCGATGGAGGGGGATAACTAC 136
Cedecea_neteri_strain_GTC1717       GGACGGGTGAGTAATGTCTGGGGATCTGCCTGATGGAGGGGGATAACTAC 149
Serratia_marcescens_subsp._sak      GGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC 150
Raoultella_planticola_strain_A      GGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC 139
Klebsiella_oxytoca_strain_ATCC      GGACGGGTGAGTAATGTCTGGGAAACTGCCCGATGGAGGGGGATAACTAC 139
Kluyvera_cryocrescens_strain_1      GGACGGGTGAGTAATGTCTGGGAAACTGCCCGATGGAGGGGGATAACTAC 119
Buttiauxella_izardii_strain_S3      GGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC 133
Yersinia_kristensenii_strain_A      GGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC 129
Dickeya_zeae_strain_CFBP_2052       GGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC 126
                                    **********************.*:***** *.**********

E.Coli_16SrRNA                      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGAC 150
Escherichia_coli_strain_U_5/41      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGCACAAAGAGGGGGAC 197
Shigella_flexneri_strain_ATCC_      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGAC 187
Enterobacter_cowanii_strain_88      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGAC 168
Salmonella_bongori_strain_BR_1      TGGAAACGGTGGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGAC 194
Cronobacter_muytjensii_strain_      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGTGGGGGAC 159
Trabulsiella_guamensis              TGGAAACGGTAGCTAATACCGCATAACGTCTTCGGACCAAAGTGGGGGAC 171
Citrobacter_farmeri_strain_CDC      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGAC 187
Pantoea_dispersa_strain_LMG260      TGGAAACGGTAGCTAATACCGCATAACGTCGCRAGACCAAAGTGGGGGAC 100
Pectobacterium_cypripedii_stra      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGTGGGGGAC 186
Erwinia_papayae_strain_CFBP_11      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGTGGGGGAC 186
Cedecea_neteri_strain_GTC1717       TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGAC 199
Serratia_marcescens_subsp._sak      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGAC 200
Raoultella_planticola_strain_A      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGTGGGGGAC 189
Klebsiella_oxytoca_strain_ATCC      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGAC 189
Kluyvera_cryocrescens_strain_1      TGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGTGGGGGAC 169
Buttiauxella_izardii_strain_S3      TGGAAACGGTAGCTAATACCGCATAACGTCTTCGGACCAAAGAGGGGGAC 183
Yersinia_kristensenii_strain_A      TGGAAACGGTAGCTAATACCGCATGACCTCGCAAGAGCAAAGTGGGGGAC 179
Dickeya_zeae_strain_CFBP_2052       TGGAAACGGTAGCTAATACCGCATAACGTCTTCGGACCAAAGAGGGGGAC 176
```

```
E.Coli_16SrRNA                      CTTCGGGCCTCTTGCCATCGGATGTGCCCAGATGGGATTAGCTAGTAGGT 200
Escherichia_coli_strain_U_5/41      CTTAGGGCCTCTTGCCATCGGATGTGCCCAGATGGGATTAGCTAGTAGGT 247
Shigella_flexneri_strain_ATCC_      CTTCGGGCCTCTTGCCATCGGATGTGCCCAGATGGGATTAGCTAGTAGGT 237
Enterobacter_cowanii_strain_88      CTTCGGGCCTCTTGCCATCAGATGTGCCCAGATGGGATTAGCTAGTAGGT 218
Salmonella_bongori_strain_BR_1      CTTCGGGCCTCTTGCCATCAGATGTGCCCAGATGGGATTAGCTTGTTGGT 244
Cronobacter_muytjensii_strain_      CTTCGGGCCTCATGCCATCAGATGTGCCCAGATGGGATTAGCTAGTAGGT 209
Trabulsiella_guamensis              CTTCGGGCCTCATGCCATCAGATGTGCCCAGATGGGATTAGCTAGTAGGT 221
Citrobacter_farmeri_strain_CDC      CTTCGGGCCTCTTGCCATCGGATGTGCCCAGATGGGATTAGCTTGTTGGT 237
Pantoea_dispersa_strain_LMG260      CTTCGGGCCTCACACCATCGGATGTGCCCAGATGGGATTAGCTAGTAGGT 150
Pectobacterium_cypripedii_stra      CTTCGGGCCTCACACCATCGGATGAACCCAGATGGGATTAGCTAGTAGGT 236
Erwinia_papayae_strain_CFBP_11      CTTCGGGCCTCACACCATCGGATGTGCCCAGATGGGATTAGCTAGTTGGT 236
Cedecea_neteri_strain_GTC1717       CTTCGGGCCTCTTGCCATCAGATGAACCCAGATGGGATTAGCTAGTAGGT 249
Serratia_marcescens_subsp._sak      CTTCGGGCCTCTTGCCATCAGATGTGCCCAGATGGGATTAGCTAGTAGGT 250
Raoultella_planticola_strain_A      CTTCGGGCCTCATGCCATCAGATGTGCCCAGATGGGATTAGCTAGTAGGT 239
Klebsiella_oxytoca_strain_ATCC      CTTCGGGCCTCTTGCCATCGGATGTGCCCAGATGGGATTAGCTTGTAGGT 239
Kluyvera_cryocrescens_strain_1      CTTCGGGCCTCACACCATCGGATGTGCCCAGATGGGATTAGCTAGTAGGT 219
Buttiauxella_izardii_strain_S3      CTTCGGGCCTCTTGCCATCAGATGTGCCCAGATGGGATTAGCTAGTAGGT 233
Yersinia_kristensenii_strain_A      CTTCGGGCCTCACGCCATCGGATGTGCCCAGATGGGATTAGCTAGTAGGT 229
Dickeya_zeae_strain_CFBP_2052       CTTCGGGCCTCTTGCCATCGGATGTGCCCAGATGGGATTAGCTAGTAGGT 226
                                    ***.*******: .*****.****:.******************:**:***

E.Coli_16SrRNA                      GGGGTAACGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 250
Escherichia_coli_strain_U_5/41      GGGGTAACGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 297
Shigella_flexneri_strain_ATCC_      GGGGTAACGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 287
Enterobacter_cowanii_strain_88      GGGGTAACGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 268
Salmonella_bongori_strain_BR_1      GGGGTAACGGCCCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 294
Cronobacter_muytjensii_strain_      GGGGTAACGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 259
Trabulsiella_guamensis              GGGGTAACGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 271
Citrobacter_farmeri_strain_CDC      GAGGTAACGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 287
Pantoea_dispersa_strain_LMG260      GGGGTAATGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 200
Pectobacterium_cypripedii_stra      GGGGTAACGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 286
Erwinia_papayae_strain_CFBP_11      GAGGTAACGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 286
Cedecea_neteri_strain_GTC1717       GNGGTAATGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 299
Serratia_marcescens_subsp._sak      GGGGTAATGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 300
Raoultella_planticola_strain_A      GGGGTAATGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 289
Klebsiella_oxytoca_strain_ATCC      GAGGTAACGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 289
Kluyvera_cryocrescens_strain_1      GGGGTAATGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 269
Buttiauxella_izardii_strain_S3      GGGGTAATGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 283
Yersinia_kristensenii_strain_A      GGGGTAATGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 279
Dickeya_zeae_strain_CFBP_2052       GGGGTAAAGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGA 276
                                    * ***** *** ****:**********************************

E.Coli_16SrRNA                      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 300
Escherichia_coli_strain_U_5/41      CCAGCAACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 347
Shigella_flexneri_strain_ATCC_      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 337
Enterobacter_cowanii_strain_88      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 318
Salmonella_bongori_strain_BR_1      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 344
Cronobacter_muytjensii_strain_      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 309
Trabulsiella_guamensis              CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 321
Citrobacter_farmeri_strain_CDC      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 337
Pantoea_dispersa_strain_LMG260      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 250
Pectobacterium_cypripedii_stra      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 336
Erwinia_papayae_strain_CFBP_11      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 336
Cedecea_neteri_strain_GTC1717       CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 349
Serratia_marcescens_subsp._sak      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 350
Raoultella_planticola_strain_A      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 339
Klebsiella_oxytoca_strain_ATCC      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 339
Kluyvera_cryocrescens_strain_1      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 319
Buttiauxella_izardii_strain_S3      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 333
Yersinia_kristensenii_strain_A      CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 329
Dickeya_zeae_strain_CFBP_2052       CCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA 326
                                    ******.*******************************************

