RESEARCH ARTICLE

Effect of hydrophobic and hydrogen bonding interactions on the potency of ß-alanine analogs of G-protein coupled glucagon receptor inhibitors

Pushyaraga P. Venugopal | Bratin K. Das | E. Soorya | Debashree Chakraborty [ORCID]

Biophysical and Computational Chemistry Laboratory, Department of Chemistry, National Institute of Technology Karnataka, Surathkal, Mangalore, India

Correspondence
Debashree Chakraborty, Department of Chemistry, National Institute of Technology Karnataka, Surathkal, Mangalore, Karnataka 575025, India.
Email: debashree@nitk.edu.in

Peer Review
The peer review history for this article is available at https://publons.com/publon/10.1002/prot.25807.

Abstract

G-protein coupled glucagon receptors (GCGRs) play an important role in glucose homeostasis and pathophysiology of Type-II Diabetes Mellitus (T2DM). The allosteric pocket located at the trans-membrane domain of GCGR consists of hydrophobic (TM5) and hydrophilic (TM7) units. Hydrophobic interactions with the amino acid residues present at TM5, found to facilitate the favorable orientation of antagonist at GCGR allosteric pocket. A statistically robust and highly predictive 3D-QSAR model was developed using 58 β-alanine based GCGR antagonists with significant variation in structure and potency profile. The correlation coefficient (R^2) and cross-validation coefficient (Q^2) of the developed model were found to be 0.9981 and 0.8253, respectively at the PLS factor of 8. The analysis of the favorable and unfavorable contribution of different structural features on the glucagon receptor antagonists was done by 3D-QSAR contour plots. Hydrophobic and hydrogen bonding interactions are found to be main dominating non-bonding interactions in docking studies. Presence of highest occupied molecular orbital (HOMO) in the polar part and lowest unoccupied molecular orbital (LUMO) in the hydrophobic part of antagonists leads to favorable protein-ligand interactions. Molecular mechanics/generalized born surface area (MM/GBSA) calculations showed that van der Waals and nonpolar solvation energy terms are crucial components for thermodynamically stable binding of the inhibitors. The binding free energy of highly potent compound was found to be −63.475 kcal/mol; whereas the least active compound exhibited binding energy of −41.097 kcal/mol. Further, five 100 ns molecular dynamics simulation (MD) simulations were done to confirm the stability of the inhibitor-receptor complex. Outcomes of the present study can serve as the basis for designing improved GCGR antagonists.

1 | INTRODUCTION

Type 2 diabetes mellitus (T2DM) is known to be a chronic metabolic disorder, which is mainly caused due to improper lifestyle and genetics. Rapid growth of T2DM and lack of proper medication for this disease has become one of the real problems in recent days. The medical condition is characterized by elevated hepatic glucose level in the blood due to deregulated signal transduction by G-protein coupled glucagon receptor (GCGR) in hepatocytes. T2DM is also reported to increase the risk of diabetic-related complications such as weight loss, blindness, kidney failure, amputations, and cardiovascular diseases.

Glucagon is a small peptide hormone consists of 29 amino acid residues,[1] secreted by the α-cells of pancreatic islets and known to activate GCGR. In normal conditions, the attenuation of insulin inhibitory effect and release of glucagon from pancreas, increase the hepatic glucose level in blood by glycogenolysis during fasting. The combined activities of both insulin and glucagon have a crucial role in glucose homeostasis in the human body. Glucagon activated hepatic glucagon receptors (GCGR) transduces the activation of adenylate cyclase and initiates cAMP (cyclic adenosine monophosphate) production. This process ultimately, ends up with the expression of enzymes responsible for gluconeogenesis and glycogenolysis.[2] Therefore, blocking the activation of GCGR is believed to be an efficient way to control the abnormal hepatic glucose production in T2DM patients. Despite of several medications available, the development of new improved therapeutics has been hindered due to lack of structural

details of GCGR. Recently, the publications of X-ray crystal structure of GCGR provided an opportunity to design improved therapeutics by both structure and ligand based drug design for this receptor class.[1,3]

GCGR is a 62 kDa protein which belongs to class B family of G-protein coupled receptor (GPCR) superfamily. They are mainly found in the liver and kidney cells. However, they are also expressed in heart, adipose tissue, spleen, and adrenal glands. GCGR consists of an extracellular domain (ECD) and a trans-membrane domain (TMD) with a stalk region connecting both the domains. The ECD consists of a common α-β-β-α fold similar to the ECD of other class B GCGRs.[3] The TMD features the canonical seven trans-membrane helical bundles (TM1-TM7) of G-proteins.[3] Many small molecule antagonists/ inhibitors with varying potency and structural features have been reported as GCGR antagonists and these inhibitors range from glucagon neutralizing antibodies to small molecular antagonists. Recently a novel allosteric pocket outside the seven transmembrane domains have been reported which provides a scope to design improved therapeutics or antagonists against hyperglycemia. According to literature, some small molecule antagonists accommodate themselves to the extra-helical allosteric site located at the external surface of TMD.[1] The TM6-helix of TMD is found to divide the binding site into two regions: One is a hydrophobic cleft toward TM5 and the other is a polar end toward TM7.[1,4] A schematic representation of GCGR allosteric pocket with a hypothetical antagonist is illustrated in Figure 1. During allosteric inhibition, the conformational changes at the stalk region of GCGR found to alter the relative orientation of the ECD, TMD, and cause the inactivation of this serpentine receptor. Small
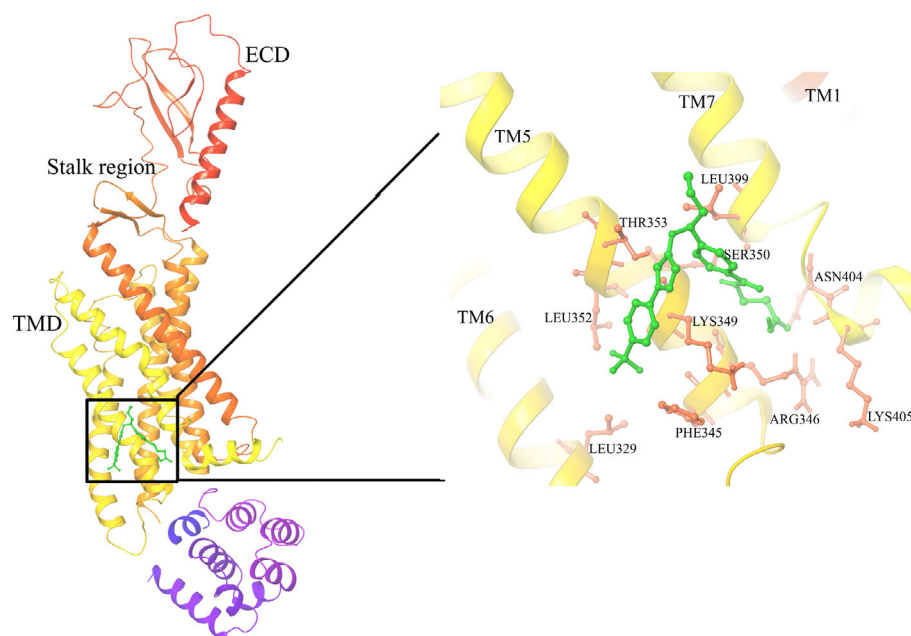
FIGURE 1 Secondary-structure representation of human GCGR (PDB ID: 5XEZ) with a hypothetical ligand showing the trans-membrane domain (TMD), the stalk region and the extracellular domain (ECD). An enlarged view of the extra-helical ligand-binding site of GCGR with the major amino acid residues at the catalytic pocket was shown next to it [Color figure can be viewed at wileyonlinelibrary.com]

molecule glucagon receptor antagonism is found to be useful in diabetic therapy. Although several publications have been reported on highly potent small molecule antagonists with desirable selectivity,[5-7] only a few have been entered to the clinical trials. Therefore, there is a dire need to develop a predictive biological model comprising of structurally diverse GCGR inhibitors based on GCGR crystal structure to improve the efficacy and safety of GCGR selective inhibitors.

Virtual screening (VS) is known to be a powerful in-silico approach to filter a large number of small molecules for new hits with desired properties, which can be subjected to experimental testing. Among the VS approaches, 3D-QSAR is a ligand-based method to correlate chemical structure and biological property by a statistically significant regression technique. Recently, the evolution of 3D-QSAR method helped significantly to screen a large number of data sets comprising of diverse scaffolds using advanced machine learning techniques.[8] Additionally, structure-based molecular docking is a well-known technique to probe the interaction of small molecules at the catalytic pocket of an enzyme. Therefore, docking is an important tool for characterizing the behavior of drug candidate and elucidating the fundamental biological processes.[9] The behavior of complex at the electronic level can also be explained through density functional theory (DFT) calculations.[10-13] Along with the information of drug-receptor interactions, it is important to quantify the binding energy of such complexes. A range of computational approaches including free energy perturbation (FEP), thermodynamic integration (TI), linear interaction energies (LIE), molecular-mechanics generalized born surface area (MM/GBSA), molecular-mechanics Poisson-Boltzmann surface area have been adopted to estimate free energies. Among these methods, MM-GBSA approach evolved to be a widely used method to compute free energy of binding of a protein-ligand complex. In addition to the above-mentioned VS strategy, molecular dynamics simulation is believed to be a crucial tool for confirming the stability of a drug candidate inside the binding site of a receptor.

In this present work, a series of experimentally tested pyrazole ethers and aminopyrazole compounds of ß-alanine were subjected to 3D-QSAR and molecular docking to screen the active antagonists, which can inhibit GCGR significantly. Further, we carried out DFT calculation to find out the preferred HOMO-LUMO distribution of the ligand and the binding site of the protein. Free energy calculations were performed for all the inhibitors and the protein complex to find out the deciding factor contributing to the stability of the complex. Moreover, five independent 100 ns MD simulations were carried out to see the dynamics of the most potent inhibitor at the protein-binding site. We also performed ADME/Toxicity calculation to confirm safe administration of top-scored hits into the human body. To the best of our knowledge, such combined in-silico study to investigate the crucial counterparts of allosteric inhibitors of GCGR is presented here for the first time. The rest of the paper has been organized as follows. In Section 2, we discussed the methodology and simulation protocol implemented. The result obtained from the calculation are discussed and illustrated in Section 3. Finally, the important achievements and conclusion drawn from this study are highlighted in Section 4. This study provides insightful information on the crucial structural features required to develop potential antagonists of GCGR.

