ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Towards the Prediction of a Three-Dimensional Structure of Human GLP-1 Receptor

A Ranganadha Reddy[1]*, Sreedhara R Voleti[2], and Ch Lakshmi Padma[1]

[1]School of Biotechnology, Vignan University, Vadlamudi, Guntur 522213, Andhra Pradesh, India
[2]Institute of Life Sciences, University of Hyderabad, Hyderabad 500046, Andhra Pradesh, India

ABSTRACT

Glucose is the primary source of energy. Glucose levels in the body are monitored by various agents, such as insulin and glucagon. Glucagon Like Peptide-1 (GLP-1) is an incretin hormone which binds to the Glucagon Like Peptide -1 Receptor (GLP-1R), a Family-B GPCR, and helps in the maintenance of blood glucose homeostasis. In the face of demand for novel therapeutics to treat type -2 diabetes and the fact that the GLP-1R is a well-validated target, there is a clear need to develop more understanding of the regulation and function of this receptor. A complete crystal structure for GLP-1R is not available - thus, keeping in view regarding its role and importance in human health, the aim of this project is to build a complete homology model structure by making use of template based ab-initio and hybrid methodologies like PRIME, Bhageerath, and Bhageerath-H softwares.

Keywords: GLP-1R,Bhageerath, Bhageerath-H, Validation, SCHRÖDINGER

Corresponding author

ISSN: 0975-8585

INTRODUCTION

Diabetes mellitus, often simply referred to as diabetes, is a group of metabolic diseases in which a person has high blood sugar, either because the body does not produce enough insulin, or because cells do not respond to the insulin that is produced [1]. There are three main types of diabetes: Type-1 diabetes,Type-2 diabetes,Gestational diabetes.The search for new and effective therapies for Type-2 diabetes has led to the identification of a novel therapeutic target, the "incretin" hormones, which play a role in mediating glucose homeostasis via effects on glucagon and insulin secretion from pancreatic islet α- and β-cells, respectively. Glucagon-like peptide-1 (GLP-1) and Glucose-dependent Insulinotropic Polypeptide (GIP) are hormones secreted by the enteroendocrine cells of the gut in response to the ingestion of nutrients. These incretin hormones, so called because they increase insulin secretion, are key modulators of pancreatic islet hormone secretion and, thus, glucose homeostasis. Moreover GLP-1 decreases glucagon secretion. Studies demonstrating that incretin activity is impaired in type-2 diabetes have led to investigations into incretin-based new therapies such as incretin-mimetics, analogues of GLP-1 (i.e, a long-acting GLP-1 receptor agonist) as Exenatide, a long-acting GLP-1 analogue (liraglutide) and inhibitors of Dipeptidyl peptidase-IV (DPP-IV), a ubiquitous enzyme that rapidly inactivates both GLP-1 and GIP, increase active levels of these hormones and, in doing so, improve islet function and glycaemic control in type-2 diabetes[2]. GIP is a single 42 amino acid peptide derived from the processing of a 153 amino acid precursor, whose 10 kilo base-spanning gene is located on chromosome 17 in humans. GIP is secreted in a single bioactive form by K cells and released from the upper small intestine (duodenum and proximal jejunum), in response to the oral ingestion of carbohydrates and lipids. GLP-1 is a product of the proglucagon gene, spanning 10 kilobases and located on the long arm of chromosome 2, that encodes not only GLP-1 but also glucagon, GLP-2 and other proglucagon derived peptides. GLP-1 is extensively and rapidly degraded by the enzyme Dipeptidyl-peptidase 4 (DPP-4), that cleaves the biologically active forms at the position 2 alanine (N-terminal), resulting in inactive or weak antagonist peptide fragments. In a high glucose state, as shown in the Figure-1, stimulated insulin secretion by β-cells acts on the insulin receptor on the surface of α-cells and then suppresses glucagon secretion in a paracrine manner. In a low glucose state, decreased insulin secretion by β-cells is recognized by α-cells as a reduction of insulin signaling in α-cells through insulin receptors, and α-cells then increase glucagon secretion in response. GLP-1 directly suppresses glucagon secretion by α-cells through slight increase of cAMP followed by inhibition of N-type Ca^{2+} channels. GLP-1 also potentiates insulin secretion by β-cells and then suppresses glucagon secretion through insulin effects on α-cells. Glucose stimulates insulin secretion by β-cells and suppresses glucagon from α-cells through insulin effects, while glucose can stimulate glucagon secretion by α-cells.

Similar to native GLP-1, extraneous GLP-1R agonists control blood glucose through regulation of islet function, principally by the stimulation of insulin secretion and inhibition of glucagon secretion in a glucose-dependent manor. So, GLP-1R is found to be a valid novel target for the treatment of Type-2 Diabetes.

ISSN: 0975-8585

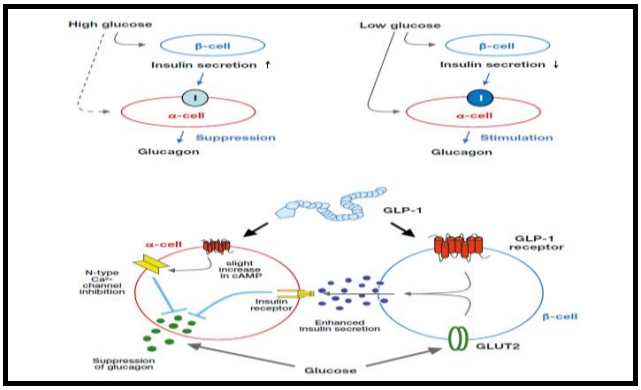


Figure-1: Regulation of glucagon and insulin secretions by GLP-1

According to a recent report by market analysis firm Navigant Consulting, "GPCRs (G Protein Coupled Receptors) are among the most heavily investigated drug targets in the pharmaceutical industry". These proteins are active in just about every organ system and present a wide range of opportunities as therapeutic targets in areas including cancer, cardiac dysfunction, diabetes, central nervous system disorders, obesity, inflammation, and pain. Consequently, GPCRs are prominent components of pipelines in small and large drug companies alike, and many drug discovery firms focus exclusively on these receptors [3]. Based on certain key sequences, GPCRs can be divided into three major subfamilies, receptors related to rhodopsin (type-A/Family-A), receptors related to the calcitonin receptor (type-B/Family-B), and receptors related to the metabotropic receptors (type-C/Family-C) [4]. The class of family-B GPCRs includes receptors for moderate-sized peptides that are involved in regulating important endocrine and neuroendocrine functions. In addition to the prototype member, the secretin receptor, structurally related mammalian class B GPCRs include the calcitonin and calcitonin receptor-like, corticotropin-releasing factor (CRF), gastric inhibitory peptide, glucagon, glucagon-like peptide, growth hormone releasing hormone, parathyroid hormone (PTH), pituitary adenylate cyclase activating peptide, and vasoactive intestinal polypeptide receptors. Although these GPCRs share a similar seven transmembrane domain topology with their family-A counterparts, there is virtually no amino acid conservation between the two groups of proteins. The relative paucity in current knowledge on ligand interaction and receptor activation of class B GPCRs is reflected by the difficulty in identifying synthetic, small molecule drugs that either mimic or block the function of endogenous peptide ligands[5].GLP-1R comes under Family-B GPCRs. This family of receptors is characterized by a relatively long N-terminal tail and a network of three conserved cysteine disulfide bridges which stabilize the N-terminal structure (Figure 2b). The present study will focus on the GLP-1R in this family.