E.Coli_16SrRNA                      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 349
Escherichia_coli_strain_U_5/41      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCNGCGT 396
Shigella_flexneri_strain_ATCC_      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 386
Enterobacter_cowanii_strain_88      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 367
Salmonella_bongori_strain_BR_1      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 393
Cronobacter_muytjensii_strain_      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 358
Trabulsiella_guamensis              GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 370
Citrobacter_farmeri_strain_CDC      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 386
Pantoea_dispersa_strain_LMG260      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 299
Pectobacterium_cypripedii_stra      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 385
Erwinia_papayae_strain_CFBP_11      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 385
Cedecea_neteri_strain_GTC1717       GTGGGGAATATTGCACAATGGGCGCAAGCTCTGATGCAGCCATGCCGCGT 399
Serratia_marcescens_subsp._sak      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 399
Raoultella_planticola_strain_A      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 388
Klebsiella_oxytoca_strain_ATCC      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 388
Kluyvera_cryocrescens_strain_1      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 368
Buttiauxella_izardii_strain_S3      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT

Yersinia_kristensenii_strain_A      GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 378
Dickeya_zeae_strain_CFBP_2052       GTGGGGAATATTGCACAATGGGCGCAAGC-CTGATGCAGCCATGCCGCGT 375
                                    **************************** *************** ****
```

```
E.Coli_16SrRNA                      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGG 399
Escherichia_coli_strain_U_5/41      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGG 446
Shigella_flexneri_strain_ATCC_      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGG 436
Enterobacter_cowanii_strain_88      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGC 417
Salmonella_bongori_strain_BR_1      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGC 443
Cronobacter_muytjensii_strain_      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGAGGAGGAAGGG 408
Trabulsiella_guamensis              GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGT 420
Citrobacter_farmeri_strain_CDC      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGG 436
Pantoea_dispersa_strain_LMG260      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGC 349
Pectobacterium_cypripedii_stra      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGG 435
Erwinia_papayae_strain_CFBP_11      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGG 435
Cedecea_neteri_strain_GTC1717       GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGAGGAGGAAGGC 449
Serratia_marcescens_subsp._sak      GTGTGAAGAAGGCCTTCGGGTTGTAAAGCACTTTCAGCGAGGAGGAAGGT 449
Raoultella_planticola_strain_A      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGAGGAGGAAGGC 438
Klebsiella_oxytoca_strain_ATCC      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGG 438
Kluyvera_cryocrescens_strain_1      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGAGGAGGAAGGC 418
Buttiauxella_izardii_strain_S3      GTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGAGGAGGAAGGC 432
Yersinia_kristensenii_strain_A      GTGTGAAGAAGGCCTTCGGGTTGTAAAGCACTTTCAGCGAGGAGGAAGGG 428
Dickeya_zeae_strain_CFBP_2052       GTGTGAAGAAGGCCTTCGGGTTGTAAAGCACTTTCAGCGGGGAGGAAGGG 425
                                    **.************************ **********.*********

E.Coli_16SrRNA                      AGTAAAGTTAATACCTTTGCTCATTGACGTTACCCGCAGAAGAAGCACCG 449
Escherichia_coli_strain_U_5/41      AGTAAAGTTAATACCTTTGCTCATTGACGTTACCCGCAGAAGAAGCACCG 496
Shigella_flexneri_strain_ATCC_      AGTAAAGTTAATACCTTTGCTCATTGACGTTACCCGCAGAAGAAGCACCG 486
Enterobacter_cowanii_strain_88      G-ATGGGTTAATAACCGCGTCGATTGACGTTACCCGCAGAAGAAGCACCG 466
Salmonella_bongori_strain_BR_1      GHCAAGGTTAATAACCTTGTTGATTGACGTTACCCGCAGAAGAAGCACCG 493
Cronobacter_muytjensii_strain_      GTTAAGGTTAATAACCTTGGNCATTGACGTTACTCGCAGAAGAAGCACCG 458
Trabulsiella_guamensis              GTTGTGGTTAATAACCAGAGCAATTGACGTTACCCGCAGAAGAAGCACCG 470
Citrobacter_farmeri_strain_CDC      GTTATGGTTAATAACCTTANCCATTGACGTTACCCGCAGAAGAAGCACCG 486
Pantoea_dispersa_strain_LMG260      GGTGAGGTTAATAACCTYGCCGATTGACGTTACCCGCAGAAGAAGCACCG 399
Pectobacterium_cypripedii_stra      GTTGAGTTTAATACGCTCAATCATTGACGTTACCCGCAGAAGAAGCACCG 485
Erwinia_papayae_strain_CFBP_11      AGTGGAGTTAATAACTTCATTCATTGACGTTACCCGCAGAAGAAGCACCG 485
Cedecea_neteri_strain_GTC1717       NTTAAGGTTAATAACCTTNGNGATTGACGTTACTCGCAGAAGAAGCACCG 499
Serratia_marcescens_subsp._sak      GGTGAACTTAATACGCTCATCAATTGACGTTACTCGCAGAAGAAGCACCG 499
Raoultella_planticola_strain_A      GTTAAGGTTAATAACCTTAGCGATTGACGTTACTCGCAGAAGAAGCACCG 488
Klebsiella_oxytoca_strain_ATCC      AGTGAGGTTAATAACCTTATTCATTGACGTTACCCGCAGAAGAAGCACCG 488
Kluyvera_cryocrescens_strain_1      ATTAAGGTTAATAACCTTAGTGATTGACGTTACTCGCAGAAGAAGCACCG 468
Buttiauxella_izardii_strain_S3      ATTGTGGTTAATAACCACAGTGATTGACGTTACTCGCAGAAGAAGCACCG 482
Yersinia_kristensenii_strain_A      TTCAGTGTTAATAGCACTGAGCATTGACGTTACTCGCAGAAGAAGCACCG 478
Dickeya_zeae_strain_CFBP_2052       GGCAGGCTTAATACGTCTGTTCATTGACGTTACCCGCAGAAGAAGCACCG 475
                                           ******         *********** ****************

E.Coli_16SrRNA                      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 499
Escherichia_coli_strain_U_5/41      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 546
Shigella_flexneri_strain_ATCC_      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 536
Enterobacter_cowanii_strain_88      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 516
Salmonella_bongori_strain_BR_1      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 543
Cronobacter_muytjensii_strain_      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 508
Trabulsiella_guamensis              GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 520
Citrobacter_farmeri_strain_CDC      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 536
Pantoea_dispersa_strain_LMG260      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 449
Pectobacterium_cypripedii_stra      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 535
Erwinia_papayae_strain_CFBP_11      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 535
Cedecea_neteri_strain_GTC1717       GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 549
Serratia_marcescens_subsp._sak      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 549
Raoultella_planticola_strain_A      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 538
Klebsiella_oxytoca_strain_ATCC      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 538
Kluyvera_cryocrescens_strain_1      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 518
Buttiauxella_izardii_strain_S3      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 532
Yersinia_kristensenii_strain_A      GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 528
Dickeya_zeae_strain_CFBP_2052       GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAAT 525
                                    **************************************************

E.Coli_16SrRNA                      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGTCAGATGT 549
Escherichia_coli_strain_U_5/41      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGTCAGATGT 596
Shigella_flexneri_strain_ATCC_      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGTCAGATGT 586
Enterobacter_cowanii_strain_88      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTCAAGTCGGATGT 566
Salmonella_bongori_strain_BR_1      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTCAAGTCGGATGT 593
Cronobacter_muytjensii_strain_      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTCAAGTCGGATGT 558
Trabulsiella_guamensis              CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTCAAGTCGGATGT 570
Citrobacter_farmeri_strain_CDC      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTCAAGTCGGATGT 586
Pantoea_dispersa_strain_LMG260      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTTAAGTCAGATGT 499
Pectobacterium_cypripedii_stra      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTCAAGTCGGATGT 585
Erwinia_papayae_strain_CFBP_11      CGGAATGACTGGGCGTAAAGCGCACGCAGGCGGTCTGTTAAGTCAGATGT 585
Cedecea_neteri_strain_GTC1717       CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGTCGGATGT 599
Serratia_marcescens_subsp._sak      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGTCAGATGT 599
Raoultella_planticola_strain_A      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGTCAGATGT 588
Klebsiella_oxytoca_strain_ATCC      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTCAAGTCGGATGT

Kluyvera_cryocrescens_strain_1      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTCAAGTCGGATGT 568
Buttiauxella_izardii_strain_S3      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTCAAGTCGGATGT 582
Yersinia_kristensenii_strain_A      CGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGTCAGATGT 578
Dickeya_zeae_strain_CFBP_2052       CGGAATGACTGGGCGTAAAGCGCACGCAGGCGGTCTGTTAAGTTGGATGT 575
                                    ****** ************************** *** **** .*****

E.Coli_16SrRNA                      GAAATCCCCGGGCTCAACCTGGGAACTGCATCTGATACTGGCAAGCTTGA 599
Escherichia_coli_strain_U_5/41      GAAATCCCCGGGCTCAACCTGGGAACTGCATCTGATACTGGCAAGCTTGA 646
```

```
Shigella_flexneri_strain_ATCC_      GAAATCCCCGGGCTCAACCTGGGAACTGCATCTGATACTGGCAAGCTTGA 636
Enterobacter_cowanii_strain_88      GAAATCCCCGGGCTCAACCTGGGAACTGCATCCGAAACTGGCAGGCTTGA 616
Salmonella_bongori_strain_BR_1      GAAATCCCCGGGCTCAACCTGGGAACTGCATCCGAAACTGGCAGGCTTGA 643
Cronobacter_muytjensii_strain_      GAAATCCCCGGGCTCAACCTGGGAACTGCATCCGAAACTGGCAGGCTTGA 608
Trabulsiella_guamensis              GAAATCCCCGGGCTCAACCTGGGAACTGCATCCGAAACTGGCAGGCTTGA 620
Citrobacter_farmeri_strain_CDC      GAAATCCCCGGGCTCAACCTGGGAACTGCATTCGAAACTGGCAGGCTTGA 636
Pantoea_dispersa_strain_LMG260      GAAATCCCCGGGCTTAACCTGGGAACTGCATTTGAAACTGGCAGGCTTGA 549
Pectobacterium_cypripedii_stra      GAAATCCCCGGGCTTAACCTGGGAACTGCATTCGAAACTGGCAGGCTAGA 635
Erwinia_papayae_strain_CFBP_11      GAAATCCCCGGGCTTAACCTGGGAACTGCATTTGAAACTGGCAGGCTTGA 635
Cedecea_neteri_strain_GTC1717       GAAATCCCCGGGCTCAACCTGGGAACTGCATTCGAAACTGGCAAGCTTGA 649
Serratia_marcescens_subsp._sak      GAAATCCCCGGGCTCAACCTGGGAACTGCATTTGAAACTGGCAAGCTAGA 649
Raoultella_planticola_strain_A      GAAATCCCCGGGCTCAACCTGGGAACTGCATTTGAAACTGGCAAGCTTGA 638
Klebsiella_oxytoca_strain_ATCC      GAAATCCCCGGGCTCAACCTGGGAACTGCATTCGAAACTGGCAGGCTGGA 638
Kluyvera_cryocrescens_strain_1      GAAATCCCCGGGCTCAACCTGGGAACTGCATTCGAAACTGGCAGGCTAGA 618
Buttiauxella_izardii_strain_S3      GAAATCCCCGGGCTCAACCTGGGAACTGCATTCGAAACTGGCAGGCTAGA 632
Yersinia_kristensenii_strain_A      GAAATCCCCGCGCTTAACGTGGGAACTGCATTTGAAACTGGCAAGCTAGA 628
Dickeya_zeae_strain_CFBP_2052       GAAATCCCCGGGCTTAACCTGGGAACTGCATTCAAAACTGACAGGCTAGA 625
                                    ********** *** *** ************  .*:****.**.*** **

E.Coli_16SrRNA                      GTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 649
Escherichia_coli_strain_U_5/41      GTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 696
Shigella_flexneri_strain_ATCC_      GTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 686
Enterobacter_cowanii_strain_88      GTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 666
Salmonella_bongori_strain_BR_1      GTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 693
Cronobacter_muytjensii_strain_      GTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 658
Trabulsiella_guamensis              GTCTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 670
Citrobacter_farmeri_strain_CDC      GTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 686
Pantoea_dispersa_strain_LMG260      GTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 599
Pectobacterium_cypripedii_stra      GTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 685
Erwinia_papayae_strain_CFBP_11      GTCTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 685
Cedecea_neteri_strain_GTC1717       GTCTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 699
Serratia_marcescens_subsp._sak      GTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 699
Raoultella_planticola_strain_A      GTCTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 688
Klebsiella_oxytoca_strain_ATCC      GTCTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 688
Kluyvera_cryocrescens_strain_1      GTCTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 668
Buttiauxella_izardii_strain_S3      GTCTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 682
Yersinia_kristensenii_strain_A      GTCTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 678
Dickeya_zeae_strain_CFBP_2052       GTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG 675
                                    **** ********************************************

E.Coli_16SrRNA                      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACGAAGACTGACG 699
Escherichia_coli_strain_U_5/41      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACGAAGACTGACG 746
Shigella_flexneri_strain_ATCC_      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACGAAGACTGACG 736
Enterobacter_cowanii_strain_88      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACGAAGACTGACG 716
Salmonella_bongori_strain_BR_1      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACGAAGACTGACG 743
Cronobacter_muytjensii_strain_      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACGAAGACTGACG 708
Trabulsiella_guamensis              ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCTGGACAAAGACTGACG 720
Citrobacter_farmeri_strain_CDC      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACGAAGACTGACG 736
Pantoea_dispersa_strain_LMG260      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACGAAGACTGACG 649
Pectobacterium_cypripedii_stra      ATCTGGAGGAATACCGGTGGCGAAGGGCGGCCCCCTGGACGAAGACTGACG 735
Erwinia_papayae_strain_CFBP_11      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACAAAGACTGACG 735
Cedecea_neteri_strain_GTC1717       ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACAAAGACTGACG 749
Serratia_marcescens_subsp._sak      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACGAAGACTGACG 749
Raoultella_planticola_strain_A      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACAAAGACTGACG 738
Klebsiella_oxytoca_strain_ATCC      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACAAAGACTGACG 738
Kluyvera_cryocrescens_strain_1      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCCTGGACAAAGACTGACG 718
Buttiauxella_izardii_strain_S3      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCTGGACAAAGACTGACG 732
Yersinia_kristensenii_strain_A      ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCTGGACAAAGACTGACG 728
Dickeya_zeae_strain_CFBP_2052       ATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCTGGACGAAGACTGACG 725
                                    *************************  * *********.**********

E.Coli_16SrRNA                      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 749
Escherichia_coli_strain_U_5/41      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 796
Shigella_flexneri_strain_ATCC_      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 786
Enterobacter_cowanii_strain_88      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 766
Salmonella_bongori_strain_BR_1      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 793
Cronobacter_muytjensii_strain_      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 758
Trabulsiella_guamensis              CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 770
Citrobacter_farmeri_strain_CDC      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 786
Pantoea_dispersa_strain_LMG260      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 699
Pectobacterium_cypripedii_stra      CTCAGTTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 785
Erwinia_papayae_strain_CFBP_11      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 785
Cedecea_neteri_strain_GTC1717       CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 799
Serratia_marcescens_subsp._sak      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 799

Raoultella_planticola_strain_A      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 788
Klebsiella_oxytoca_strain_ATCC      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 788
Kluyvera_cryocrescens_strain_1      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 768
Buttiauxella_izardii_strain_S3      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 782
Yersinia_kristensenii_strain_A      CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 778
Dickeya_zeae_strain_CFBP_2052       CTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC 775
                                    ***** ********************************************

E.Coli_16SrRNA                      CACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 799
Escherichia_coli_strain_U_5/41      CACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 846
Shigella_flexneri_strain_ATCC_      CACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 836
Enterobacter_cowanii_strain_88      CACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 816
```