2 | MATERIALS AND METHODS

2.1 | Dataset and preparation of 3-D structures of ligand

In the present study, a dataset consisting of 58 ß-alanine based glucagon receptor antagonist were selected from recent experimental reports.[14-16] The structural details of the inhibitors considered in this study are illustrated in Figure 2. All the drug candidates of the dataset reported to shared same assay procedure. The experimental inhibitory constant K_i was converted into pK_i (negative logarithm of K_i) for the ease of further analysis. The 3D structures of the glucagon receptor antagonists were constructed using the builder panel in Maestro graphical user interface (GUI). The partial charges were ascribed and possible ionization states were generated at pH 7.4 to mimic the experimental assay condition. Further, the geometry of the resulted structures was
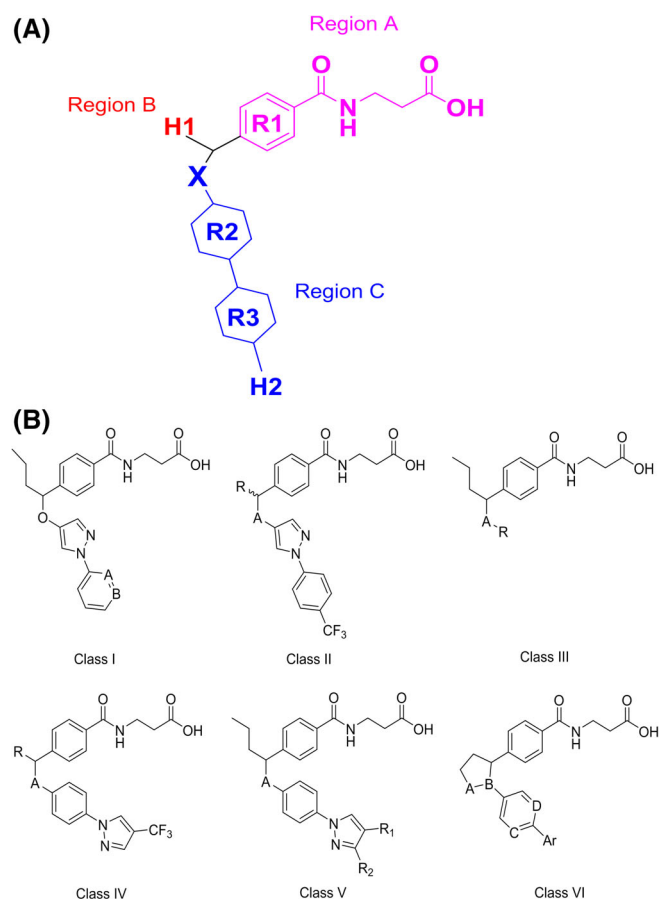
FIGURE 2 A, Skeletal structure of glucagon receptor antagonist: Region A represented in pink color is the polar region containing R1 ring, Region B represented in red color is the alkyl side chain which is hydrophobic in nature (H1) and Region C represented in blue color is the hydrophobic core containing R2, R3 ring, and H2 side group. B, Detailed structure of six classes of GCGR antagonists [Color figure can be viewed at wileyonlinelibrary.com]

optimized by semi-empirical PM3[17] and then by B3LYP/6-31G(d,p) level,[18,19] respectively, using Gaussian09 package.[20] The resulted structures were then used for computational studies.

2.2 | 3D-QSAR modeling

In the present study, Phase[21] module of Schrödinger was employed to develop 3D-QSAR model of β-alanine based GCGR inhibitors. Atom-based 3D-QSAR model are reported to be efficient in explaining true structure activity relationship rather than pharmacophore based 3D-QSAR model. Atom-based QSAR model considers ligand feature beyond the pharmacophoric sight thereby enabling to predict possible steric clashes with the receptor. Prior to 3D-QSAR modeling, all the ligands were aligned using flexible shape-based alignment tool in Phase module. The entire dataset was divided into training-set and test-set constituting of 44 and 14 compounds, respectively (based on standard 3:1 ratio), using "Automated Random Selection"[22] option in Phase. Care was taken to include the most active and inactive molecules in the training set.[21,23] A statistically significant model was generated by using Partial Least Square regression method with a grid spacing of 1 Å. The optimal PLS factor was taken as 8 (N/5, where N is the number of molecules in the training set), as the use of higher factor leads to over-fitting of data.[24] 3D contour plots were analyzed for understanding the effect of spatial arrangement of structural features at ligand sites on glucagon receptor antagonism. Further, the accuracy of the developed 3D-QSAR model in predicting the biological activity was validated by an external test set.

2.3 | Molecular docking procedure

The co-crystal structure of full-length GCGR with a negative allosteric modulator (NNC0640; PDB ID: 5XEZ, resolution: 3 Å)[3] was retrieved from RCSB Protein Data Bank. Prior to docking the 3D structure of 5XEZ were refined with Prime[25] and missing atoms were added. Appropriate ionization was confirmed by adding hydrogen bond corresponding to pH of 7.4. Automated software, Autodock (v4.2.6) was employed to dock pyrazole ether and aminopyrazole derivatives at the allosteric binding site of the protein. Autodock GUI[26] was used to prepare the protein coordinate suitable for docking procedure. The protein was prepared by removing water, membrane lipids and cocrystallized ligands. Gasteiger charges[27] were added. A 3D grid was prepared with a dimension of 30 Å × 46 Å × 30 Å having a spacing of 0.375 Å at the allosteric pocket located to the TMD of 5XEZ using Auto-grid program.[26] Lamarckian Genetic Algorithm[28] was used and the runs were set to 100 in order to search all the possible ligand-binding conformations at the allosteric pocket. Molecular docking was performed with an initial population of 150 and a number of 2.5×10^6 energy evaluations were carried out. The reliability of docking was confirmed by measuring RMSD[29] between the co-crystal and re-docked ligand. The best-docked conformations of individual inhibitors at the 5XEZ allosteric pocket were retained for further analysis.

2.4 | DFT calculation

Single point energy (SPE) calculations using self-consistent field (SCF) approach[30] were performed using Gaussian09 package[20] to explain antagonist bound receptor in electronic level. The structures were optimized using B3LYP level,[31,32] 6-31G(d,p) basis set[19] and then energies were determined using SPE calculations. The positions of HOMO-LUMO orbitals of selected molecules based on their biological activity were analyzed to study the binding interaction at quantum level. The HOMO-LUMO energy gap indicates to the chemical reactivity of molecules. The HOMO-LUMO densities over the binding site residues was analyzed to study the ligand-binding mechanism. The N and C-terminals of the amino acid residues were capped with N-acetyl group and N-methyl amide group, respectively, using Protein Preparation Wizard: Maestro.[33] The SPE calculation for the amino acid residues Leu329, Phe345, Arg346, Lys349, Ser350, Leu352, Thr353, Leu399, Asn404, and Lys405 were performed using B3LYP level and 6-31G(d,p) basis set. The cube files for visualizing HOMO and LUMO orbitals were created by using Cubegen utility in Gaussian09 software.

2.5 | MM/GBSA calculation

Computational methods comprising molecular mechanics energy and implicit solvent methods are known to be a widely used technique in free energy calculation. Their performance and applicability have been reported in several protein-substrate systems.[34,35] These methods are reported to be computationally efficient methods to calculate binding free energies.[34] The binding free energy of the docked ligands inside the trans-membrane binding pocket of the complexes was calculated by employing molecular-mechanics/generalized born surface area (MM/GBSA) approach,[34] incorporating OPLS_2005 force field.[36] An implicit membrane was placed surrounding the TMD of GCGR using default options of prime and local optimization sampling algorithm. Simulations were carried out using VSGB solvation model[37] (dielectric constant, ϵ = 80) with input ligand partial charges. The absolute free energy of the receptor-antagonist complexes was calculated; next the separate free energies of individual receptor and ligand were computed using an OPLS_2005 force field to determine binding affinity of the antagonists according to MM/GBSA approach. This provides a quantitative comparison of the binding strength of the glucagon receptor antagonists to the GCGR allosteric site.

2.6 | Molecular dynamics simulation protocol

Atomistic molecular dynamics simulation was employed to confirm the stability of highly active antagonist (**Compound 20**) at the allosteric pocket located at the TMD of GCGR. The MD simulation of complex **20** with GCGR (PDB ID: 5XEZ) was carried out with OPLS_2005 force field[36] in explicit solvent SPC (simple point-charge) water model[38] using Desmond software. POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) membranes was properly placed by

defining trans-membrane residues from M137 to W418.[3] The entire system was solvated in a periodic orthorhombic box with 20 Å buffer region between protein atoms and box sides to fill with water. All the surface and interior water molecules near the protein and the membrane were automatically removed by the system builder module of Desmond. The volume of the protein-ligand complex was found to be 1665354 Å^3. The system was neutralized by adding counter ions. The total number of atoms in the solvated system was 151 379 including 8362 number of atoms of complex 5XEZ/**Compound 20**. The system was minimized to a gradient threshold of 25 kcal/mol Å^{-1} using Steepest Descent algorithm and the iteration steps during minimization were kept as 2000 until a convergence threshold of 1.0 kcal/mol/Å was attained. For long-range electrostatic interactions, smooth Particle Mesh Ewald method[39] was used with a tolerance of $1 \times e^{-9}$ and for short-range electrostatic interactions a cut-off radius of 9.0 Å was applied. Reversible reference system propagator algorithm (RESPA) integrator algorithm[40,41] was applied with time steps of 2 fs for bonded, 2 fs for "near" nonbonded and 6 fs for "far" nonbonded interactions. A 5 ns MD run with NVT ensemble was carried out to equilibrate the system at 300 K with a time step of 2 fs. Noose-Hoover thermostat[42] was chosen to maintain the system temperature and thermostat relaxation time was kept at the interval of 200 ps, with a time step of 2 fs. Next, NPT equilibration was performed for 5 ns with a time step of 2 fs at 300 K and 1 atm using Noose-Hoover thermostat[40] (thermostat relaxation time = 200 ps) and Martyna-Tobias-Klein barostat[43] (barostat relaxation time = 200 ps). During the equilibration phase the heavy atoms of 5XEZ and **Compound 20** were restrained. Lastly, five independent 100 ns production MD (removing restrain on solute heavy atom) were carried out using NPT ensemble (T = 300 K, thermostat relaxation time = 200 ps; P = 1 atm, barostat relaxation time = 200 ps). The velocity and trajectory data of the simulation was retrieved at every 20 ps (number of frames = 5000) and visualization of the 3-D structures and the trajectories were done by using Maestro GUI.