As shown in the Figure-3, glucose signaling pathway results in the generation of ATP and an increase in the ATP: ADP ratio, and the GLP-1 receptor (GLP-1R)-mediates cAMP PKA pathways to effect closure of ATP-sensitive potassium channels (K-ATP) consisting of the inward rectifier Kir6.2 and the sulfonylurea receptor SUR1. The closure of these channels results in a rise in the resting potential (depolarization) of the β-cell, leading to opening of voltage-sensitive calcium channels (L-type VDCC). A major component of the depolarizing current is carried by NSCCs that import Na^+ (and Ca^{2+}). In response to activation of NSCC and influx of Na^+ there is import of Ca^{2+} by the Na^+/Ca^{2+}

exchanger (Na:Ca Exch). Release of intracellular membrane stores of calcium (Ca^{2+} stores) is induced by intracellular free Ca^{2+}, so called calcium-induced calcium release. The influx of Ca^{2+} through the open-end L-type VDCC triggers vesicular insulin secretion by the process of exocytosis. Phosphorylation of vesicular (granule) proteins by PKA may also trigger insulin secretion. Repolarization of the β-cell is achieved by opening of calcium-sensitive potassium channels (Ca-K). It is believed that the GLP-1 receptor is coupled to a stimulatory G-protein (Gs) and a calcium-calmodulin-sensitive adenylate cyclase[6].

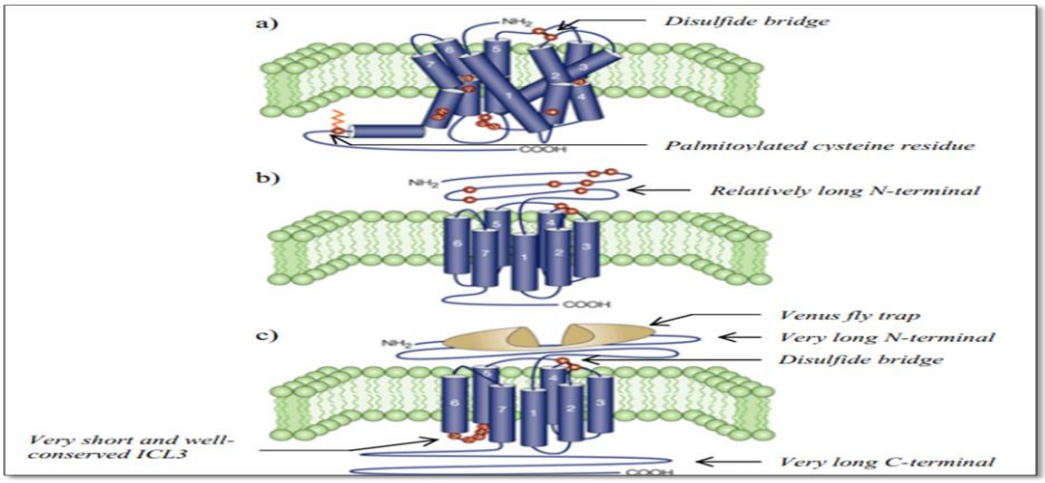


Figure-2: Molecular structures of GPCRs: Family-A, Family-B and Family-C

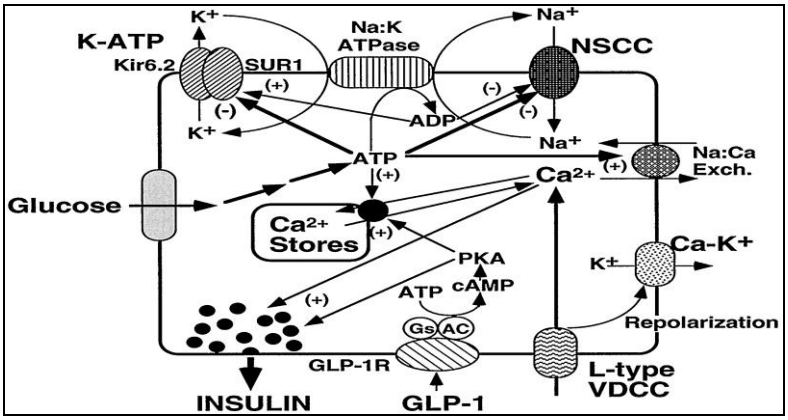


Figure-3: Mechanisms of insulin secretion in response to glucose and GLP-1

The GLP-1 receptor belongs to class B of the G-protein-coupled receptors, a sub Family-Characterized by a large N-terminal extracellular ligand binding domain. Exendin-4 and GLP-1 are 50% identical, and exendin-4 is a full agonist with similar affinity and potency for the GLP-1 receptor. Recently, the crystal structure of the GLP-1 receptor extracellular domain in complex with the competitive antagonist Exendin-4 (9-39) was crystallized. The isolated extracellular domain binds exendin-4 with much higher affinity than the endogenous agonist GLP-1.The crystal structure of the extracellular domain in complex with GLP-1 to 2.1Åresolution was solved [7]. The available partially crystal structures of GLP-1R are 3IOL, 3C5T, 3C59. Since N-Terminus of the receptor majorly involves in the

ISSN: 0975-8585

ligand binding and there is no completely crystallized structure for the GLP-1Receptor, homology modeling technique was employed to build the complete structure. The gene for the human GLP-1R is localized to the short arm of chromosome 6 (6p21) and encodes a sequence of 463 amino acids. Like other members of Family-B GPCRs, the GLP-1R is predicted by UniProt (http://expasy.org/uniprot/P43220) to contain a large hydrophilic, extracellular (N-terminal) domain of 145 residues with a putative signal peptide; seven hydrophobic transmembrane domains (TM1 to TM7) that are connected by hydrophilic extracellular and intracellular loops (ECL1, ECL2, ECL3 and ICL1, ICL2, and ICL3) and an intracellular (C-terminal) domain. The initial model of the amino terminal domain of the GLP-1R was generated using X-ray structure of amino terminus of this receptor complexed with the GLP-1 peptide (PDB ID 3IOL) as template. Extracellular amino terminal domain of GLP-1R is stabilized by Pairs of disulfide bonds linking six conserved cysteine residues and within ECL 1 also there are conserved cysteine residues which form a disulfide bond. Figure-4 represents a snake plot of the sequence of GLP-1R receptor. The disulfide bridges are among the residues C46-C71, C62-C104, C85-C126 and C226-C296 [8].