```
Salmonella_bongori_strain_BR_1      CACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 843
Cronobacter_muytjensii_strain_      CACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 808
Trabulsiella_guamensis              CACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 820
Citrobacter_farmeri_strain_CDC      CACGCCGTAAACGATGTCTATTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 836
Pantoea_dispersa_strain_LMG260      CACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGYGTGGCTT 749
Pectobacterium_cypripedii_stra      CACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 835
Erwinia_papayae_strain_CFBP_11      CACGCCGTAAACGATGTCGATTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 835
Cedecea_neteri_strain_GTC1717       CACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 849
Serratia_marcescens_subsp._sak      CACGCTGTAAACGATGTCGATTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 849
Raoultella_planticola_strain_A      CACGCTGTAAACGATGTCGACTTGGAGGTTGTTCCCTTGAGGAGTGGCTT 838
Klebsiella_oxytoca_strain_ATCC      CACGCTGTAAACGATGTCGACTTGGAGGTTGTTCCCTTGAGGAGTGGCTT 838
Kluyvera_cryocrescens_strain_1      CACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 818
Buttiauxella_izardii_strain_S3      CACGCCGTAAACGATGTCGACTTGGAGGTTGTTCCCTTGAGGAGTGGCTT 832
Yersinia_kristensenii_strain_A      CACGCTGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 828
Dickeya_zeae_strain_CFBP_2052       CACGCTGTAAACGATGTCGATTTGGAGGTTGTGCCCTTGAGGCGTGGCTT 825
                                    ***** ************ * *********** ********* *******

E.Coli_16SrRNA                      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGA----------------- 832
Escherichia_coli_strain_U_5/41      CCGGANNTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 896
Shigella_flexneri_strain_ATCC_      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 886
Enterobacter_cowanii_strain_88      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 866
Salmonella_bongori_strain_BR_1      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 893
Cronobacter_muytjensii_strain_      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 858
Trabulsiella_guamensis              CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 870
Citrobacter_farmeri_strain_CDC      CCGGAGCTAACGCGTTAAATAGACCGCCTGGGGAGTACGGCCGCAAGGTT 886
Pantoea_dispersa_strain_LMG260      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 799
Pectobacterium_cypripedii_stra      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 885
Erwinia_papayae_strain_CFBP_11      CCGTAGCTAACGCGTTAAATCGACCGCCTGGGGAGTACGGCCGCAAGGTT 885
Cedecea_neteri_strain_GTC1717       CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 899
Serratia_marcescens_subsp._sak      CCGGAGCTAACGCGTTAAATCGACCGCCTGGGGAGTACGGCCGCAAGGTT 899
Raoultella_planticola_strain_A      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 888
Klebsiella_oxytoca_strain_ATCC      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 888
Kluyvera_cryocrescens_strain_1      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 868
Buttiauxella_izardii_strain_S3      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 882
Yersinia_kristensenii_strain_A      CCGGAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTT 878
Dickeya_zeae_strain_CFBP_2052       CCGGAGCTAACGCGTTAAATGGACCGCCTGGGGAGTACGGCCGCAAGGTT 875
                                    *** *. ***********.* ***********.

E.Coli_16SrRNA                      ----------------------------------------------------
Escherichia_coli_strain_U_5/41      AAAACTCAAATGAATTGACGGGGGCCG-CACAAGCGGTGGAGCATGTGGT 945
Shigella_flexneri_strain_ATCC_      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 936
Enterobacter_cowanii_strain_88      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 916
Salmonella_bongori_strain_BR_1      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 943
Cronobacter_muytjensii_strain_      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 908
Trabulsiella_guamensis              AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGTT 920
Citrobacter_farmeri_strain_CDC      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 936
Pantoea_dispersa_strain_LMG260      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 849
Pectobacterium_cypripedii_stra      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 935
Erwinia_papayae_strain_CFBP_11      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 935
Cedecea_neteri_strain_GTC1717       AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAG-ATGTGGT 948
Serratia_marcescens_subsp._sak      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 949
Raoultella_planticola_strain_A      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 938
Klebsiella_oxytoca_strain_ATCC      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 938
Kluyvera_cryocrescens_strain_1      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 918
Buttiauxella_izardii_strain_S3      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 932
Yersinia_kristensenii_strain_A      AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 928
Dickeya_zeae_strain_CFBP_2052       AAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT 925


E.Coli_16SrRNA                      ----------------------------------------------------
Escherichia_coli_strain_U_5/41      TTAATT-CGATGCAACGCGAAGAACCTTACCTGGTCTTGACATCCACGGA 994
Shigella_flexneri_strain_ATCC_      TTAATT-CGATGCAACGCGAAGAACCTTACCTGGTCTTGACATCCACGGA 985
Enterobacter_cowanii_strain_88      TTAATT-CGATGCAACGCGAAGAACCTTACCTGGTCTTGACATCCACAGA 965
Salmonella_bongori_strain_BR_1      TTAATT-CGATGCAACGCGAAGAACCTTACCTGGTCTTGACATCCACAGA 992
Cronobacter_muytjensii_strain_      TTAATT-CGATGCAACGCGAAGAACCTTACCTGGTCTTGACATCCAGAGA 957
Trabulsiella_guamensis              TTAATTTCGATGCAACGCGAAGAACCTTACCTGGTCTTGACACCCACAGA 970
Citrobacter_farmeri_strain_CDC      TTAATT-CGATGCAACGCGAAGAACCTTACCTGGTCTTGACATCCACAGA 985
Pantoea_dispersa_strain_LMG260      TTAATT-CGATGCAACGCGAAGAACCTTACCTGGCCTTGACATCCAGAGA 898
Pectobacterium_cypripedii_stra      TTAATT-CGATGCAACGCGAAGAACCTTACCTGGCCTTGACATCCACAGA 984
Erwinia_papayae_strain_CFBP_11      TTAATT-CGATGCAACGCGAAGAACCTTACCTGGCCTTGACATCCACGGA

Citrobacter_farmeri_strain_CDC      TGGATCAGAATGCCACGGTGAATACGTTCCCGGGCCTTGTACACACCGCC 1384
Pantoea_dispersa_strain_LMG260      TAGATCAGAATGCTACGGTGAATACGTTCCCGGGCCTTGTACACACCGCC 1297
Pectobacterium_cypripedii_stra      TAGATCAGAATGCTACGGTGAATACGTTCCCGGGCCTTGTACACACCGCC 1383
Erwinia_papayae_strain_CFBP_11      TGGATCAGAATGCCACGGTGAATACGTTCCCGGGCCTTGTACACACCGCC 1383
Cedecea_neteri_strain_GTC1717       TAGATCAGAATGCTACGGTGAATACGTTCCCGGGCCTTGCACACACCGCC 1396
Serratia_marcescens_subsp._sak      TAGATCAGAATGCTACGGTGAATACGTTCCCGGGCCTTGTACACACCGCC 1397
Raoultella_planticola_strain_A      TAGATCAGAATGCTACGGTGAATACGTTCCCGGGCCTTGTACACACCGCC 1386
Klebsiella_oxytoca_strain_ATCC      TGGATCAGAATGCCACGGTGAATACGTTCCCGGGCCTTGTACACACCGCC 1386
Kluyvera_cryocrescens_strain_1      TAGATCAGAATGCTACGGTGAATACGTTCCCGGGCCTTGTACACACCGCC 1366
Buttiauxella_izardii_strain_S3      TAGATCAGAATGCTACGGTGAATACGTTCCCGGGCCTTGTACACACCGCC 1380
Yersinia_kristensenii_strain_A      TAGATCAGAATGCTACGGTGAATACGTTCCCGGGCCTTGTACACACCGCC 1377
Dickeya_zeae_strain_CFBP_2052       TAGATCAGAATGCTACGGTGAATACGTTCCCGGGCCTTGTACACACCGCC 1373