2.7 | ADME/toxicity prediction

The evaluation of pharmacokinetic properties of top-scored hits obtained from a VS protocol is believed to be an important step of in-silico drug discovery process. It is known that the major concern for failure of drug candidates in clinical trial is poor pharmacokinetics.[44] Therefore, inclusion of ideal pharmacokinetic properties in previous stages of drug discovery is crucial to pass the clinical trial easily. With this aim, the absorption, distribution, metabolism, excretion, toxicity (ADMET) properties and different physically significant descriptors of the top scored inhibitors of GCGR obtained from the present study were predicted using Qikprop module (Schrödinger Release 2018-4: QikProp, Schrödinger, LLC, NY, 2018). The reliability of such predictions has already been reported and benchmarked.[45] Qikprop program employs the method of Jorgensen to predict the pharmacokinetics properties of drug-like molecules.[46] The drug-likeness of the top scored GCGR antagonists was also evaluated using Lipinski's rule of five.[47] Other physiochemical properties such as QPlogBB, QPlogK$_{hsa}$, QPPCaco, and QPPMDCK predict the permeability and binding at blood brain barrier, human serum albumin, Caco-2 cell, and so on.
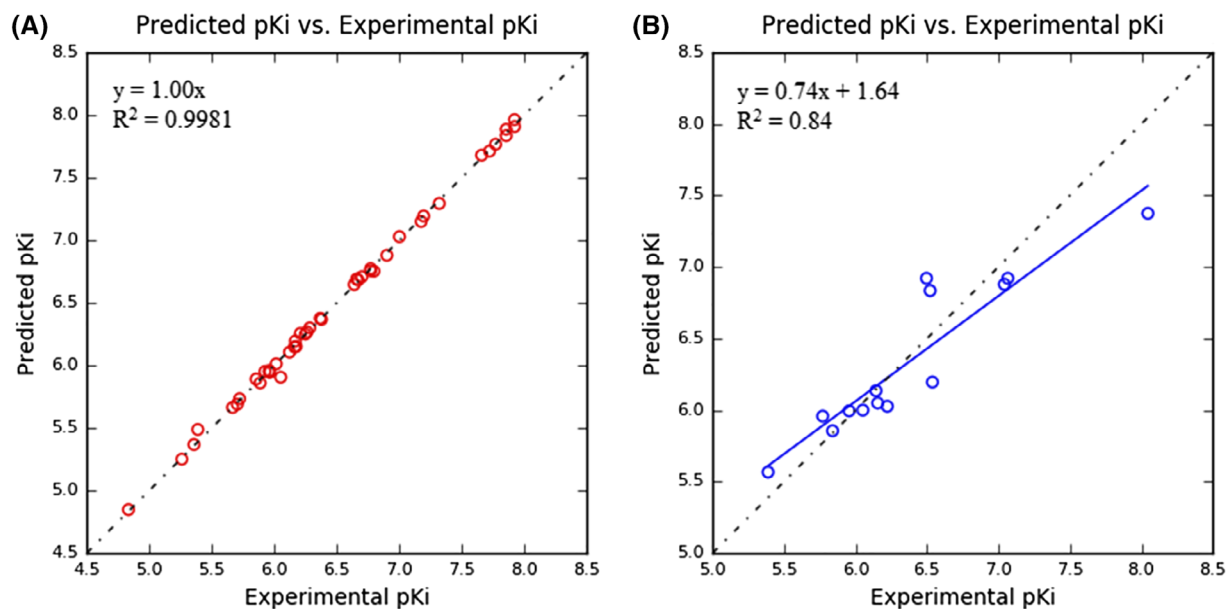
3 | RESULTS AND DISCUSSION

3.1 | 3D-QSAR analysis

In this study, a 3D-QSAR model was developed using 44 training set and 14 test set compounds of GCGR antagonists which yields optimum statistics in terms of correlation coefficient (R^2, measures internal consistency) and cross-validation coefficient (Q^2, measures internal predictability). The value of cross-validation coefficient (Q^2) for test set compounds and regression coefficient (R^2) for training set compounds are found to be 0.8253 and 0.9981, respectively. The model was developed with a PLS factor of 8 in order to avoid the risk of over-fitting of data and to achieve a significant statistical correlation between experimental activity and predicted activity. PLS regression statistics of generated 3D-QSAR model shown in Table 1, exhibited good statistical stability. The greater confidence of the model is indicated from the high Pearson-r value of 0.9176 and F value of 2336.6 with smaller P value (2.19×10^{-45}). Further, an acceptably low standard deviation (SD) value of 0.0376 and root mean square error (RMSE = 0.27) indicates the predictability and reliability of the generated model. The scatter plots of experimental pK_i values

TABLE 1 PLS regression summary of generated 3D-QSAR model

PLS	SD	R^2	F	P	Stability	RMSE	Q^2	Pearson-r
1	0.5064	0.5935	61.3	9.68×10^{-10}	0.667	0.38	0.6540	0.8149
2	0.2730	0.8847	157.3	5.84×10^{-20}	0.296	0.29	0.8009	0.9114
3	0.1613	0.9607	326.0	3.88×10^{-28}	0.235	0.30	0.7897	0.8973
4	0.0908	0.9879	793.2	8.93×10^{-37}	0.164	0.29	0.7909	0.9002
5	0.0725	0.9925	999.7	3.21×10^{-39}	0.154	0.28	0.8076	0.9097
6	0.0609	0.9948	1185.6	1.01×10^{-40}	0.144	0.27	0.8216	0.9172
7	0.0497	0.9966	1527.3	1.51×10^{-42}	0.150	0.27	0.8249	0.9189
8	0.0376	0.9981	2336.6	2.19×10^{-45}	0.152	0.27	0.8253	0.9176

Abbreviations: F, variance ratio; P, significance level of variance ratio; Pearson-r, square of correlation coefficient for test set; Q^2, cross validated correlation coefficient for test set; R^2, regression coefficient; RMSE, root mean square error; SD, SD of regression.

(A) Predicted pKi vs. Experimental pKi

$y = 1.00x$
$R^2 = 0.9981$

(B) Predicted pKi vs. Experimental pKi

$y = 0.74x + 1.64$
$R^2 = 0.84$

FIGURE 3 Scatter plot between the experimental activity (pK_i) vs 3D-QSAR predicted activity. A, The training dataset (unfilled red circles). B, The test dataset (unfilled blue circle). The best fitted equation for the scatter plot of test set compounds is given as $y = 0.74x + 1.64$ ($R^2 = 0.84$) [Color figure can be viewed at wileyonlinelibrary.com]

vs phase predicted values (Figure 3) for training set and test set compounds showed a strong linear correlation and Phase predicted activities were documented in Table 2. In Figure 2, we showed the skeletal structure of the inhibitors.

The efficiency of the model was validated by predicting the activity of 15 external test set compounds.[6,7] The predicted activity of external test set is tabulated in Table S1. The scatter plot of experimental activity vs predicted activity of external test set (Figure S1-A) showed a good linear correlation with a R^2 value of 0.83. It is believed that a 3D-QSAR model with R^2 greater than 0.5 have good predictability and reliability of the generated model.[48] The prediction errors of developed 3D-QSAR model found to be distributed randomly near the zero line (Figure S1-B) which denotes the absence of systematic errors due to biased calculation.

3.2 | Contour plot analysis

The effect of spatial arrangements of structural determinants on GCGR inhibition is analyzed by visualizing 3D-QSAR contour plots. The nature and position of substitution groups are found to have crucial role in defining the activity of chosen antagonists. It is evident from the present study that the hydrophobic/nonpolar groups, hydrogen bond donor groups and electron-withdrawing group have major contribution to the final 3D-QSAR model. The positive contributions are shown in blue cubes and the negative contributions are indicated by red cubes. For better visualization, the favorable and unfavorable interactions for each feature are mapped over **Compound 20** and illustrated in Figure 4. The contour plots of the compounds discussed below are shown in Figure S2.

3.2.1 | Effect of hydrophobic core

The hydrophobic character of β-alanine analogs is found to play an important role in GCGR inhibition. Figure 4A illustrates the favorable positions of hydrophobic groups on R2, R3, H1, and H2 regions of the inhibitors. Among the class I compounds, **Compound 12** (pK_i = 6.9) and **Compound 11** (pK_i = 6.67) are found to have high potency due to the substitution of hydrophobic —CF_3 group[49] and ethyl group at 4′ position of R3 ring, respectively. Among class IV compounds, the acyclic hydrophobic substituents at H1 site found to increase the biological activity than their cyclic analogs. For example, **Compound 40** (pK_i = 6.769) with tertiary-butyl group is found to be more potent than **Compound 37** (pK_i = 6.155) having cyclobutyl group. Similarly, **Compound 36** (pK_i = 6.143) with *iso*-propyl group is found to be more potent than **Compound 35** (pK_i = 5.886) having cyclo-propyl group. The replacement of heteroatoms with —CH and —CH_2 group at the cyclic linkage of class VI inhibitors (**Compound 58**, pK_i = 7.854) enhances the hydrophobic character which contributes to its higher potency.

3.2.2 | Effect of hydrogen bond donor group
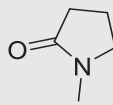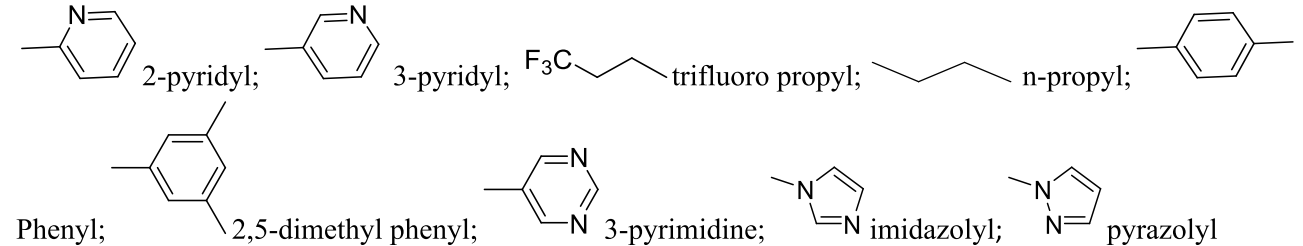
The blue cubes indicate the favorable position of hydrogen bond donor groups (Figure 4B). The presence of amino linkage (—NH) instead of ether group (—O—) as X at hydrophobic region found to favor the molecule to be highly active as antagonists. This explains the higher biological activity of aminopyrazole derivatives (**Compound 20**, pK_i = 8.046; **Compound 17**, pK_i = 7.721; **Compound 33**, pK_i = 7.921) than the pyrazole ether derivatives (**Compound 13**, pK_i = 6.367; **Compound 31**, pK_i = 7.194).