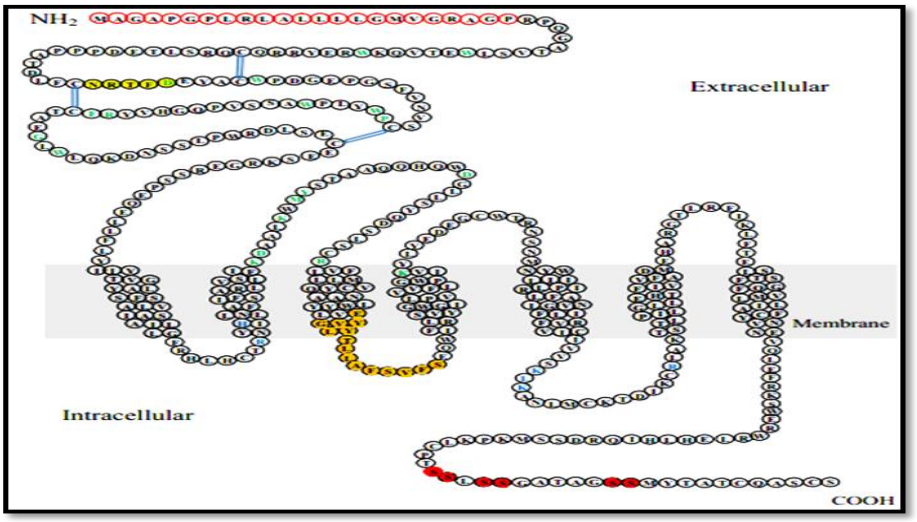


Figure-4: Snake plot of GLP-1R

Bhageerath is an Ab initio protein structure prediction Web interface server. It is developed in India at IIT Delhi by Prof. B. Jayaram's group; the web server predicts five native-like candidate structures for the protein, there is a limitation for the input sequence, query's length should be less than 100 amino acids. Bhageerath-H is hybrid Web interface server which takes use of both template and Ab initio calculations for the prediction of the three dimensional structure of a given proteins primary sequence. It is also developed by the Scfbio team of IIT Delhi; the web server predicts five native-like candidate structures for the protein. This Bhageerath-H also predicts 5 native like pictures for the query sequence. It does not have limitations for the input sequence, so any number of residues could be submitted to this software. Bhageerath and Bhageerath-H are the only software coming out of India, which predicts the protein folding of a given primary sequence. The validation of these two softwares was also incorporated into the project so that the results obtained from these servers could be used for the construction of the model [7][11].

ISSN: 0975-8585

COMPUTATIONAL TOOLS

All calculations were carried out in Maestro v9.2 installed in Cadd-WS3 machine under 64-bit centos operating system placed in CADD department, Institute of Life Sciences. The machine was built up with:

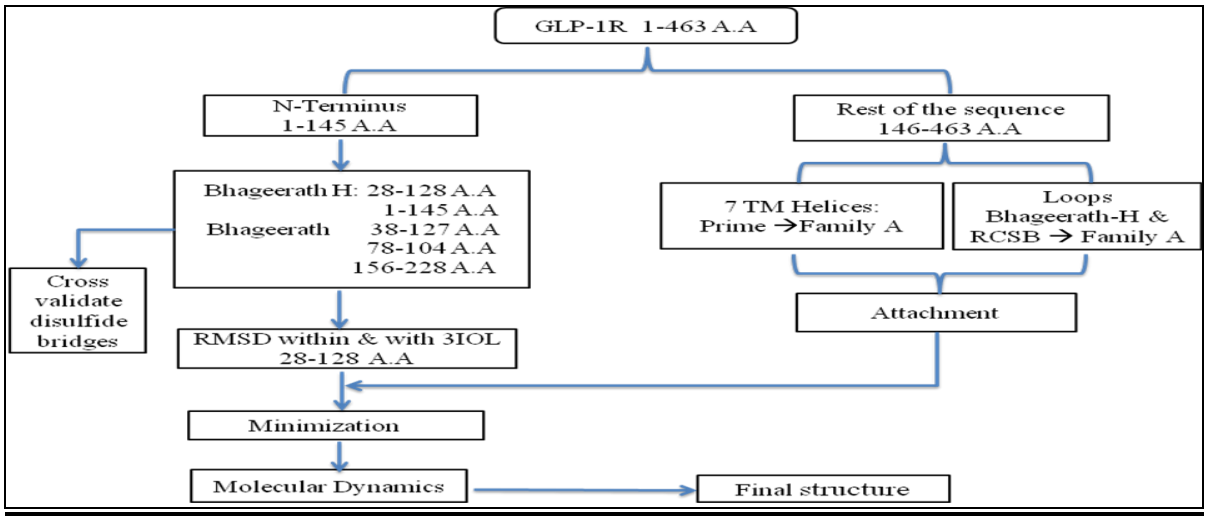
A) 4 cores and 8 processers with Intel Xenon CPU E5620 @ 2.40GHZ
B) 16 GB RAM
C) NVidia Qudvo FX3800 Graphical Process Unit (GPU)
D) The PROCHECK analysis provides an idea of the stereochemical quality of all protein chains in a given PDB structure.They highlight regions of the proteins which appear to have unusal geometry and provide an overall assessment of the structure as a whole.
E) Other Servers

- Primary sequence of the GLP-1R was retrieved from Swiss Prot (accession number P43220) from the ExPASy (Expert Protein Analysis System) proteomics serves of the Swiss Institute of Bioinformatics.
- Homology search for GLP-1R was carried out using BLAST software.

F) The crystal structure for GLP-1R (PDB ID: 3IOL) was obtained from PDB database RCSB.

PROCEDURE

To build a homology model of GLP-1 Receptor and to validate the Bhageerath and Bhageerath-H protein structure prediction softwares, the following methodology was followed as shown in the flowchart-1.



Flowchart-1: Workflow to achieve the project objectives

BLAST Searches

The Basic Local Alignment Search Tool (BLAST) finds regions of local similarity between protein or nucleotide sequences. The program compares nucleotide or protein sequences to sequence in a database and calculates the statistical significance of the matches. The target sequence i.e., GLP-1 Receptor (UniProt ID: P43220) was searched against the protein database by using BLAST tool. From the BLAST results, I observed that three

proteins (PDB IDs: 3IOL, 3C5T & 3C59) are showing the maximum identity with the target sequence. Among the three proteins obtained, 3IOL was selected for further proceedings.