E.Coli_16SrRNA                      ----------------------------------------------------
Escherichia_coli_strain_U_5/41      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACTT-CG 1442
```

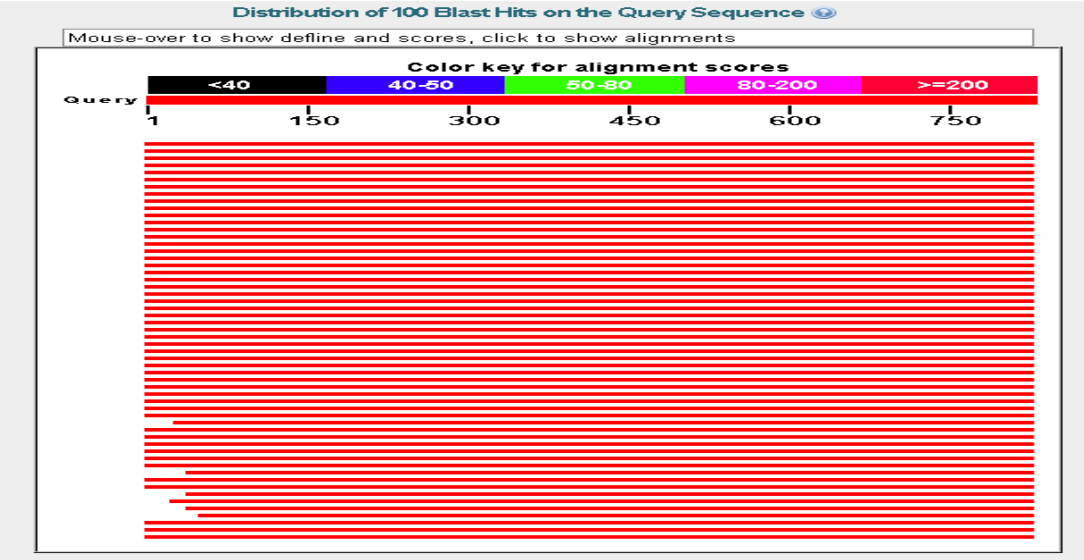
```
Shigella_flexneri_strain_ATCC_      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1434
Enterobacter_cowanii_strain_88      --------------------------------------------------
Salmonella_bongori_strain_BR_1      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1441
Cronobacter_muytjensii_strain_      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1407
Trabulsiella_guamensis              CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1416
Citrobacter_farmeri_strain_CDC      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1434
Pantoea_dispersa_strain_LMG260      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1347
Pectobacterium_cypripedii_stra      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1433
Erwinia_papayae_strain_CFBP_11      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1433
Cedecea_neteri_strain_GTC1717       CGT---------------------------------------------- 1399
Serratia_marcescens_subsp._sak      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1447
Raoultella_planticola_strain_A      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1436
Klebsiella_oxytoca_strain_ATCC      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1436
Kluyvera_cryocrescens_strain_1      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1416
Buttiauxella_izardii_strain_S3      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1430
Yersinia_kristensenii_strain_A      CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCG 1427
Dickeya_zeae_strain_CFBP_2052       CGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTC- 1422

E.Coli_16SrRNA                      --------------------------------------------------
Escherichia_coli_strain_U_5/41      GGAGGGCG---------------------------------------- 1450
Shigella_flexneri_strain_ATCC_      GGAGGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAA 1484
Enterobacter_cowanii_strain_88      --------------------------------------------------
Salmonella_bongori_strain_BR_1      GGAGGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAA 1491
Cronobacter_muytjensii_strain_      GGAGGGCGCTTACCACTT------------------------------- 1425
Trabulsiella_guamensis              GGAGGGCGCTTACCACTTGTGATTA------------------------- 1441
Citrobacter_farmeri_strain_CDC      GGAGGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAA 1484
Pantoea_dispersa_strain_LMG260      --------------------------------------------------
Pectobacterium_cypripedii_stra      GGAGGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAA 1483
Erwinia_papayae_strain_CFBP_11      GGA---------------------------------------------- 1436
Cedecea_neteri_strain_GTC1717       --------------------------------------------------
Serratia_marcescens_subsp._sak      GGAGGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAA 1497
Raoultella_planticola_strain_A      --------------------------------------------------
Klebsiella_oxytoca_strain_ATCC      --------------------------------------------------
Kluyvera_cryocrescens_strain_1      GGAGGGCGCTTACCACTTTGTGATTCATGATT------CATTGGTATCA- 1459
Buttiauxella_izardii_strain_S3      GGAGGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAA 1480
Yersinia_kristensenii_strain_A      GGAGGGCGCTTACCACTTTGTGATTCATGACTGG---------------- 1461
Dickeya_zeae_strain_CFBP_2052       --------------------------------------------------