TABLE 2 Detailed structure, experimental activity (Exp pK_i), predicted activity (Pred pK_i), and residual activity of glucagon receptor antagonists

Cpd	X	H1	R2	R3	H2	Exp pK_i	Pred pK_i	Residual activity
Class I: Pyrazole ether series of inhibitors								
1	O	n-propyl	pyrazolyl	Phenyl	2—Me	4.833	4.846	0.013
2	O	n-propyl	pyrazolyl	Phenyl	3—OMe	5.666	5.662	−0.004
3	O	n-propyl	pyrazolyl	Phenyl	4—OMe	5.703	5.689	−0.014
4[a]	O	n-propyl	pyrazolyl	Phenyl	4—CF_3	5.839	5.854	0.015
5	O	n-propyl	pyrazolyl	Phenyl	4—CF_3	5.924	5.951	0.027
6[a]	O	n-propyl	pyrazolyl	Phenyl	3—Me	5.955	5.993	0.038
7	O	n-propyl	pyrazolyl	Phenyl	3—Cl	5.963	5.944	−0.019
8	O	n-propyl	pyrazolyl	Phenyl	4—Me	6.014	6.011	−0.003
9[a]	O	n-propyl	pyrazolyl	Phenyl	4—Cl	6.222	6.025	−0.197
10	O	n-propyl	pyrazolyl	Phenyl	3—Et	6.255	6.249	−0.006
11	O	n-propyl	pyrazolyl	2-pyridyl	4—Et	6.670	6.681	0.011
12	O	n-propyl	pyrazolyl	3-pyridyl	4—CF_3	6.900	6.878	−0.022
Class II: R-substituted pyrazole ethers and aminopyrazoles series								
13	O	n-propyl $_{(+)}$	pyrazolyl	Phenyl	4—CF_3	6.367	6.375	0.008
14[a]	O	n-propyl $_{(−)}$	pyrazolyl	Phenyl	4—CF_3	7.041	6.878	−0.163
15	NH	n-propyl $_{(−)}$	pyrazolyl	Phenyl	4—CF_3	7.174	7.148	−0.026
16	NH	trifluoro propyl $_{(−)}$	pyrazolyl	Phenyl	4—CF_3	7.319	7.293	−0.026
17	NH	n-propyl $_{(+)}$	pyrazolyl	Phenyl	4—CF_3	7.721	7.711	−0.010
18	NH	cyclopentyl$_{(−)}$	pyrazolyl	Phenyl	4—CF_3	7.770	7.765	−0.005
19	NH	trifluoro propyl $_{(+)}$	pyrazolyl	Phenyl	4—CF_3	7.921	7.907	−0.014
20[a]	NH	cyclopentyl$_{(+)}$	pyrazolyl	Phenyl	4—CF_3	8.046	7.374	−0.672
Class III: Ether and amino derivatives with different hydrophobic core								
21	O	n-propyl	Pyrimidine	Pyrazolyl	4—CF_3	5.959	5.948	−0.011
22	O	n-propyl	Phenyl	Pyrazolyl	4—CF_3	6.051	5.999	−0.052
23	NH	n-propyl	Phenyl	Imidazolyl	4—CF_3	6.169	6.193	0.024
24	NH	n-propyl	2-pyridyl	Pyrazolyl	4—CF_3	6.174	6.150	−0.024
25	O	n-propyl	3-pyridyl	Pyrazolyl	4—CF_3	6.638	6.645	0.007
26[a]	O	n-propyl	Phenyl	Imidazolyl	4—CF_3	6.658	6.689	0.031
27	NH	n-propyl	3-pyridyl	Pyrazolyl	4—CF_3	6.670	6.772	0.102
28	O	n-propyl	3-MePh	Pyrazolyl	4—CF_3	6.699	6.707	0.008
29	NH	n-propyl	Phenyl	Pyrazolyl	4—CF_3	6.796	6.751	−0.045
30	NH	n-propyl	3-pyrimidine	Pyrazolyl	4—CF_3	7.000	7.027	0.027
31	O	n-propyl	3,5-diMePh (-S)	Pyrazolyl	4—CF_3	7.194	7.192	−0.002
32	O	n-propyl	3,5-diMePh (+R)	Pyrazolyl	4—CF_3	7.854	7.835	−0.019
33	NH	n-propyl	3,5-diMePh	Pyrazolyl	4—CF_3	7.921	7.963	0.042
Class IV: Ethers substituted with different alkyl side chains								
34	O	Ethyl	Phenyl	Pyrazolyl	4—CF_3	5.854	5.889	0.035
35	O	cyclo-Propyl	Phenyl	Pyrazolyl	4—CF_3	5.886	5.858	−0.028
36[a]	O	iso- Propyl	Phenyl	Pyrazolyl	4—CF_3	6.143	6.134	−0.009
37[a]	O	cyclo-Butyl	Phenyl	Pyrazolyl	4—CF_3	6.155	6.048	−0.107
38	O	iso- Butyl	Phenyl	Pyrazolyl	4—CF_3	6.260	6.266	0.006
39	O	cyclo-Pentyl	Phenyl	Pyrazolyl	4—CF_3	6.377	6.366	−0.011
40	O	tert-Butyl	Phenyl	Pyrazolyl	4—CF_3	6.769	6.754	−0.015

TABLE 2 (Continued)

Cpd	X	H1	R2	R3	H2	Exp pK_i	Pred pK_i	Residual activity
Class V: Ethers with substitution on heterocyclic ring								
41	O	*n*-propyl	Phenyl	Pyrazolyl	3—C(O) NMe$_2$	5.187	5.249	0.062
42	O	*n*-propyl	Phenyl	Pyrazolyl	3—OMe	5.357	5.366	0.009
43	O	*n*-propyl	Phenyl	Pyrazolyl	H	5.387	5.486	0.099
44[a]	O	*n*-propyl	Phenyl	Pyrazolyl	3—CN	5.387	5.565	0.178
45	O	*n*-propyl	Phenyl	Pyrazolyl	3—Me	5.721	5.732	0.011
46[a]	O	*n*-propyl	Phenyl	Pyrazolyl	3—F	5.769	5.957	0.188
47	O	*n*-propyl	Phenyl	Pyrazolyl	3—Cl	5.959	5.904	−0.055
48	O	*n*-propyl	Phenyl	Pyrazolyl	3—CF$_3$, 4—Me	5.959	5.957	−0.002
49	O	*n*-propyl	Phenyl	Pyrazolyl	—(CH$_2$)$_4$—	6.161	6.146	−0.015
50[a]	O	*n*-propyl	Phenyl	Pyrazolyl	—CH=CH—CH=CH—	6.538	6.194	−0.344
Class VI: N- and C-linked 5-membered cyclic compounds								
51		(see structure)	Phenyl	Phenyl	4—CF$_3$	6.121	6.104	−0.017
52		Cyclobutyl	3-pyridyl	Pyrazolyl	4—CF$_3$	6.208	6.254	0.046
53		*N*-methyl pyrrolidine	Phenyl	Phenyl	4—CF$_3$	6.258	6.298	0.040
54[a]		Cyclobutyl	2,5-diMePh$_{trans}$	Pyrazolyl	4—CF$_3$	6.496	6.920	0.424
55[a]		Cyclobutyl	Phenyl	Pyrazolyl	4—CF$_3$	6.521	6.834	0.313
56[a]		Cyclobutyl	2,5-diMePh$_{cis}$	Pyrazolyl	4—CF$_3$	7.066	6.920	−0.146
57		Cyclobutyl	Phenyl	Phenyl	4—CF$_3$	7.658	7.678	0.020
58		Cyclobutyl	3-pyrimidine	Phenyl	4—CF$_3$	7.854	7.885	0.031

2-pyridyl; 3-pyridyl; F$_3$C trifluoro propyl; n-propyl; Phenyl; 2,5-dimethyl phenyl; 3-pyrimidine; imidazolyl; pyrazolyl

Note: +/− indicates the enantiomers of the compounds; +R/−S indicates the absolute configuration of chiral centers of the compounds; Residual activity is defined as the difference between predicted activity and experimental activity.
[a]Defines the test set compounds considered for 3D-QSAR analysis.

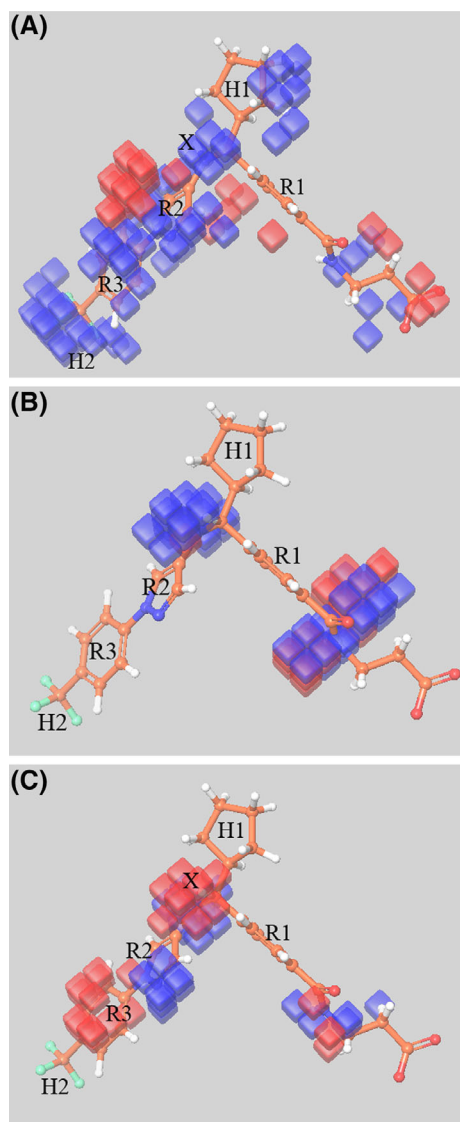### 3.2.3 | Effect of electron withdrawing group

The favorable and unfavorable spatial arrangements of electron withdrawing groups are displayed in Figure 4C. The red cubes over R3 ring indicate the unfavorable positions of electron withdrawing groups in the molecule (Figure 4C). It is known that the presence of electron withdrawing groups alter the electron density over the ring making it more polar, thereby reducing the hydrophobicity of the molecule. The potency of class V compounds is found to be less due to the presence of —CONMe$_2$, —CN, —F and —Cl groups at H2 position of R3 ring. **Compound 41** (pK_i = 5.187), **Compound 44** (pK_i = 5.387), **Compound 46** (pK_i = 5.769), and **Compound 47** (pK_i = 5.959) are found to have electron-withdrawing substituents at H2 position which make them less potent. Slight increment in the potency was observed by the replacement of electron withdrawing groups at H2 by cyclic moiety (**Compound 50**; pK_i = 6.538).