Protein Preparation (PrepWiz)

A typical PDB structure consists of heavy atoms, waters, cofactors, metal ions and can be multimeric. The structure generally has no information on bond orders, topologies, or formal atomic charges. So, 3IOL (from the PDB) must be prepared by using protein preparation wizard (PrepWiz) of Schrödinger software. Protein preparation ensures that the 3IOL protein structure was properly assigned with bond orders and correct number of hydrogens to make the structure compatible with the OPLS (Optimized Potential for Liquid Simulations) forcefields [9].

The protein structure obtained and prepared was just a part of the N-Terminus - hence, in order to build the remaining structure GLP-1 Receptor was divided into 3 parts:

N-Terminus, Transmembrane helices & loops, and C-Terminus

In the process of building the homology model Bhageerath and Bhageerath H softwares were used.

Verifying Bhageerath-H

For the validation of Bhageerath-H a part of N-terminus sequence (28-128 amino acids) was submitted as input to Bhageerath-H. Five models were obtained as output and their RMSD (Root Mean Square Deviation) values with the query sequence (3IOL) were calculated. The model with low RMSD with query was selected.

Verification of Bhageerath

For the validation of Bhageerath a part of N-terminus sequence (38-127) was submitted as input. Five models were obtained as output and their RMSDs with respect to the query sequence were calculated. As the results were not found to be satisfactory, two more queries were submitted. One of the queries was from 156 to 248 amino acid residues of GLP-1R i.e., the end of first helix to the starting of third helix. Another query was from 78-104 amino acid residues i.e., the beta sheet part of the GLP-1R sequence was also submitted. The obtained results were analyzed.

N-Terminus Building

Since Bhageerath-H can take a query sequence of more than 100 amino acid residues and as the length of the N-terminus sequence is 145 amino acid residues, Bhageerath-H was selected for the building of N-terminus. Five models were obtained from Bhageerath-H of them one was selected.

Transmembrane Helices Construction

Since there is no completely crystallized structure of Family-B GPCR till date, transmembrane helices were constructed based on the crystal structures of Family-A GPCRs. For that some of the known Family-A GPCRs were collected.

List of the PDB IDs of Family-A GPCRs that were considered for the transmembrane generation of GLP-1R are:

1. Rhodopsin: 1F88, 1GZH, 1HZX, 1L9H, 1U19, 3DQB;
2. β2 Adrenergic: 2R4R, 2R4S, 2RH1, 3P06, 3KJ6, 3SN6, 3PDS, 3D4S, 3NY8, 3NY9, 3NYA, 3P0G;
3. Adenosine: 2YDV, 3RFM, 3EM2, 3QAK, 3PWH, 2RDD, 3REY;
4. CXCR4: 3ODU, 3OEO, 3OE8, 3OE9, 3OE6;
5. Tukey β1 adrenergic: 2VT4, 2YCW, 2YCX, 2YCY, 2YCZ;
6. β1 Adrenergic: 2Y00, 2Y01, 2Y02, 2Y03, 2Y04;
7. D3 receptor: 3BBL

Sequence alignment of the query sequence with the selected Family-A GPCRs was performed using the multiple sequence viewer tools in Schrödinger software. Based on the percentage identity and similarity with GLP-1R, 3POG, 3QAK, 3DQB were selected. The individual transmembrane helices were built based on the maximum identity shown among the three Family-A GPCRs (3DQB, 3POG, and 3QAK) using the PRIME module of Schrödinger software v9.2.

Loop Building

The FASTA sequence of all the intracellular and extracellular loops were submitted to Bhageerath-H, which generated five different models for each loop. Of all these models the best predicted loops were selected.

C-Terminus Building

For the construction of C-terminus, the amino acid sequence from 409 to 463 was give as an input to Bhageerath software, which generated five models of which one best model was selected.

Building the Final Model

The final model was built by connecting the N-terminus generated from Bhageerath-H, seven transmembrane helices generated by PRIME module, loops generated by Bhageerath-H and C-terminus generated from Bhageerath software, using the connect and fuse tool of Maestro software.

Loop Refinement

All extracellular and intracellular loops of the model were selected and subjected to loop refinement using PRIME module. Upon completion of loop refinement 5-different conformers for each loop would be generated, and the best loop structure were selected and incorporated into the model based on several parameters.

Structure Minimization

For energy minimization of the generated models, MacroModel module was utilized. The default parameters were used i.e. OPLS_2005 was selected as the forcefield, convergence threshold was set to 0.05 and 15000 iterations were specified. The job got completed immediately after the convergence had reached.

Molecular Dynamics

MacroModel dynamics was performed to the structure resulted from minimization. In the Molecular Dynamics all hydrogen atoms, ions, and water molecules were first subjected to 500 steps of energy minimization by steepest descent algorithm to remove undesired close vander Waals contacts. The temperature of the system was kept 323K. Time step value altered to 1.0 fs used in the integration of the equations of motion during the simulation. Equilibration time kept default value 1.0 ps used to determine the length of the settling down period at the start of the simulation. Number of structures to sample the receptor protein structure (GLP-1R) was given 100. The convergence of simulation was analyzed in terms of the potential energy gradient and the convergence threshold is 0.05 is a default value. Force field OPLS_2005 was used with a constant dielectric of 1.0. Solvent was used to run the protein system in solvent phase. The energy equilibrated molecular structure was computed upon completion of the Molecular Dynamics simulation, and the averaged structure was energy re-minimized to eliminate errors associated out of averaging exercise for obtaining the final model [10].

RESULTS AND DISCUSSION

Sequence Retrieval and Data Collection

The details of the protein GLP-1R like amino acid lengths of N-and C-terminus, helix lengths etc, were obtained from Protein knowledgebase UniProtKB/Swiss-Prot database with accession number P43220 (GLP-1R). The details of sequence annotation are shown in **Figure-5**.

Bhageerath-H

Of all those 5-models, one model showed maximum homology that maintained all the disulfide bonds as shown in the Figure 5 (A). RMSDs of the obtained models resulted from Bhageerath-H with the 3IOL (partially crystallized N-term structure of GLP-1R) were calculated. It is with a great degree of confidence, all models lie within 0.2 Å from the template 3IOL as shown in Figure (6B).

ISSN: 0975-8585

It was also found that the disulfide bridges as expected were conserved as shown in Figure-4. Since Bhageerath-H uses template (in this case, 3IOL), the expected results turned out to be the predicted outcome.