E.Coli_16SrRNA                      ----------------------------------
Escherichia_coli_strain_U_5/41      ----------------------------------
Shigella_flexneri_strain_ATCC_      GGTA------------------------------- 1488
Enterobacter_cowanii_strain_88      ----------------------------------
Salmonella_bongori_strain_BR_1      GGTAACC---------------------------- 1498
Cronobacter_muytjensii_strain_      ----------------------------------
Trabulsiella_guamensis              ----------------------------------
Citrobacter_farmeri_strain_CDC      GGGTAACCGTAGGGGAACCTGCGGCTG-------- 1511
Pantoea_dispersa_strain_LMG260      ----------------------------------
Pectobacterium_cypripedii_stra      GGTAACCGTAGG----------------------- 1495
Erwinia_papayae_strain_CFBP_11      ----------------------------------
Cedecea_neteri_strain_GTC1717       ----------------------------------
Serratia_marcescens_subsp._sak      GGTAACCGTAGGGGAACCTGCGGCTGGATCACCTC 1532
Raoultella_planticola_strain_A      ----------------------------------
Klebsiella_oxytoca_strain_ATCC      ----------------------------------
Kluyvera_cryocrescens_strain_1      ----------------------------------
Buttiauxella_izardii_strain_S3      GGTAACCGTAGGGGAACC----------------- 1498
Yersinia_kristensenii_strain_A      ----------------------------------
Dickeya_zeae_strain_CFBP_2052       ----------------------------------
```

APPENDIX II

TABLE 6 :-BLAST RESULT



Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
NR_026331.1	Shigella flexneri strain ATCC 29903 16S ribosomal RNA, partial sequen	1515	1515	99%	0.0	99%	
NR_026332.1	Shigella dysenteriae strain ATCC 13313 16S ribosomal RNA, partial se	1507	1507	99%	0.0	99%	
NR_027549.1	Escherichia fergusonii ATCC 35469 16S ribosomal RNA, partial sequen	1506	1506	99%	0.0	99%	
NR_024570.1	Escherichia coli strain U 5/41 16S ribosomal RNA, partial sequence	1493	1493	99%	0.0	99%	
NR_025569.1	Escherichia albertii strain Albert 19982 16S ribosomal RNA, partial seq	1454	1454	99%	0.0	98%	
NR_044059.1	Cronobacter muytjensii strain E603; ATCC 51329 16S ribosomal RNA,	1428	1428	99%	0.0	98%	
NR_025566.1	Enterobacter cowanii strain 888-76 16S ribosomal RNA, partial sequer	1419	1419	99%	0.0	97%	
NR_041699.1	Salmonella bongori strain BR 1859 16S ribosomal RNA, partial sequenc	1399	1399	99%	0.0	97%	
NR_042154.1	Enterobacter hormaechei ATCC 49162 strain CIP 103441 16S ribosom	1397	1397	99%	0.0	97%	
NR_044061.1	Cronobacter turicensis z3032 strain z3032 16S ribosomal RNA, partial	1389	1389	99%	0.0	97%	
NR_024861.1	Citrobacter farmeri strain CDC 2991-81 16S ribosomal RNA, partial se	1386	1386	99%	0.0	97%	
NR_041527.1	Citrobacter youngae strain GTC 1314 16S ribosomal RNA, partial sequ	1382	1382	99%	0.0	97%	
NR_044060.1	Cronobacter malonaticus strain E825; CDC 1058-77; API 76-2121 16S	1376	1376	99%	0.0	97%	
NR_044076.1	Cronobacter sakazakii strain ATCC 29544 16S ribosomal RNA, partial s	1376	1376	99%	0.0	97%	
NR_028686.1	Citrobacter sedlakii strain I-75 16S ribosomal RNA, partial sequence	1373	1373	99%	0.0	96%	
NR_028685.1	Citrobacter rodentium strain DO 14784 16S ribosomal RNA, partial seq	1371	1371	99%	0.0	96%	
NR_024640.1	Enterobacter asburiae strain JCM6051 16S ribosomal RNA, partial seq	1369	1369	99%	0.0	96%	
NR_044371.1	Salmonella enterica subsp. houtenae strain DSM 9221 16S ribosomal I	1367	1367	99%	0.0	96%	
NR_044370.1	Salmonella enterica subsp. indica strain DSM 14848 16S ribosomal RN	1365	1365	99%	0.0	96%	
NR_041696.1	Salmonella enterica subsp. arizonae strain ATCC 13314 16S ribosomal	1365	1365	99%	0.0	96%	
NR_044977.1	Enterobacter cancerogenus strain LMG 2693 16S ribosomal RNA, part	1363	1363	99%	0.0	96%	
NR_024641.1	Enterobacter gergoviae strain JCM1234 16S ribosomal RNA, partial se	1362	1362	99%	0.0	96%	
NR_028688.1	Citrobacter murliniae strain CDC 2970-59 16S ribosomal RNA, partial s	1360	1360	99%	0.0	96%	
NR_028875.1	Enterobacter pyrinus strain B43 16S ribosomal RNA, complete sequen	1356	1356	99%	0.0	96%	
NR_044372.1	Salmonella enterica subsp. salamae strain DSM 9220 16S ribosomal RI	1356	1356	99%	0.0	96%	
NR_028687.1	Citrobacter braakii strain 167 16S ribosomal RNA, partial sequence	1354	1354	99%	0.0	96%	
NR_024996.1	Raoultella planticola strain ATCC 33531 16S ribosomal RNA, partial se	1352	1352	99%	0.0	96%	
NR_041749.1	Klebsiella oxytoca strain ATCC 13182 16S ribosomal RNA, partial sequ	1352	1352	99%	0.0	96%	
NR_040930.1	Cedecea neteri strain GTC1717 16S ribosomal RNA, partial sequence	1351	1351	99%	0.0	96%	
NR_036886.1	Serratia marcescens subsp. sakuensis strain KRED 16S ribosomal RNA	1349	1349	99%	0.0	96%	
NR_028894.1	Citrobacter freundii strain DSM 30039 16S ribosomal RNA, partial sequ	1349	1349	99%	0.0	96%	
NR_042349.1	Enterobacter ludwigii strain : EN-119 = DSMZ 16688 = CIP 108491 16	1347	1347	99%	0.0	96%	
NR_041697.1	Citrobacter gillenii strain CDC 4693-86 16S ribosomal RNA, partial seq	1343	1343	99%	0.0	96%	
NR_028803.1	Kluyvera cryocrescens strain 12993 16S ribosomal RNA, partial seque	1341	1341	99%	0.0	96%	
NR_024862.1	Citrobacter werkmanii strain CDC 0876-58 16S ribosomal RNA, partial	1341	1341	99%	0.0	96%	
NR_044385.1	Serratia nematodiphila strain DZ0503SBS1 16S ribosomal RNA, partial	1338	1338	99%	0.0	96%	

NR_037102.1	Pectobacterium cacticida strain 1-12 16S ribosomal RNA, partial sequ	1304	1304	99%	0.0	95%
NR_042893.1	Trabulsiella guamensis 16S ribosomal RNA, partial sequence	1303	1303	94%	0.0	96%
NR_041970.1	Erwinia amylovora strain DSM 30165 16S ribosomal RNA, partial seque	1303	1303	99%	0.0	95%
NR_041833.1	Yersinia ruckeri strain ATCC 29473 16S ribosomal RNA, partial sequen	1299	1299	99%	0.0	95%
NR_025161.1	Yersinia rohdei strain ATCC 43380 16S ribosomal RNA, partial sequenc	1299	1299	99%	0.0	95%
NR_027546.1	Yersinia mollaretii ATCC 43969 16S ribosomal RNA, partial sequence	1299	1299	99%	0.0	95%
NR_042062.1	Serratia liquefaciens strain CIP 103238 16S ribosomal RNA, partial sec	1299	1299	97%	0.0	96%
NR_025341.1	Serratia proteamaculans strain DSM 4543 16S ribosomal RNA, partial :	1299	1299	99%	0.0	95%
NR_025340.1	Serratia grimesii strain DSM 30063 16S ribosomal RNA, partial sequenc	1299	1299	99%	0.0	95%
NR_025331.1	Buttiauxella izardii strain S3/2-161 16S ribosomal RNA, partial sequen	1297	1297	99%	0.0	95%
NR_025330.1	Buttiauxella gaviniae strain S1/1-984 16S ribosomal RNA, partial sequ	1297	1297	99%	0.0	95%
NR_026045.1	Pantoea ananatis strain 1846 16S ribosomal RNA, partial sequence	1297	1297	99%	0.0	95%
NR_027545.1	Yersinia intermedia ATCC 29909 16S ribosomal RNA, partial sequence	1293	1293	99%	0.0	95%
NR_041832.1	Yersinia enterocolitica strain ATCC 9610 16S ribosomal RNA, partial se	1293	1293	99%	0.0	95%
NR_041923.1	Dickeya zeae strain CFBP 2052 16S ribosomal RNA, partial sequence	1293	1293	99%	0.0	95%
NR_041924.1	Dickeya dieffenbachiae strain CFBP 2051 16S ribosomal RNA, partial s	1293	1293	99%	0.0	95%
NR_041976.1	Erwinia rhapontici strain DSM 4484 16S ribosomal RNA, partial sequen	1293	1293	99%	0.0	95%
NR_041974.1	Erwinia mallotivora strain DSM 4565 16S ribosomal RNA, partial sequer	1293	1293	99%	0.0	95%
NR_026046.1	Pectobacterium betavasculorum strain ATCC 43762 16S ribosomal RN	1293	1293	99%	0.0	95%
NR_028893.1	Buttiauxella warmboldiae strain NSW 326 16S ribosomal RNA, partial s	1291	1291	99%	0.0	95%
NR_036919.1	Buttiauxella noackiae strain NSW 11 16S ribosomal RNA, partial seque	1291	1291	99%	0.0	95%