3.3 | Molecular docking

Molecular docking study provides information about the interaction present between the protein and ligand in the protein-ligand complex. The stable binding pose of 58 glucagon receptor antagonists along with their binding energy was predicted using molecular docking simulations. The molecules were divided into six classes (Figure 2) for the better analysis of binding interaction of ligand with the protein. The results were tabulated (Table S2) in the increasing order of pK_i value for each class of ligands. The score obtained from docking results showed a good correlation with the experimental biological activity.

FIGURE 4 3D-QSAR contour plots visualized in the context of favorable and unfavorable positions A, hydrophobic groups, B, hydrogen bond donor groups, and C, electron withdrawing groups [Color figure can be viewed at wileyonlinelibrary.com]

The energy terms contributing to the docking energy of the molecules were given in Table S2. From the study, it was found that ligand interaction takes place in the region of amino acid residues from Arg346 to Lys405, which covers the binding site of the protein. The inhibitor orients in such a way that the polar region is aligned toward TM7 and the hydrophobic part is aligned toward TM5 as mentioned in the literature.[1] All the complexes were found to have the same binding pocket, with different stabilization energy. **Compound 20** showed the highest potency with a least docking score of −8.25 kcal/mol and **Compound 1** showed the least potency with a docking score of −5.33 kcal/mol. Hydrogen bonding and hydrophobic interactions between the ligand and the protein plays a crucial role in the stabilization of protein-ligand complex. The interaction of all the ligands mainly comprises hydrogen-bonding network with residues Arg346, Ser350, Leu399, Asn404, and Lys405.

Among class I ligands, **Compound 12** was found to be the most potent inhibitor with lowest docking energy of −7.79 kcal/mol. The polar region found to stabilize the complex by forming hydrogen bonds between —C═O and the residues Arg346, Ser350, Asn404, and Lys405 (Figure5B [i,ii]).The introduction of electronegative CF_3 group at *para* position of R3 ring resulted in hydrophobic interaction[49] with the amino acid residues which stabilizes the hydrophobic core of ligand. The least potency of **Compound 1** was attributed due to the presence of large number of steric clashes than the attractive nonbonding interactions. The presence of methyl group at *ortho* position does not favor the orientation of hydrophobic core towards TM5 (Figure5A [i,ii]).

Among class II ligands, **Compound 20** is found to be highly potent with lowest docking energy of −8.25 kcal/mol. The introduction of cyclopentyl group as H1 stabilizes the protein-ligand complex (Figure5C [i,ii]). Varying the hydrophobic core of class III inhibitors changes the potency and stability of protein-ligand complexes. **Compound 32** with a pK_i value of 7.854 showed an additional hydrogen bonding interaction between —NH and the residues Ser350 and Leu399 (Figure5D [i,ii]).The introduction of different alkyl side chains to class IV ether series showed less difference in its binding energy. The docking studies showed that the side chains favors the orientation of hydrophobic and polar cleft of inhibitors toward the TM5 and TM7 membranes, respectively. The alkyl chain at H1 position of **Compound 40** favors the formation of a stable protein-ligand complex than their cyclic analogs (Figure5E [i,ii]). The complexity of the substituents on the heterocyclic ring of class V inhibitors has an impact on the orientation and binding energy of the complexes. The **Compound 49** and **Compound 50** contains cyclic substituents which support the orientation of hydrophobic cleft toward TM5 and polar cleft toward TM7 favoring hydrogen bond formation (Figure5F [i,ii]). The heteroatoms at cyclic linkage (H1 region) of class VI inhibitors do not favor the formation of stable protein-ligand complex. The replacement of heteroatoms with —CH_2 and —CH groups in H1 region favors the hydrophobicity of the molecules and forms a protein-ligand complex with least binding energy (**Compound 56**, **Compound 57**, **Compound 58**). Additionally, **Compound 58** showed hydrogen bonding between —NH and Leu399 and π-cation interaction with Lys349 (Figure5G [i,ii]). Docking studies showed that the proper orientation of hydrophobic region of the molecule toward TM5 is essential for the formation of hydrogen bonds with Arg346, Asn404, and Lys405 amino acids to stabilize protein-ligand complex.

The docking results were validated by redocking the cocrystallized ligand NNC-0640 to the binding site of glucagon class B G-protein-coupled receptor (PDB ID: 5XEZ). The root mean square deviation (RMSD) value of redocked and X-ray crystal structure of ligand was calculated. The docking pose and the interactions obtained after redocking showed good agreement with the literature[3] with an RMSD value of 0.82 Å[29] (Figure S3).

3.4 | Frontier molecular orbital analysis

DFT calculations were carried out for selected ligands to find out the electronic properties of the molecules such as HOMO-LUMO
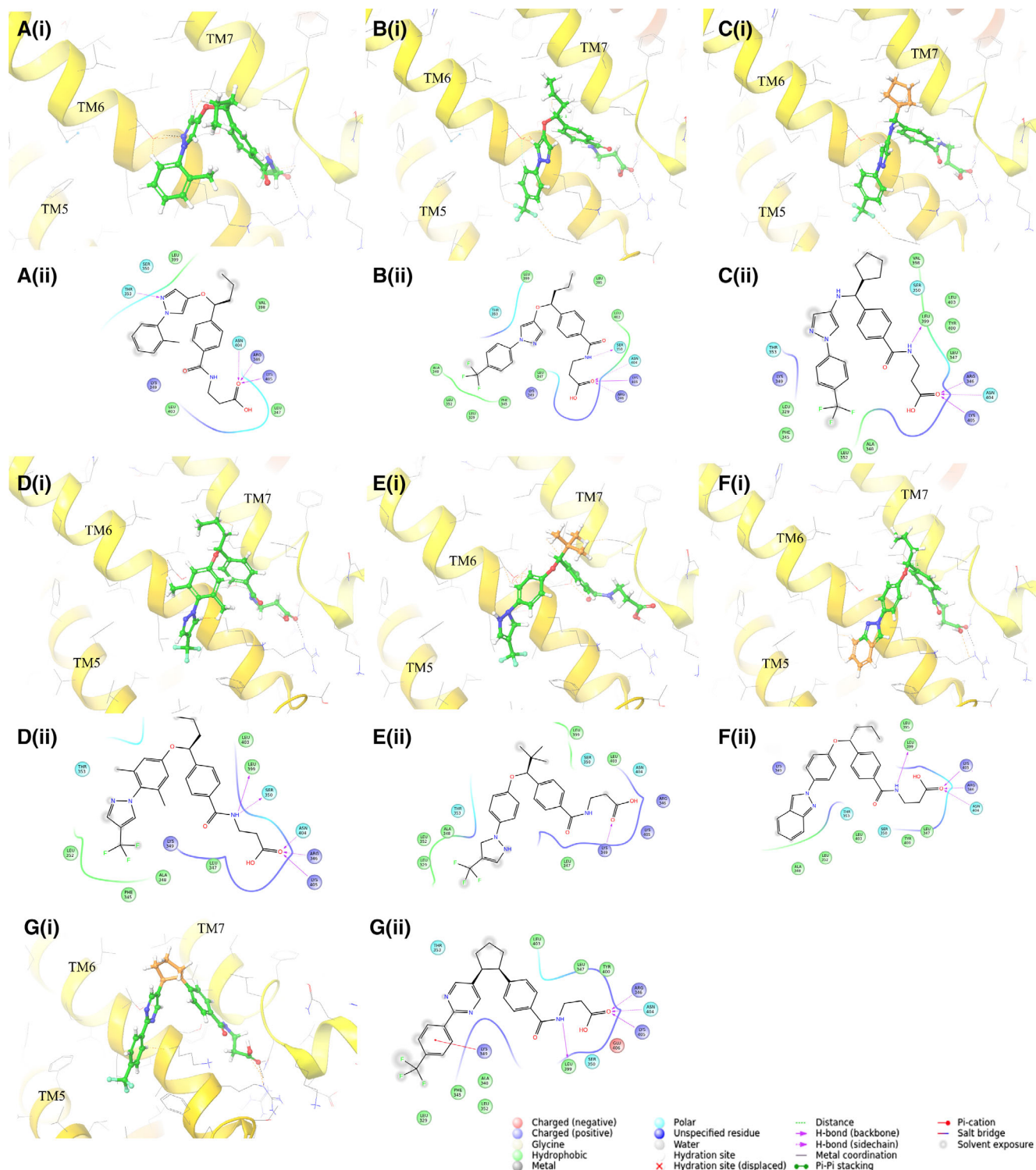
FIGURE 5 Binding pose of lowest-energy conformation of inhibitors bound to glucagon receptor and its 2-D ligand interaction diagrams are shown. A (i,ii), B (i,ii), C (i,ii), D (i,ii), E (i,ii), F (i,ii), and G (i,ii) corresponds to Compound 1, Compound 12, Compound 20, Compound 32, Compound 40, Compound 50, and Compound 58, respectively

energies, energy gap and dipole moment. Three ligands (active, moderately active, and inactive) from each class were chosen for energy calculations. The electronic interactions of the molecules play an important role in its pharmacological effects. The position of HOMO-LUMO orbital is responsible for the electron transfer in a chemical reaction and the energy gap value represents its chemical reactivity. The electronic properties of active, moderately active and inactive compounds from each class were shown in Table 3. The value of HOMO ranges from −5.495 to −6.620 eV, LUMO ranges from −1.045 to −1.720 eV and the energy gap ranges from 4.076 to

5.368 eV. The dipole moment of the ligands ranges from 1.7147 to 9.5281 Debye. The low HOMO-LUMO energy gap of 4.076 eV for the potent ligand (**Compound 20**) among all the classes indicates the high chemical reactivity and low kinetic stability. The lowest potent ligand (**Compound 1**) among all the classes has a higher energy gap of 5.368 eV, indicates low reactivity and comparatively high stability among inhibitors.