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
Signal peptide	1 – 23	23	Potential		
Chain	24 – 463	440	Glucagon-like peptide 1 receptor		PRO_0000012835
Regions					
Topological domain	24 – 145	122	Extracellular Potential		
Transmembrane	146 – 168	23	Helical; Name=1; Potential		
Topological domain	169 – 176	8	Cytoplasmic Potential		
Transmembrane	177 – 196	20	Helical; Name=2; Potential		
Topological domain	197 – 227	31	Extracellular Potential		
Transmembrane	228 – 252	25	Helical; Name=3; Potential		
Topological domain	253 – 264	12	Cytoplasmic Potential		
Transmembrane	265 – 288	24	Helical; Name=4; Potential		
Topological domain	289 – 303	15	Extracellular Potential		
Transmembrane	304 – 329	26	Helical; Name=5; Potential		
Topological domain	330 – 351	22	Cytoplasmic Potential		
Transmembrane	352 – 372	21	Helical; Name=6; Potential		
Topological domain	373 – 387	15	Extracellular Potential		
Transmembrane	388 – 408	21	Helical; Name=7; Potential		
Topological domain	409 – 463	55	Cytoplasmic Potential		
Amino acid modifications					
Modified residue	341	1	ADP-ribosylcysteine		
Modified residue	348	1	ADP-ribosylarginine		
Glycosylation	63	1	N-linked (GlcNAc...) Potential		
Glycosylation	82	1	N-linked (GlcNAc...) Potential		
Glycosylation	115	1	N-linked (GlcNAc...) Potential		
Disulfide bond	46 ↔ 71		Ref.11 Ref.13		
Disulfide bond	62 ↔ 104		Ref.11 Ref.13		
Disulfide bond	85 ↔ 126		Ref.11 Ref.13		
Disulfide bond	226 ↔ 296		Ref.11 Ref.13		

Figure-5: Structural details of GLP-1R with residue numbering from the Swissprot Sequence Database

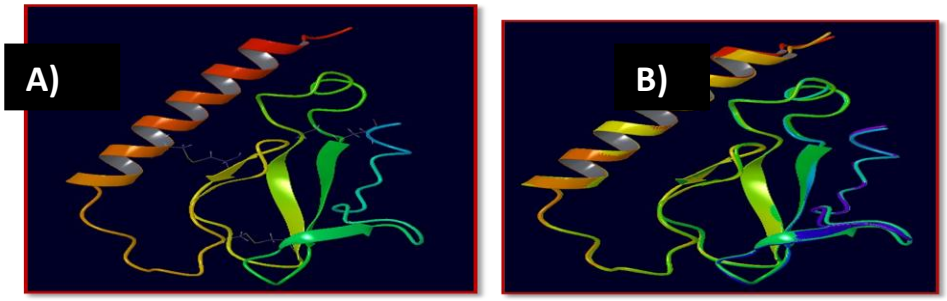


Figure-6: (A) The best model obtained from Bhageerath-H (B) Superposition of that model on 3IOL

Bhageerath

For the verification of Bhageerath software a 38-127 amino acid residues of 3IOL were submitted (38-127 instead of 28-128 due to the limitation of handling of number of residues in Bhageerath as input). Since Bhageerath doesn't use any template, the expected ab initio protein folding from first principles would be less close to the crystal structure template, 3IOL. As observed, the obtained results (five models, as shown in Figure-7) are far from crystal structure, although some encouraging results bring confidence in use of Bhageerath as shown in table 1

ISSN: 0975-8585

Figure-7: The superposition of all the Bhageerath results with 3IOL

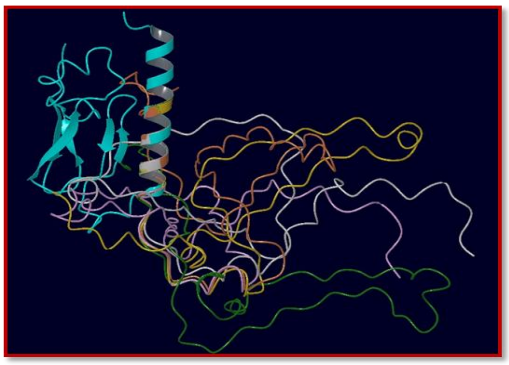


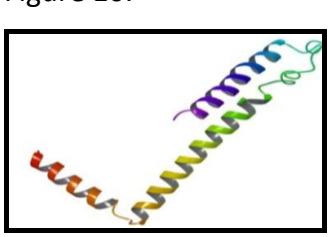
Table 1: The RMSDs of the structures obtained from Bhageerath with 3IOL

Models obtained from Bhageerath	RMSD (Å) with 3IOL
Model-1	18.7
Model-2	19.5
Model-3	17
Model-4	18.4
Model-5	18.5

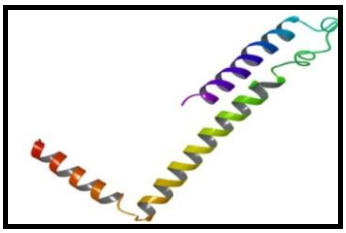
Figure-7: The superposition of all the Bhageerath results with 3IOL The resulted models did not maintain the disulfides which were expected to be maintained. Although, the RMSD's for predicted structures are far from the crystal structure, all models correctly predict the helical part of the sequence structure. To investigate deeper, we further submitted a helix-loop-helix-loop-helix structure sequence of the same protein (GLP-1R) shown as below (Figure-8). To our surprise, predictions of tertiary structures were spot-on for the given primary sequence, although the models vary with respect to fold patterns giving us the current success of the Bhageerath in prediction of protein-structure folding.

Figure-8: The sequence of helix loop helix loop helix part of the TMs submitted to Bhageerath

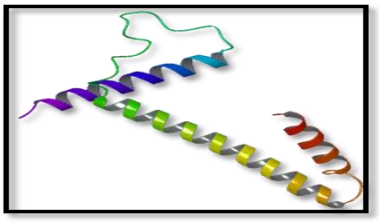
The 'helix-loop-helix-loop-helix' part of the given query was predicted pretty well. The results obtained are in good agreement with respect to RMS deviation as shown in the Figure 9 and Figure 10.