NR_025328.1	Buttiauxella brennerae strain S1/6-571 16S ribosomal RNA, partial sec	1291	1291	99%	0.0	95%
NR_041968.1	Buttiauxella agrestis ATCC 33320 strain DSM 4586 16S ribosomal RNA	1291	1291	99%	0.0	95%
NR_042422.1	Erwinia tasmaniensis Et1/99 strain : Et1/99 16S ribosomal RNA, partia	1290	1290	99%	0.0	95%
NR_042485.1	Yersinia similis strain : Y228 16S ribosomal RNA, partial sequence	1288	1288	99%	0.0	95%

NR_041980.1	Serratia marcescens subsp. marcescens ATCC 13880 strain DSM 301	1338	1338	99%	0.0	96%
NR_043679.1	Enterobacter pulveris strain 601/05 16S ribosomal RNA, partial sequer	1336	1336	99%	0.0	96%
NR_028993.1	Enterobacter kobei strain CIP 105566 16S ribosomal RNA, partial sequ	1336	1336	99%	0.0	96%
NR_044799.1	Raoultella ornithinolytica strain CIP 103 364 16S ribosomal RNA, partia	1336	1336	96%	0.0	97%
NR_037084.1	Klebsiella pneumoniae subsp. rhinoscleromatis strain R-70 16S riboson	1336	1336	99%	0.0	96%
NR_024643.1	Enterobacter aerogenes strain JCM1235 16S ribosomal RNA, partial se	1336	1336	99%	0.0	96%
NR_024883.1	Kluyvera georgiana strain ATCC 51603 16S ribosomal RNA, partial seq	1334	1334	99%	0.0	95%
NR_044978.1	Enterobacter cloacae subsp. dissolvens strain LMG 2683 16S ribosom	1334	1334	99%	0.0	96%
NR_044058.1	Cronobacter dublinensis subsp. lausanensis strain E515 16S ribosomal	1332	1332	99%	0.0	96%
NR_042748.1	Erwinia papayae strain CFBP 11606 16S ribosomal RNA, partial sequer	1332	1332	99%	0.0	96%
NR_041927.1	Escherichia vulneris strain ATCC 33821 16S ribosomal RNA, partial sec	1332	1332	95%	0.0	97%
NR_028912.1	Enterobacter cloacae strain 279-56 16S ribosomal RNA, partial seque	1332	1332	99%	0.0	96%
NR_041750.1	Klebsiella pneumoniae subsp. ozaenae strain ATCC11296 16S ribosom	1330	1330	99%	0.0	96%
NR_044057.1	Cronobacter dublinensis subsp. lactaridi strain E464 16S ribosomal RN	1328	1328	95%	0.0	97%
NR_044373.1	Salmonella enterica subsp. diarizonae strain DSM 14847 16S ribosoma	1327	1327	97%	0.0	96%
NR_044062.1	Cronobacter dublinensis subsp. dublinensis strain DES187 16S ribosom	1327	1327	95%	0.0	97%
NR_043883.1	Pantoea dispersa strain LMG2603 16S ribosomal RNA, partial sequenc	1327	1327	93%	0.0	97%
NR_037085.1	Raoultella terrigena strain 84 16S ribosomal RNA, partial sequence	1327	1327	99%	0.0	95%
NR_044800.1	Pantoea stewartii subsp. stewartii strain ATCC 8199 16S ribosomal RN	1325	1325	99%	0.0	95%
NR_044979.1	Erwinia psidii strain LMG 7034 16S ribosomal RNA, partial sequence	1323	1323	99%	0.0	95%
NR_025243.1	Cedecea davisae strain DSM 4568 16S ribosomal RNA, partial sequenc	1319	1319	93%	0.0	97%
NR_028802.1	Kluyvera intermedia strain 256 16S ribosomal RNA, partial sequence	1319	1319	99%	0.0	95%
NR_026049.1	Erwinia persicina strain HK 204 16S ribosomal RNA, partial sequence	1319	1319	99%	0.0	95%
NR_043980.1	Pantoea citrea strain JCM 8882 16S ribosomal RNA, partial sequence	1315	1315	99%	0.0	95%
NR_041973.1	Pectobacterium cypripedii strain DSM 3873 16S ribosomal RNA, partia	1315	1315	99%	0.0	95%
NR_026047.1	Pectobacterium wasabiae strain SR91 16S ribosomal RNA, partial sequ	1315	1315	99%	0.0	95%
NR_024642.1	Enterobacter amnigenus strain JCM1237 16S ribosomal RNA, partial se	1315	1315	99%	0.0	95%
NR_041972.1	Erwinia chrysanthemi strain DSM 4610 16S ribosomal RNA, partial seq	1314	1314	99%	0.0	95%
NR_036794.1	Klebsiella pneumoniae strain DSM 30104 16S ribosomal RNA, complete	1314	1314	99%	0.0	95%
NR_042356.1	Serratia ureilytica strain : NiVa 51 16S ribosomal RNA, partial sequenc	1310	1310	99%	0.0	95%
NR_037110.1	Serratia odorifera strain PADG 1073 16S ribosomal RNA, partial sequer	1310	1310	99%	0.0	95%
NR_025635.1	Klebsiella variicola strain F2R9 16S ribosomal RNA, partial sequence	1308	1308	99%	0.0	95%
NR_044980.1	Pectobacterium carotovorum strain LMG 2386 16S ribosomal RNA, par	1308	1308	99%	0.0	95%
NR_044976.1	Enterobacter nimipressuralis strain LMG 10245 16S ribosomal RNA, pa	1306	1306	99%	0.0	95%
NR_025159.1	Yersinia kristensenii strain ATCC 33638 16S ribosomal RNA, partial se	1304	1304	99%	0.0	95%
NR_027544.1	Yersinia frederiksenii ATCC 33641 16S ribosomal RNA, partial sequenc	1304	1304	99%	0.0	95%
NR_041831.1	Yersinia aldovae strain ATCC 35236 16S ribosomal RNA, partial sequer	1304	1304	99%	0.0	95%
NR_037112.1	Serratia proteamaculans strain 4364 16S ribosomal RNA, partial seque	1304	1304	99%	0.0	95%
NR_041978.1	Pantoea agglomerans strain DSM 3493 16S ribosomal RNA, partial seq	1304	1304	99%	0.0	95%

References

1. Balasuriya, BW Nileeka, and HP VasanthaRupasinghe. "Plant flavonoids as angiotensin converting enzyme inhibitors in regulation of hypertension." *Functional Foods in Health and Disease* 1.5 (2011): 172-188.

2. Danihelová, Martina, MiroslavVeverka, and Ernest Šturdík. "Inhibition of pathophysiological proteases with novel quercetin derivatives." *ActaChimicaSlovaca* 6.1 (2013): 115-122.

3. Kitchen, Douglas B., et al. "Docking and scoring in virtual screening for drug discovery: methods and applications." *Nature reviews. Drug discovery* 3.11 (2004): 935.

4. Chen, Kuan-Chung, and Calvin Yu-Chian Chen. "In Silico identification of potent PPAR-agonists from traditional Chinese medicine: a bioactivity prediction, virtual screening, and molecular dynamics study." *Evidence-Based Complementary and Alternative Medicine* 2014 (2014).

5. Rubio, Sara, et al. "Phenylbenzopyrones structure-activity studies identify betuletol derivatives as potential antitumoral agents." *European journal of pharmacology* 548.1 (2006): 9-20.

6. Chen, Kuan-Chung, and Calvin Yu-Chian Chen. "In Silico identification of potent PPAR-agonists from traditional Chinese medicine: a bioactivity prediction, virtual screening, and molecular dynamics study." *Evidence-Based Complementary and Alternative Medicine* 2014 (2014).

7. Cao, Tru, et al. "Advances in Knowledge Discovery and Data Mining 19th Pacific-Asia Conference, PAKDD 2015, Ho Chi Minh City, Vietnam, May 19-22, 2015, Proceedings, Part I." *Conference proceedings PAKDD*. 2015.

8. Manner, Suvi, et al. "Systematic exploration of natural and synthetic flavonoids for the inhibition of Staphylococcus aureus biofilms." *International journal of molecular sciences* 14.10 (2013): 19434-19451.

9. Cos, Paul, et al. "Structure– activity relationship and classification of flavonoids as inhibitors of xanthine oxidase and superoxide scavengers." *Journal of Natural Products* 61.1 (1998): 71-76.

10. Raj, U., et al. "In Silico Identification of Novel Glucagon Receptor Antagonist for the Treatment of Type 2 Diabetes Mellitus." *EnzEng* 4.128 (2015): 2.

11. Danihelová, Martina, et al. "Antioxidant action and cytotoxicity on HeLa and NIH-3T3 cells of new quercetin derivatives." *Interdisciplinary toxicology* 6.4 (2013): 209-216.

12. Taniguchi, Sayuri, et al. "Inhibition of heparin-induced tau filament formation by phenothiazines, polyphenols, and porphyrins." *Journal of Biological Chemistry* 280.9 (2005): 7614-7623.

13. Vuorinen, Anna, et al. "Ligand-based pharmacophoremodeling and virtual screening for the discovery of novel 17β-hydroxysteroid dehydrogenase 2 inhibitors." *Journal of medicinal chemistry* 57.14 (2014): 5995-6007.

14. Kimura, Yuka, et al. "Inhibitory effects of polyphenols on human cytochrome P450 3A4 and 2C9 activity." *Food and chemical toxicology* 48.1 (2010): 429-435.

15. Kandakatla, Naresh, and GeethaRamakrishnan. "Ligand based pharmacophoremodeling and virtual screening studies to design novel HDAC2 inhibitors." *Advances in bioinformatics* 2014 (2014).