Figure 6 shows the position of HOMO-LUMO orbitals of selected ligands in each class based on its potency. It was observed that the substitution of various groups at the hydrophobic part of the inhibitors alters the topology of HOMO-LUMO orbitals in molecules. Majority of the ligand showed a well-defined separation in the position of HOMO-LUMO orbitals indicating that the energies are localized on different parts of the molecule. Among class I inhibitors, HOMO covers mainly the polar part while LUMO covers R2, R3 ring of **Compound 12**. We found HOMO electron density on the H1, R2, and R3 region and LUMO density over R1 and polar end for inactive **Compound 1**. This can be attributed due to the presence of electronegative p-CF$_3$ group at R3 ring of **Compound 12**, which shifts the electron density toward the polar region of molecule. The presence of o-CH$_3$ group at the R3 ring of **Compound 1** increases the electron density over hydrophobic region. A low energy gap of 5.097 eV shows that **Compound 12** is chemically reactive among class I compound. The HOMO orbitals were located mainly on the polar R1 region and cyclobutyl ring (H1) of **Compound 20** while LUMO orbitals were located on R2, R3 ring of hydrophobic region. The compound has a low energy gap value of 4.076 eV. Inactive **Compound 13** found to have HOMO over hydrophobic (R2 and R3) region and LUMO over polar R1 region. The active inhibitors among class III (**Compound 33**)

and class IV (**Compound 40**) compounds found to have LUMO orbital over polar R1 region and HOMO orbital mainly over hydrophobic R2 and R3 region of the molecule. The HOMO-LUMO energy gap of highly active **Compound 33** and **Compound 40** was found to be 4.517 and 4.701 eV, respectively. The active inhibitor among class V compounds (**Compound 50**) showed HOMO electron density near hydrophobic R3 and H2 region and LUMO electron density over R2, R3, and H2 region. The active inhibitor of class VI compounds (**Compound 58**) showed HOMO in the region of R1, R2, R3, and H1; while LUMO electron densities at the R2 and R3 region only. The energy gap value for **Compound 50** in class V and **Compound 58** in class VI were found to be 4.507 and 4.900 eV, respectively indicating high chemical reactivity among their respective classes. Therefore, from the above discussion it is clear that presence of HOMO electron density near the polar part, mainly over the R1 region and the hydrophobic H1 region contributes to higher potency of the inhibitor (**Compound 12**, **Compound 20**, and **Compound 58**). The LUMO electron density was found near R2 and R3 region for such cases. This can be further clarified from the HOMO-LUMO analysis of the interacting amino acid of GCGR allosteric site.

Finally, the HOMO-LUMO electron density over the interacting amino acid residues at the allosteric pocket was determined to predict the mechanism of ligand binding (Figure 7). From the literature, it is known that, HOMO orbitals of ligand interacts with the LUMO orbitals of amino acid residues at the binding site.[50] Similarly, HOMO orbitals of binding site residues interact with the LUMO orbitals of ligand during complex formation.[51] The LUMO density over hydrophobic part of ligands are responsible for its interactions with the amino acid residues Leu329, Phe345, and Ser350 having HOMO

TABLE 3 Summary of electronic properties of selected ligands

Compound		E_{HOMO} (eV)	E_{LUMO} (eV)	E_{Gap} (eV)	Dipole moment (Debye)
Class I	1	−6.428	−1.060	5.368	3.3852
	8	−6.074	−1.045	5.029	3.4788
	12	−6.478	−1.381	5.097	3.3376
Class II	13	−6.287	−1.332	4.955	4.9411
	17	−5.495	−1.187	4.308	4.6355
	20	−5.684	−1.608	4.076	7.1049
Class III	21	−6.592	−1.687	4.905	3.5415
	27	−6.215	−1.395	4.820	7.7252
	33	−5.776	−1.259	4.517	5.6430
Class IV	34	−6.092	−1.263	4.829	1.7147
	36	−6.037	−1.280	4.757	2.8992
	40	−6.158	−1.457	4.701	9.5281
Class V	41	−6.423	−1.130	5.293	3.9758
	49	−6.594	−1.237	5.357	5.9272
	50	−6.166	−1.659	4.507	5.3732
Class VI	54	−6.309	−1.068	5.241	6.3752
	55	−6.525	−1.378	5.147	7.5089
	58	−6.620	−1.720	4.900	8.6880

density. The ligands having HOMO density over the polar end of the molecules are highly potent due to its ability to form stable hydrogen bonding interactions with Lys405, Leu399, Ser350, and Arg346 having LUMO density. A well-separated HOMO and LUMO electron density over the hydrophobic part and polar end, respectively of the ligands help in stabilizing the formed complex. The HOMO over the cyclopentyl side chain and the polar part of most potent **Compound 20** interacts well with LUMO of amino acid residues Leu399 and Ser350. This interaction helps to orient the ligand to form a stable hydrogen bonding with the carbonyl oxygen of Leu399 (C=O---H—N). The amino acid residues Arg346, Asn404, Lys405 form hydrogen bond between the carbonyl oxygen of ligand (N—H---C=O) and Leu399 forms hydrogen bond with the amino group of ligand (C=O---H—N). This study confirms the importance of presence of HOMO orbital near the polar part and LUMO near the hydrophobic region of the inhibitor to form stable protein-ligand complex.

3.5 | MM/GBSA binding free energy

The quantitative measure of the binding strength of docked inhibitors to the allosteric pocket of GCGR was carried out by using MM/GBSA method. The average free energy of binding was calculated using following equation,

$$\Delta G_{bind} = \Delta G_{Complex} - \left(\Delta G_{Protein} + \Delta G_{Ligand}\right) \tag{1}$$

$\Delta G_{Complex}$ is the Gibbs free energy of bound protein, whereas $\Delta G_{Protein}$ and ΔG_{Ligand} are individual Gibbs free energy of the unbound protein and ligand molecule. In general, ΔG_{bind} to form a protein-ligand complex can be written as[52]

$$\Delta G_{bind} = \langle \Delta G_{bind-MM} \rangle + \langle \Delta G_{bind-sol} \rangle - T\langle \Delta S \rangle \tag{2}$$

$$\Delta G_{bind-MM} = \Delta G_{bind-ele} + \Delta G_{bind-vdw} + \Delta G_{bind-cov} \tag{3}$$
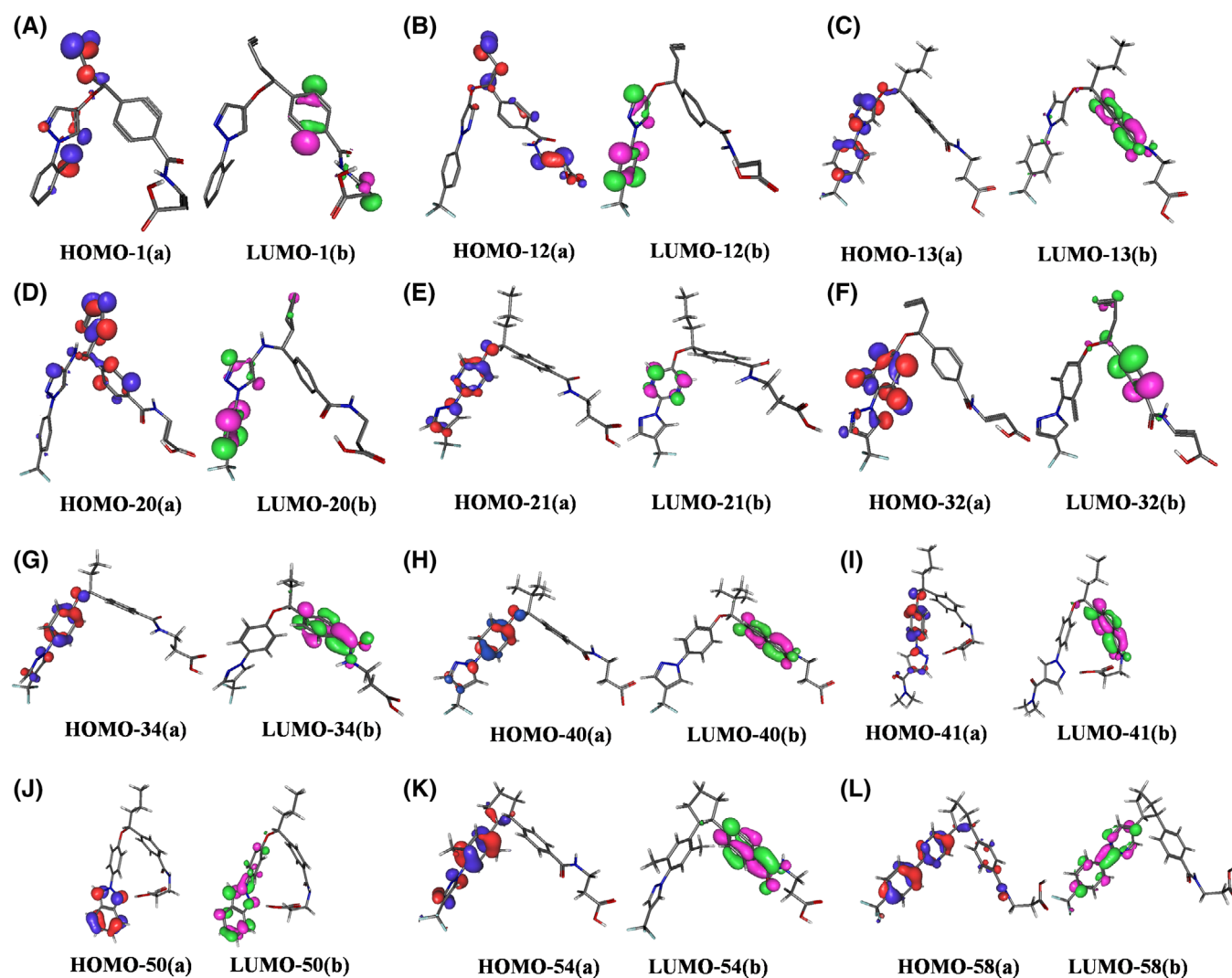
FIGURE 6 The highest occupied molecular orbital (HOMO)-lowest unoccupied molecular orbital (LUMO) orbital positions mapped on selected ligands at B3LYP/6-31G(d,p) level of calculation [Color figure can be viewed at wileyonlinelibrary.com]