Model 1

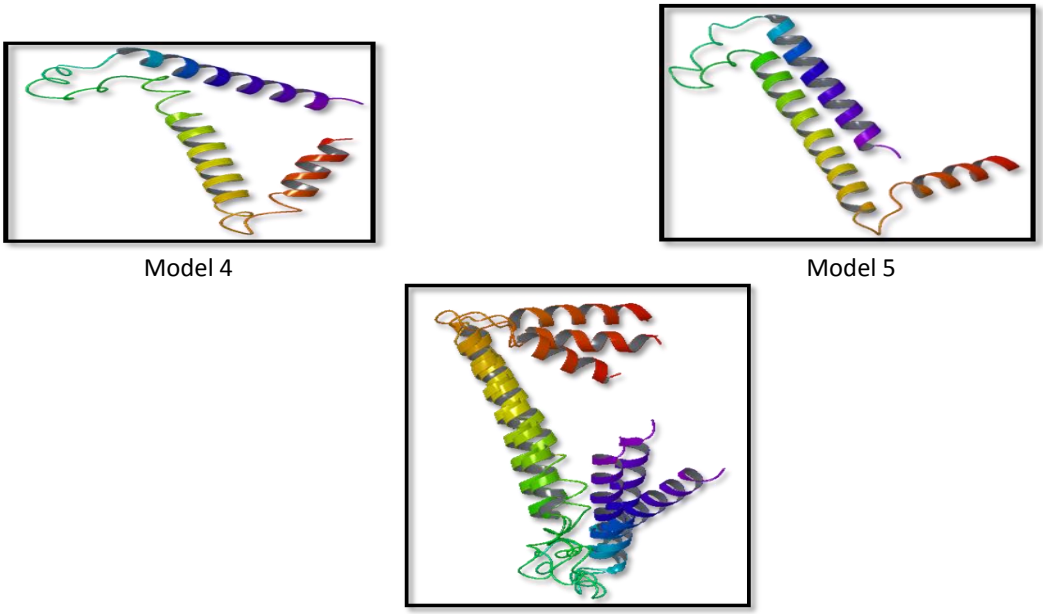


Model 2



Model 3

ISSN: 0975-8585



Model 4 Model 5

Figure-10: Superposition of obtained models

Figure- 9: Five models obtained from Bhageerath We then submitted few sequence lengths from crystal structure of 3IOL as β-sheets for a predictive fold. The sheet part of the N-terminus structure (27 amino acid sequence length i.e., from 78-104 amino acid residues) of the GLP-1R was submitted. The amino acid sequence that was submitted to Bhageerath is shown in Figure 11.

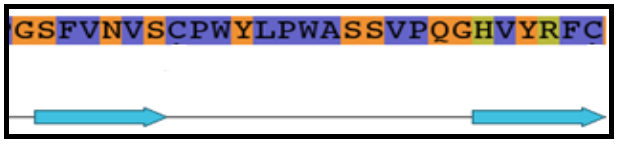
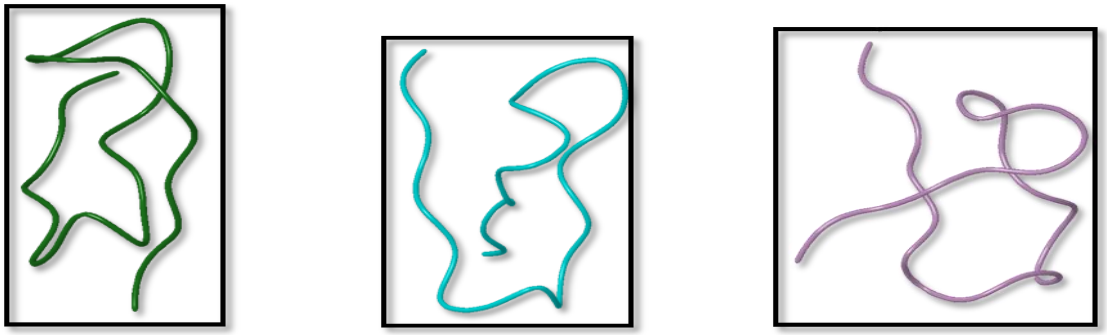


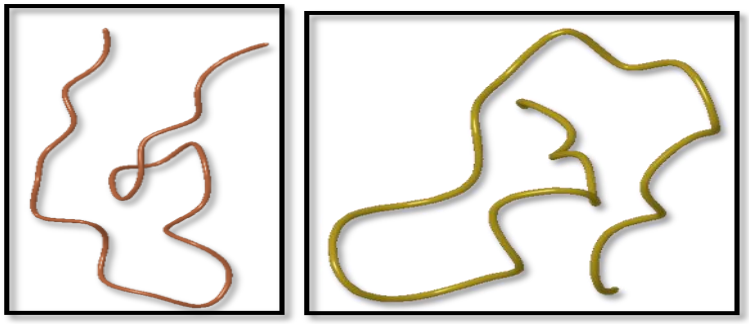
Figure-11: Sequence of 3IOL submitted for the sheet prediction in Bhageerath

The sheet part present on the N-terminus of the GLP-1R was not predicted well by the Bhageerath software. The models obtained were like loops instead of sheets (Figure-12). With a hope, we expected whether molecular dynamics could convert these loops into sheets. Hence, a molecular dynamics of 10ns was implemented in MacroModel module of Schrodinger in gas phase with no constrains. To our interest, we found that there is no change in the secondary structure from loops to sheet for the model.



Model-1 Model-2 Model-3

ISSN: 0975-8585



Model 4 Model 5

Figure-12: Five models obtained from Bhageerath protein structure prediction software

Sequence Alignment - Transmembrane Modeling

The most significant part of homology modeling process is to obtain the correct sequence alignment of target sequence with the homologues. The sequence alignment of GLP-1R with the three family-A GPCRs i.e., 3DQB, 3POG & 3QAK was performed as shown in the Figure-13 and the sequence identities and similarities for each helix position of the GLP-1R with Family-A GPCRs were tabulated in the table 2.

Table 2: Percentage identities and similarities of individual transmembrane helices of GLP-1R with family-A GPCRs

	TM-1	TM-2	TM-3	TM-4	TM-5	TM-6	TM-7
3QAK	%I = 12 %S = 22	%I=38 %S=50	%I=8 %S=25	%I=16 %S=24	%I=32 %S=49	%I=8 %S=24	%I=8 %S=27
3P0G	%I=16 %S=28	%I=12 %S=22	%I=19 %S=25	%I=12 %S=28	%I=24 %S=43	%I=26 %S=39	%I=8 %S=27
3DQB	%I=9 %S=22	%I=28 %S=34	%I=11 %S=19	%I=4 %S=20	%I=3 %S=19	%I=5 %S=21	%I=27 %S=38

The transmembrane helices were build on the basis of the maximal identity as shown in Figure-14. Individual transmembranes were built based on the maximum identity shown among the three Family-A GPCRs.

Loops generated by Bhageerath-H were joined to the helices generated, through connect and fuse option. The joined loops were again set for the refinement. Prime-Refinement is capable of refining loop structures of various lengths, and provides algorithms for different loop lengths. In addition, loops whose structure affects other loops can be cooperatively refined in pairs. All extracellular and intracellular loops are selected and subjected to loop refinement in GLP-1R. When all loops have been refined, one final structure is returned which is shown in Figure-15.