16. Matter, William F., Raymond F. Brown, and Chris J. Vlahos. "The inhibition of phosphatidylinositol 3-kinase by quercetin and analogs." *Biochemical and biophysical research communications* 186.2 (1992): 624-631.

17. Bandele, Omari J., and Neil Osheroff. "Bioflavonoids as poisons of human topoisomerase IIα and IIβ." *Biochemistry* 46.20 (2007): 6097-6108.

18. Bandele, Omari J., and Neil Osheroff. "Bioflavonoids as poisons of human topoisomerase IIα and IIβ." *Biochemistry* 46.20 (2007): 6097-6108.

19. Feher, Miklos, and Jonathan M. Schmidt. "Identifying potential binding modes and explaining partitioning behavior using flexible alignments and multidimensional scaling." *Journal of computer-aided molecular design* 15.12 (2001): 1065-1083.

20. Choi, Moon-Hee, and Hyun-Jae Shin. "Anti-melanogenesis effect of quercetin." *Cosmetics* 3.2 (2016): 18.

21. Kruzlicova, Dasa, Martina Danihelova, and MiroslavVeverka."Quantitative structure-antioxidant activity relationship of quercetin and its new synthetised derivatives" *Nova BiotechnologicaetChimica* 11.1 (2012): 37-44.

22. Materska, Małgorzata. "Quercetin and its derivatives: chemical structure and bioactivity-a review." *Polish journal of food and nutrition sciences* 58.4 (2008).

23. Danihelová, Martina, et al. "Antioxidant action and cytotoxicity on HeLa and NIH-3T3 cells of new quercetin derivatives." *Interdisciplinary toxicology* 6.4 (2013): 209-216.

24. Choiprasert, W., et al. "Quercetin, quercetrin except rutin potentially increased pirarubicin cytotoxicity by non-competitively inhibiting the P-glycoprotein-and MRP1 function in living K562/adr and GLC4/adr cells." *American Journal of Pharmacology and Toxicology* 5.1 (2010): 24-33.

25. Chen, Kuan-Chung, and Calvin Yu-Chian Chen. "In Silico identification of potent PPAR-agonists from traditional Chinese medicine: a bioactivity prediction, virtual screening, and molecular dynamics study." *Evidence-Based Complementary and Alternative Medicine* 2014 (2014).

26. Vyas, Vivek, et al. "Virtual screening: a fast tool for drug design." *ScientiaPharmaceutica* 76.3 (2008): 333-360.

27. Bagger, JonatanIssing, et al. "Glucagon antagonism as a potential therapeutic target in type 2 diabetes." *Diabetes, Obesity and Metabolism*13.11 (2011): 965-971.

28. Quesada, Ivan, et al. "Physiology of the pancreatic α-cell and glucagon secretion: role in glucose homeostasis and diabetes." *Journal of Endocrinology* 199.1 (2008): 5-19.

29. Ganji, MostafaFathi, and Mohammad SanieeAbadeh. "Using fuzzy ant colony optimization for diagnosis of diabetes disease." *Electrical Engineering (ICEE), 2010 18th Iranian Conference on*. IEEE, 2010.

30. Cambon-Thomsen, Anne, Emmanuelle Rial-Sebbag, and Bartha Maria Knoppers. "Trends in ethical and legal frameworks for the use of human biobanks." *European Respiratory Journal* 30.2 (2007): 373-382.

31. Kwon, Christina S., Enrique Seoane-Vazquez, and Rosa Rodriguez-Monguio. "FDA safety actions for antidiabetic drugs marketed in the US, 1980–2015." *International Journal of Risk & Safety in Medicine* 28.4 (2016): 197-211.

32. Bennett, Wendy L., et al. "Comparative effectiveness and safety of medications for type 2 diabetes: an update including new drugs and 2-drug combinations." *Annals of internal medicine* 154.9 (2011): 602-613.

33. McNaught, Alan D., and Alan D. McNaught. *Compendium of chemical terminology*.Vol. 1669. Oxford: Blackwell Science, 1997.

Online content

https://www.google.co.in/search?q=virtual+screening&oq=vi&aqs=chrome.0.69i59l2j69i61l3j35i39.2184j0j7&sourceid=chrome&ie=UTF-8

https://www.google.co.in/search?q=pharmacophore+modelling+wikipedia&oq=pharmaco&aqs=chrome.1.69i57j35i39l2j0l3.4879j0j7&sourceid=chrome&ie=UTF-8

https://www.diapedia.org/management/8104277114/glucagon-antagonism.

http://www.medicalnewstoday.com/info/diabetes/diabetessymptoms.php

http://www.medicalnewstoday.com/info/diabetes

https://www.diabetesresearch.org/what-is-diabetes

http://www.diabetes.ca/about-diabetes/types-of
diabeteshttps://www.diabetes.org.uk/Diabetes-the-basics/Diabetes-Symptoms/

http://www.diabetes.co.uk/diabetes-causes.html

http://www.diabetes.co.uk/diabetes-causes.html

https://www.niddk.nih.gov/health-information/diabetes/overview/symptoms-causes

http://www.webmd.boots.com/diabetes/type-2-diabetes-guide/type-2-diabetes

https://www.cornerstones4care.com/about-diabetes/diabetes-basics/different-types-of-diabetes.html

http://www.diabetes.ca/about-diabetes/types-of-diabetes

http://www.diabetes.ca/about-diabetes/types-of- See more diabetes#sthash.sIAh3dxp.dpuf

http://www.medicinenet.com/diabetes_mellitus/page3.htm

https://www.usatoday.com/story/news/2016/04/06/diabetes-rates.../82695278/

https://nutritionfacts.org/topics/diabetes/?gclid=CPHmpdyMrtQCFdOFaAodCcgCXw

http://www.diabetes.co.uk/diabetes-causes.html

https://www.niddk.nih.gov/health-information/diabetes/overview/symptoms-causes

https://www.usatoday.com/story/news/2016/04/06/diabetes-rates.../82695278/

http://www.medicalnewstoday.com/info/diabetes

https://www.diabetesresearch.org/what-is-diabetes

http://www.webmd.boots.com/diabetes/type-2-diabetes-guide/type-2-diabetes

CURRICULUM VIATE

SUBHASHREE SUBHASMITA RAUT

2nd Year M. Sc (Bioinformatics)

Department of Bioinformatics

CPGS, OUAT, Bhubaneswar

Mobile: 9692635944

Email Id:subhashree.sraut@gmail.com

CAREER OBJECTIVE

To secure a challenging position in Research environment with committed people to contribute my Bioinformatics and Computer Skills in the field of biological research to serve our nation.

ACADEMIC QUALIFICATION

Discipline	Institution	Board/University	Year of Passing	% of Marks	Class
M. Sc.(Bioinformatics)	Dept. of Bioinformatics, CPGS	OUAT,BHUBANESWAR	Continuing	OGPA: 6.23 (10 Pont Scale)	
B.Sc.(Mathematics)	DAYA VIHAR DEGREE COLLEGE,KANAS	UTKAL UNIVERSITY	2015	66.056	1st
+2 Science	KANAS COLLEGE,KANAS	CHSE	2012	52.167	2nd
10th	GOVT. HIGH SCHOOL, KANAS	HSCE	2010	61.833	1st

BIOINFORMATICS SKILLS

Area	Skill
Bioinformatics Tools	BLAST, FASTA, EMBOSS, ClustalW, Expasy, ORF finder, MODELLER, PYMOL, DISCOVERY STUDIO VISUALIZER, PATCH DOCK, PSIPRED, MEGA, AutoDock, Gromacs 4.5
Bioinformatics Databases	NCBI, SWISS-PROT, EMBL, DDBJ, PDB, PROSITE, OMIM & OMIA

COMPUTER SKILLS

Area	Skill
Operating systems	Windows, Linux
Programing language	C, JAVA, PERL, HTML, PHP, .NET
Database language	SQL, MYSQL

AREA OF INTEREST

Pharmacogenomics, Genomics, Proteomics, Molecular Modeling, Computer Aided Drug Design, System Biology

SEMINAR/WORKKSHOP ATTENDED

- Participated in a National seminar on "'Microbial Technology: Prospects and Application'' held during 25th-26th December, 2015 jointly organized by the Department of Microbiology and students Associations of Microbiologists, OUAT, Bhubaneswar, Odisha.

PERSONAL INFORMATION

Name	: Miss Subhashree Subhasmita Raut
Date of Birth	: 13th July 1995
Gender	: Female
Marital Status	: Unmarried
Nationality	: Indian
Language Proficiency	: Odia, Hindi, English,
PERMANENT ADDRESS	Miss Subhashree Subhasmita Raut D/O: Mr.Sushilkanta Raut At: Nuapada Po: Kanas Dist.: Puri Pin: 752017 Odisha

DECLARATION:

I hereby declare that the above mentioned information is correct up to my knowledge and I bear the responsibility for the correctness of the above mentioned particulars.

Place: Bhubaneswar

Date:

Signature of the Candidate