$$\Delta G_{bind-sol} = \Delta G_{bind-GB} + \Delta G_{bind-SA} \qquad (4)$$

where, ΔG_{MM} is the total molecular mechanics energy of the interactions between protein and ligand. ΔG_{MM} is obtained from internal energy (bond, angle, dihedral energies), electrostatic ($\Delta G_{bind-ele}$), and van der Waals energy ($\Delta G_{bind-vdw}$) terms. The solvation free energy ($\Delta G_{bind-sol}$) mainly comprises of electrostatic solvation energy and non-polar solvation energy terms. The electrostatic solvation free energy was calculated using generalized Born (GB) method[53] and the non-polar contribution to solvation free energy ($\Delta G_{bind-SA}$) was computed using solvent accessible surface area (SASA).[54] It is believed that implicit solvent models implicitly include entropy associated with the solvent during the calculation of solvation free energies. However, entropic contribution can also be calculated using normal mode analysis.[55] The average free energy of binding and the corresponding energy components of the bound antagonists toward the trans-membrane allosteric pocket of GCGR were calculated from one of the most stable simulated trajectories performed in implicit solvent. It is found that the binding free energy of all the antagonists considered in present study ranges from −34.675 kcal/mol to −64.18 kcal/mol (Table S3). **Compound 20** (pK_i = 8.046) displayed higher binding energy of −63.475 kcal/mol toward GCGR whereas **Compound 1** (pK_i = 4.833) found to have binding energy of −41.097 kcal/mol. The binding free energy decomposition of individual inhibitors was carried out according to Equations (2), (3), and (4) and illustrated in Table S3. The van der Waals energy terms ($\Delta G_{bind-vdw}$) and non-polar solvation energy ($\Delta G_{bind-SA}$) term for highest potent **Compound 20** is found to be −50.516 and −16.853 kcal/mol, respectively. Similarly, the $\Delta G_{bind-vdw}$ and $\Delta G_{bind-SA}$ energy terms found to favor the strong binding of all active compounds (Table S3). The major contribution of hydrophobic stabilization energy indicates the importance of benzene rings and hydrophobic residues located at region C of the inhibitors. The least potent **Compound 1** displayed remarkable decrease in van der Waals energy component ($\Delta G_{bind-vdw}$ = −36.623 kcal/mol) in comparison to other highly active inhibitors. Since the antagonists are buried inside the membrane bilayer the van der Waals and hydrophobic solvation energy terms are found to be dominating rather than electrostatic solvation energy terms. Similarly, the covalent energy terms are also disfavoring the binding of inhibitors toward GCGR (Table S3). The energy terms due to H-bond formation of all the inhibitors ranges from −0.002 kcal/mol to −1.979 kcal/mol which indicates the small contribution of electrostatic interactions for stable inhibitor binding at 5XEZ allosteric pocket. Therefore, the van der Waals ($\Delta G_{bind-vdw}$) and non-polar solvation energy ($\Delta G_{bind-SA}$) terms seems to be key contributing factor for thermodynamically stable binding of active inhibitors at the 5XEZ allosteric pocket.

3.6 | Molecular dynamics simulation

Five independent 100 ns atomistic Molecular dynamics simulation was performed in order to obtain insights into the dynamical behavior of highest potent **Compound 20** at the trans-membrane allosteric pocket of 5XEZ. The root-mean-square-deviation (RMSD) profiles of the Cα, backbone, side chain, and heavy atoms of one of the simulated trajectories are shown in Figure 8A. The RMSD value of the protein Cα was found to increase up to a value of 6.2 Å with respect to its starting coordinate (t = 0) for first 10 ns and stabilize around an average value of 5.748 Å for rest of the MD trajectories. The average
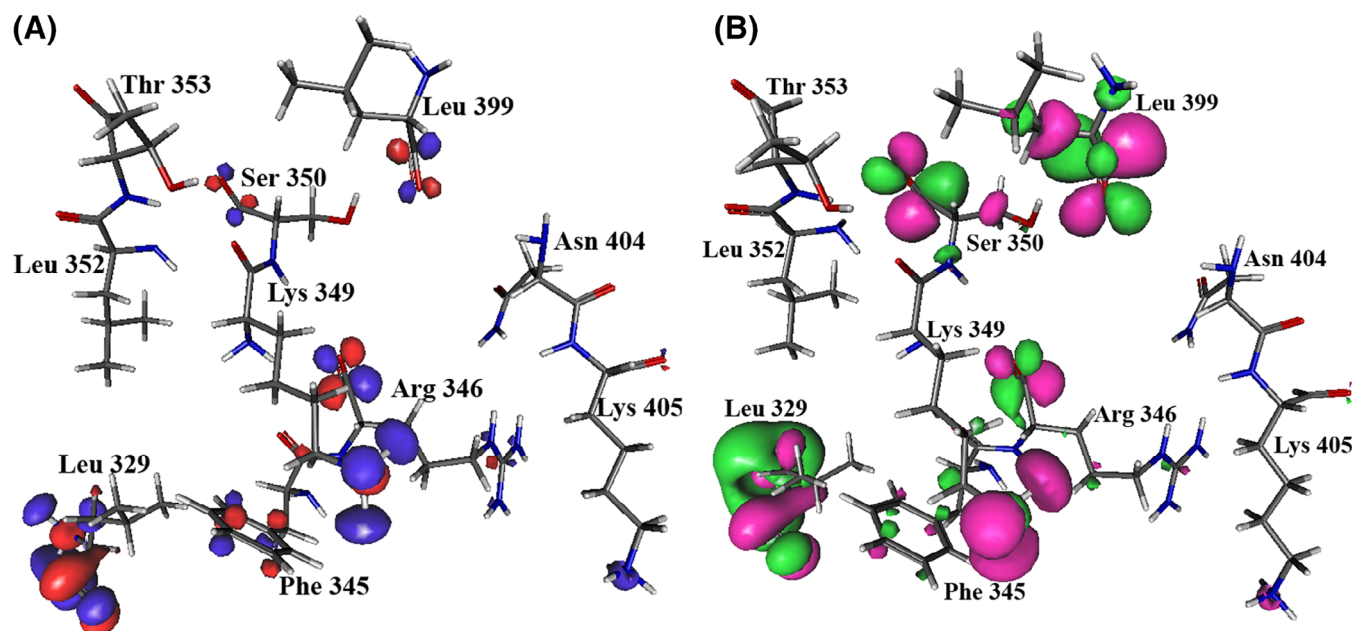

FIGURE 7 Position of (A) HOMO (B) LUMO regions of interacting amino acid residues at the allosteric pocket of GCGR [Color figure can be viewed at wileyonlinelibrary.com]

RMSD of backbone, side chains and heavy atoms are found to be 5.745, 6.657, and 6.179 Å, respectively, which indicate significant change in protein backbone compared to its crystal structure. It is evident from the RMSD of **Compound 20** that the movement of ligand copes well with the movement of amino acid residues at protein allosteric site (Figure 8A). The average RMSD values of Cα, backbone, side chains, and heavy atoms for five independent MD simulations are found to be 6.403 Å ± 0.65, 6.393 Å ± 0.64, 7.211 Å ± 0.58, and 6.785 Å ± 0.62, respectively (Figure S4 and Table S4). Further, the root mean square fluctuation (RMSF) of the backbone at the allosteric site of 5XEZ is found to be in the range of 2.548 to 4.561 Å (Figure S5) which indicates lower degree of flexibility in that region. It is clear from the above discussion that ligand movement was stable during the simulation. It is evident from Figure S5 that residues stretches including His44-Arg60, Asp70-Ala77, Lys98-Gly112, Trp415-Arg419, Arg444-Pro454 have high fluctuations and reside away from the trans-membrane allosteric site.

The key nonbonded interactions between **Compound 20** and 5XEZ during 100 ns MD simulation are illustrated in Figure S6. It is clear from Figure S6-A,B that nonbonded interactions are mainly present in the region of Arg346-Leu352 and Leu399-Glu406, whereas the region Val363-Lys381 found to be more fluctuating (Figure S5). Hence, no interactions are found between these regions. It is evident from Figure S6-A,B that Lys349 and Leu399 have major contact with the ligand throughout the simulation and are probably responsible for stabilization of **Compound 20** at 5XEZ catalytic pocket. However, Arg346, Leu352, Leu395, and Leu403 residues were found to have less interaction with the ligand throughout the MD trajectory.