ISSN: 0975-8585

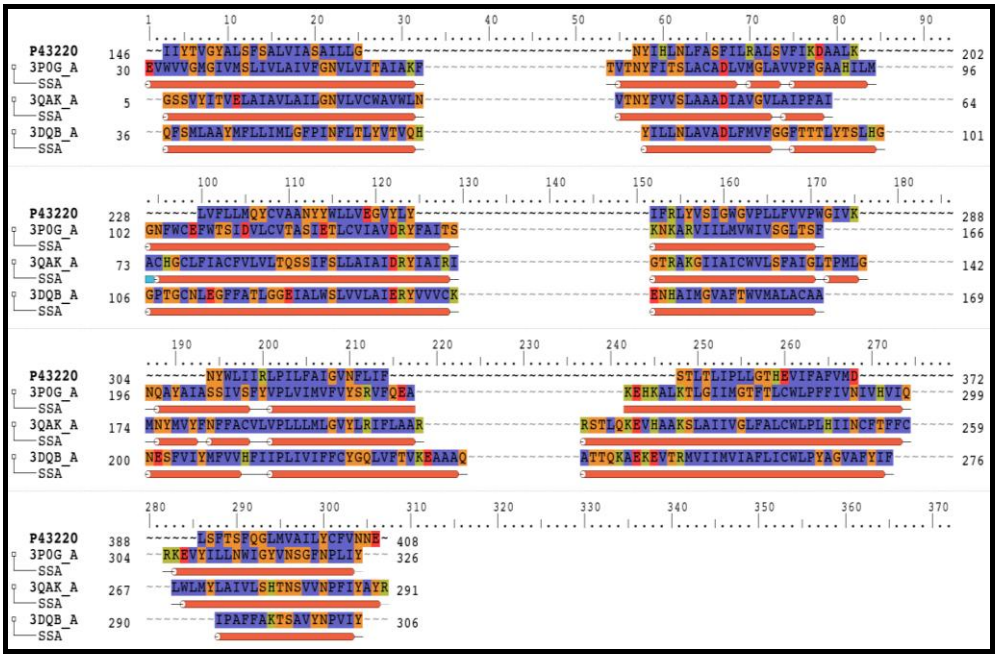


Figure-13: Sequence alignment of transmembrane regions of 3POG, 3QAK & 3DQB against GLP-1R transmembrane regions

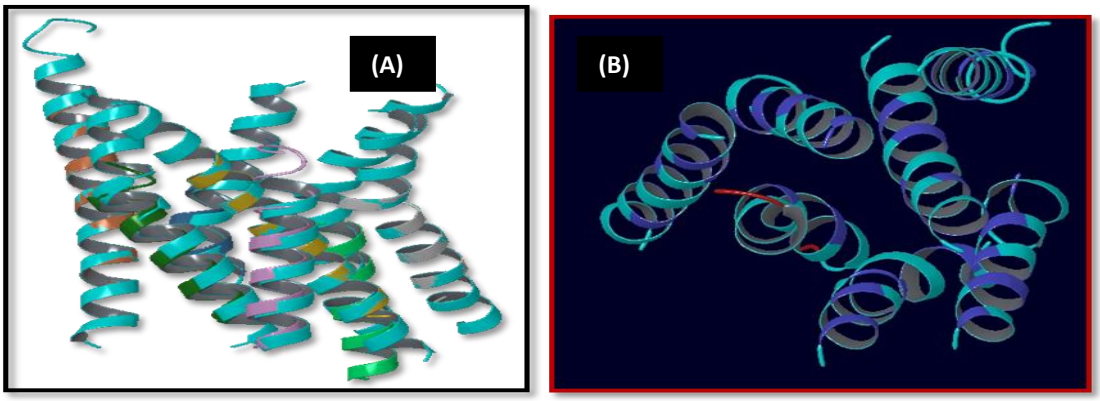


Figure-14: (A) Membrane view of helices, (B) Cellular views of the helices Joining of Loops to Helices

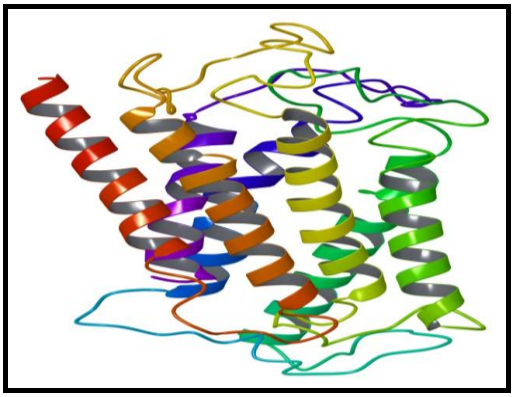


Figure-15: Transmembrane helices with loops attached Joining of N- and C- Terminals Models to the Model

The N-terminus generated by Bhageerath-H and the C-terminus predicted by Bhageerath were attached to the model with helices and loops through the connect and

fuse option in the tool bar. The characteristic structural feature of the GPCRs is the presence of disulphide bonds between two adjacent cysteine residues on different transmembrane helices or loops are treated using protein preparation wizard. The amino acid residues which are involved in extra cellular, intra cellular loops, helix sequence and disulphide linkage etc details were obtained from Protein knowledgebase UniprotKB/Swiss-prot database. According to the obtain data, the cysteine residues which are involved in disulphide linkage are 46 →71, 62→104, 85→126 and 226→296 in GLP-1R, as shown in Figure-16. .

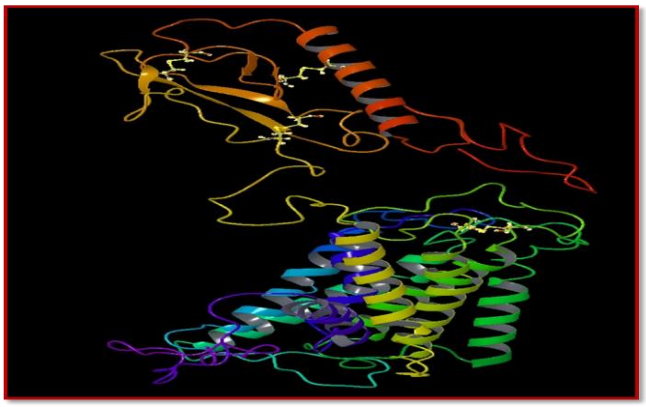


Figure-16: Transmembrane helices and N-terminus

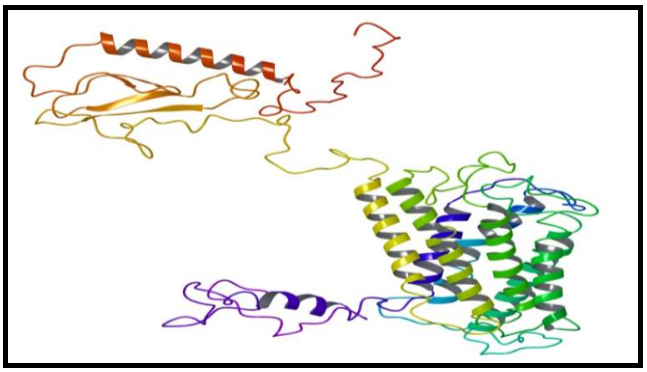


Figure-17: Final structure after molecular dynamics

Analysis of the Molecular Dynamics (MD) Simulation

MD simulation of GLP-1R was carried in order to check the stability of the model and to equilibrate the structure to get converged atoms and side chains within dynamic environmental conditions. Analysis of 10ns dynamics showed that the GLP-1R structure was stable and equilibrated. Molecular dynamics in the Macro model uses a standard constant temperature velocity-Verlet algorithm. SHAKE procedure constraints selected bond lengths to their original values. This procedure allows the use of larger time steps than unconstrained simulation. SHAKE all bonds in system with 500 steps of energy minimization by steepest descent algorithm to remove close van der Waals contacts. The parameters which are used in MacroModel molecular dynamics simulations are sufficiently used in equilibrating the models showing bad contacts and irregular loops as well as non equilibrated N- and C- terminals. The Figure-17 shows models before and after

ISSN: 0975-8585

the Molecular dynamics simulations using Macro model module of Schrodinger, version 9.2. The results and the output structures of MD simulations clearly show that the atoms on back bone and side chains are converged with sufficient equilibration time.

Assessment of the Homology Model of GLP-1R

The validation of the final model was carried out using Ramachandran plot computed with PROCHECK, program by checking the detailed residue-by-residue stereo-chemical quality of a protein structure. The PROCHECK is used for stereochemical assessment of the model.

The criteria for analysis of stereochemistry of the model includes,

1) Main chain conformation in acceptable regions of the Ramachandran plot.
2) Planar peptide bonds.
3) Side chain conformations that correspond to those in rotamer library.
4) Hydrogen bonding of polar atoms if they are buried.
5) No bad atom-atom contacts.
6) No holes inside the structure.

Ramachandran Plot

A Ramachandran plot (also known as a Ramachandran map or a Ramachandran diagram or a [Φ,Ψ] plot), developed by Gopalasamudram Narayana Ramachandran and Viswanathan Sasisekharan is a way to visualize dihedral angles Ψ and Φ of amino acid residues in protein structure. It shows the possible conformations of Φ and Ψ angles for a polypeptide. Hence, Ramachandran plot is a useful way of assessing the stereo chemical quality of a protein structure. From the main Ramachandran plot developed by PROCHECK our receptor protein GLP-1R had 78.9% of 463 amino acids in the core region of the Ramachandran plot. Residues in additionally allowed regions are 17.7%, residues in generously allowed region are 2.6%, and residues in disallowed region are 0.7%. Ramachandran plots for the obtained homology model structure of protein GLP-1R before and after molecular dynamics are shown in Figure-18 and Figure-19, respectively

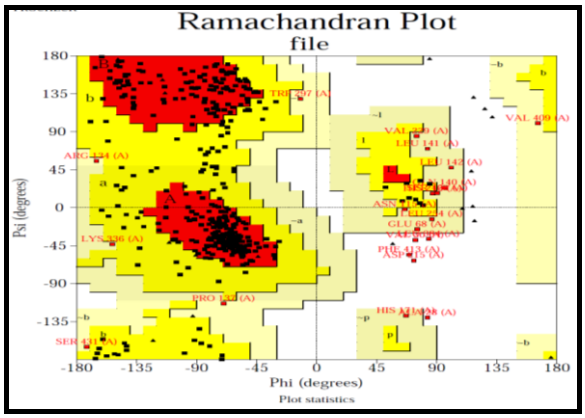


Figure-18: Ramachandran plot before dynamics

ISSN: 0975-8585

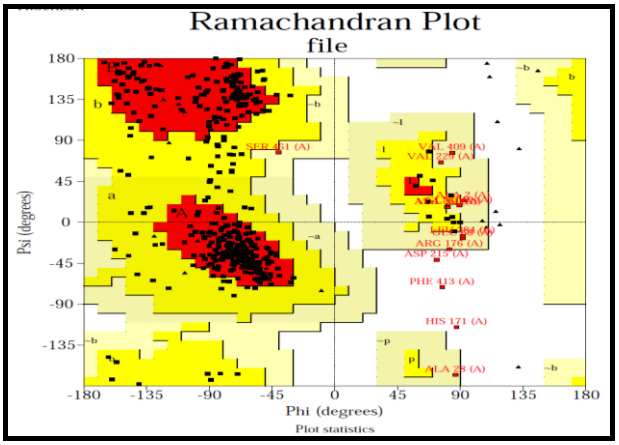


Figure-19: Ramachandran plot after dynamics

CONCLUSIONS

The 3-dimensional structure of Human GLP-1Receptor (463 amino acids) have been attempted in parts through stitching different models. The N-terminus part of GLP-1Receptor i.e, 1-145 amino acid residue length structure had generated from the hybrid protein structure prediction software Bhageerath-H, in which all disulphide linkages (46→71, 62→ 104 and 85→ 126) were maintained.

The individual transmembrane helices were built based on the maximum identity shown among the three Family-A GPCRs (3DQB, 3POG and 3QAK) using the PRIME module of Schrödinger software v9.2. The disulfide bond in the transmembrane (226→ 296) was also maintained.

The C-terminus structure (409-463 amino acids) was developed by Bhageerath- ab initio protein structure prediction software. All these parts (N-terminus, C-terminus and transmembrane helices) of the GLP-1Receptor were joined using the PRIME module.

The built structure was subjected to Macro Model minimization (until it reached 0.05 gradience) and the resultant structure was subjected to MD (for 1ns) to remove the bad contacts derived from homology modeling and to achieve a good structure. After the molecular dynamics the model was stabilized, and desired the 4-disulfide bonds among the desired cystines that needed to maintain, have stayed well.

ACKNOWLEDGEMENT

We are very greatful to Institute of Life Sciences, Hyderabad central University,Hyderabad for providing us laboratory facilities.

REFERENCES

[1] www.diabetesinformationhub.com
[2] M Virally, JF Blickle J Girard, S Halimi, D Simon, PJ Guillausseau , Diabetes Metab 2007; 33: 231–244

[3] David filmore. J Am Chem Soc 2004; 7: 24-28
[4] Ulrik Gether and Brian K. Kobilka. J Biol Chem 1998; 273: 17979 –17982
[5] Martin Beinborn. Mol Pharmacol 2006; 70: 1-4.
[6] Timothy James Kieffer and Joel Francis Habener. Endocr Rev 1999;20: 876- 913.
[7] http://www.scfbio-iitd.res.in/bhageerath/bhageerath_h.jsp
[8] Christina Rye Underwood, Patrick Garibay, Lotte Bjerre Knudsen, Sven Hastrup, Günther H. Peters, Rainer Rudolph and teffen Reedtz-Runge. J Bio Chem 2010; 285: 723-730
[9] Laurence J. Miller, Quan Chen, Polo C.-H. Lam, Delia I. Pinon, Patrick M. Sexton, Ruben Abagyan, and Maoqing Dong. J Biol Chem. 2011; 286: 15895– 15907
[10] MacroModel, version 9.7, Schrödinger, LLC, New York, NY, 2009
[11] Ch Lakshmi Padma, Sreedhara R. Voleti, A. Ranganadha Reddy, R.K.Kishore J Chem Bio Phy Sci Sec C 2013; 3(2): 1388-1394.