It is found that hydrophobic and hydrogen bonding interactions are major contributing factor for stabilizing **Compound 20** at the trans-membrane allosteric pocket of 5XEZ which is in accordance with our MM/GBSA result. Lys 349 found to exhibit π-cation interaction with the ligand benzene ring and pyrazole ring for 97% of the MD trajectory (Figure S6-C). Among the four-hydrogen bond predicted by Autodock, only one is found to be preserved during MD simulation. The carbonyl oxygen of Leu399 accepts a hydrogen bond with hydroxyl hydrogen (region A) of the ligand for 18% of the MD trajectory. The number of hydrogen bonds between **Compound 20** and 5XEZ throughout the trajectory is found to be 1 (Figure 8B). The average number of water mediated hydrogen bond with **Compound 20** is found to be 2 (Figure 8C). This further adds to the stability of the **Compound 20**. The snapshots of the simulation at each 10 ns interval are illustrated in Figure 9 to further confirm ligand stability. Region C of **Compound 20** found to move away from its initial position around 20 ns of the simulation (Figure 9C) and further stabilized by π-cation interaction formed by Lys349. The π-cation interaction between **Compound 20** and Lys349 found to be present in each snapshot displayed in Figure 9 which confirms its key role in anchoring the inhibitor at 5XEZ catalytic pocket. A well-defined, water mediated hydrogen bond network is observed between the hydrophilic part of **Compound 20** and TM7 amino acid residues from 10 ns of the simulation. Those water molecules are probably responsible for stabilizing the hydrophilic part of **Compound 20**. A low RMSD value (1.571 Å
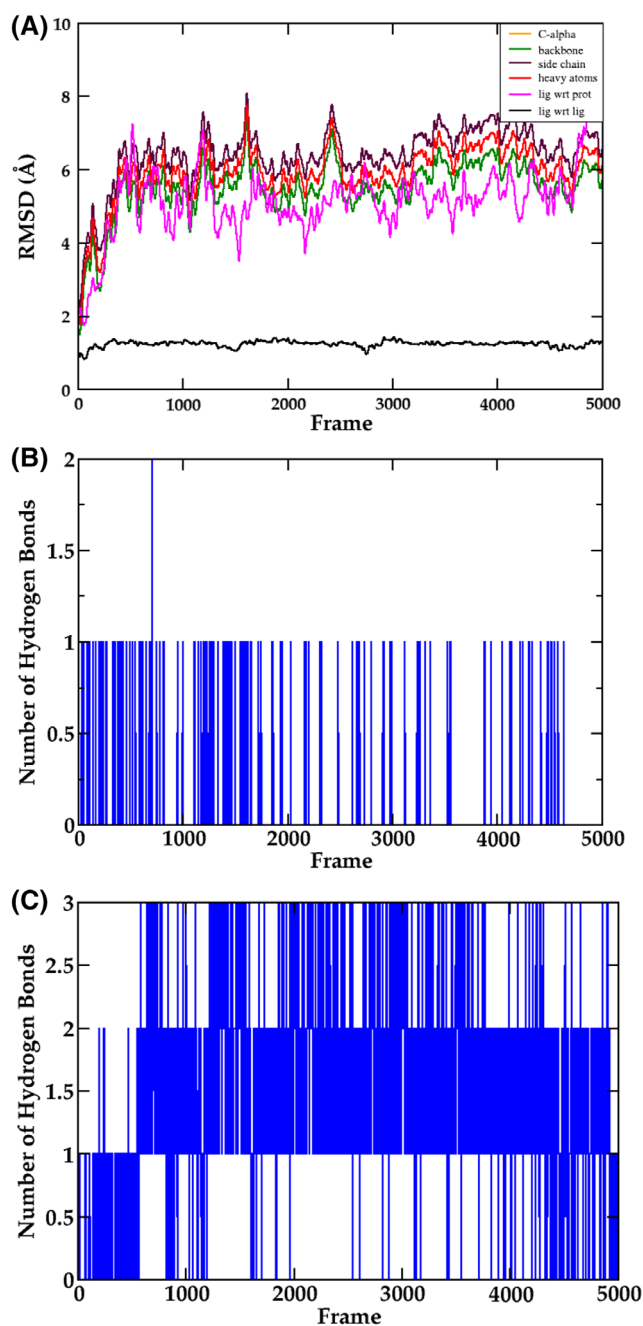
FIGURE 8 (A) Time-line representation of RMSD profile of Cα, backbone, and heavy atoms of 5XEZ with respect to its initial coordinate. The RMSD of **compound 20** with respect to protein backbone and its own starting structure was illustrated in pink and black color, respectively. Number of hydrogen bonds formed between (B) protein and ligand (C) ligand and water throughout 100 ns [Color figure can be viewed at wileyonlinelibrary.com]

± 0.57) of ligand indicates a less conformational change with respect to the initial conformation. The gyration radius (rGyr, measures the extendedness of a ligand) found to stabilize after 5 ns of the simulation with an average value of 5.612 Å ± 0.44. The SASA, polar surface area (PSA), and molecular surface area (MolSA) of ligand for most stable ligand binding are found to be in the range of 277.069 to 449.461Å², 140.24 to 180.367Å², and 439.852 to 462.211Å²,
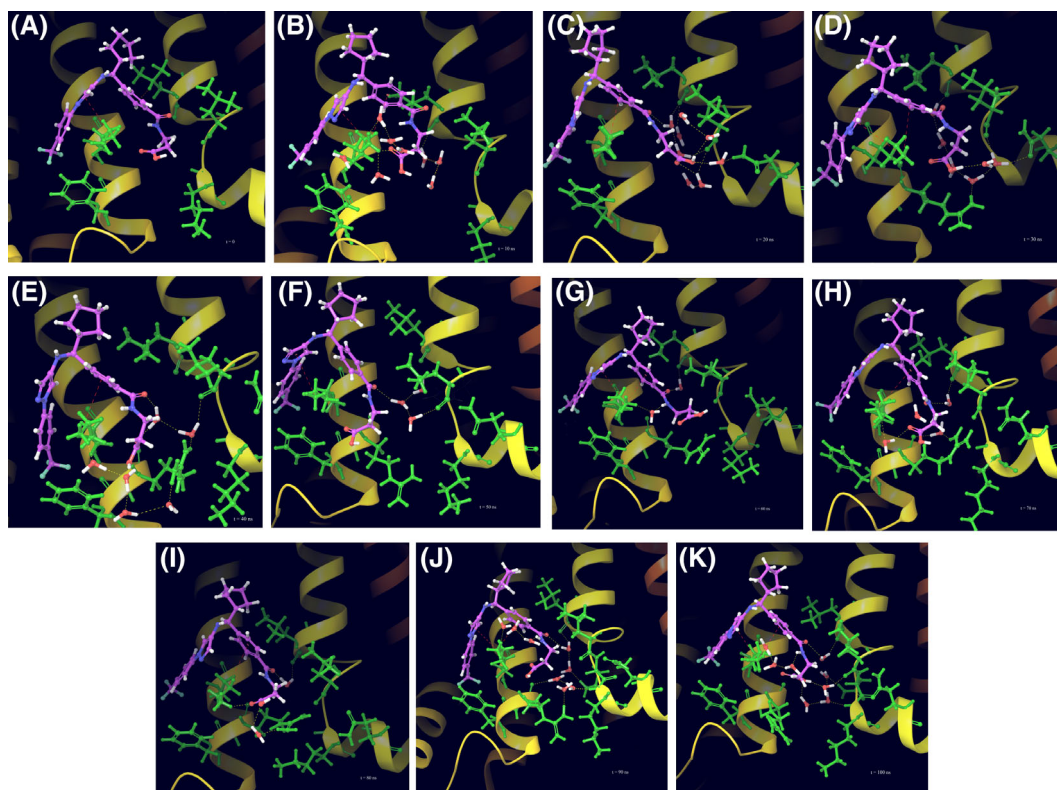
FIGURE 9 Snapshots at (A) 0 ns, (B) 10 ns, (C) 20 ns, (D)30 ns, (E) 40 ns, (F) 50 ns, (G) 60 ns, (H) 70 ns, (I) 80 ns, (J) 90 ns, and (K) 100 ns of MD trajectory are illustrated. The interacting amino-acid residues are colored in green, oxygen of water molecules are represented in red. The π-cation interaction and hydrogen bonding interactions are indicated by red and yellow, respectively [Color figure can be viewed at wileyonlinelibrary.com]

respectively, which supports the stabilization of ligand in binding site of protein during simulation (Figure S7). The details for all five simulations are given in Table S4.

3.7 | ADMET properties of active compounds

In the drug development process, the drug candidate under consideration need to possess high potency as well as good pharmacokinetic (PK) profiles (drug-like properties) in order to confirm their effectiveness and bioavailability. The permissible ranges of crucial pharmacokinetic properties and the predicted ADME/tox properties of 14 top-scored ligands are documented in Table S5. All the calculated properties were found to be in their permissible range and hence confirming their drug-like properties.

The bioavailability of top scored molecules was determined through polar surface area (PSA) analysis which measures the cell wall permeability or transport through membranes. It is believed that drug candidates with PSA less than 100 Å^2 have good absorption properties while PSA value above 140 Å^2 reported to be responsible for poor oral availability. It is clear from the result that the glucagon receptor antagonists possess PSA value in the range of 81.913 to 124.622 Å^2, which shows good oral availability. The number of non-hindered rotatable bonds was found to vary from 4 to 9. The calculated molecular weights of top-scored hits were found to be in the range of 474.482 to 528.453, which is acceptable for orally

consumable drugs. The lipophilicity of the selected compounds was addressed by QPlogP$_{o/w}$ (partition coefficient between octanol and water) which estimates the hydrophilic and hydrophobic character of any molecule. The value of QPlogP$_{o/w}$ was found to be in the range of 3.387 to 7.403 and most of the top-scored compounds are in the permissible range of −2.0 to 6.5. **Compound 32** and **Compound 57** exhibited QPlogP$_{o/w}$ value of 7.073 and 7.403, respectively. The high value of logP$_{o/w}$ indicates higher lipophilic character of those molecules. The aqueous solubility (QPlogS) was found to be in range of −9.2 to −5.5 for all the top-scored hits. The highest active **Compound 20** was found to exhibit the QPlogS value of −6.130, which is under the permissible zone. The number of hydrogen bond acceptors (4.5-7.5) and hydrogen bond donors (1, 2) were also found to vary in their permissible value. The ability of the molecule to pass through the blood/ brain barrier was defined by QPlogBB and their values found to be in acceptable range of −1.344 to −0.3. The binding of molecule to human serum albumin (QPlogK$_{hsa}$), and the present of human oral absorption were found to be in threshold limit. It is believed that orally active compound should not have more than two violations of Lipinski's rule, which is in accordance with our result (Table S5). Some of the molecules showed deviations of two parameters of Lipinski's rule of 5 due to its higher hydrophobic character. From the above discussion, it is evident that top scored compounds obtained from the present study have good oral bioavailability.

4 | CONCLUSION

In the present study, 3D-QSAR, molecular docking, DFT calculation, MM/GBSA, molecular dynamics simulation, and ADME/Toxicity studies of β-alanine analogs were performed to investigate the effect of structural determinants responsible for GCGR antagonism. The developed 3D-QSAR model gave R^2 value of 0.9981 and Q^2 value of 0.8253, indicating excellent consistency and internal predictability of the model. Contour plots obtained from 3D-QSAR model revealed the position of hydrophobic/non-polar substituents contributing to increase the inhibitory activity. Electron withdrawing groups present at R2 and R3 rings are found to have unfavorable contribution to the potency of the inhibitor. Further, molecular docking study predicted the binding pose of antagonist at the binding site of GCGR. The reliability of the docking study was confirmed by low RMSD value of 0.82 Å between the co-crystal and docked ligand. Further, the docking study suggested that the polar region of the ligand forms hydrogen-bonding network with Arg346, Ser350, Leu399, Asn404, and Lys405 amino acid residues. The presence of —CF_3 group at R3 ring and cyclopentyl group at H2 position stabilize **Compound 20** at hydrophobic region of TM5, which helps to make hydrogen bonds between the polar part of the ligand and TM7 region. The presence of o-methyl group at R3 ring of **Compound 1** found to destabilize the ligand towards TM5. It is evident from the study that the alignment of hydrophobic region towards TM5 facilitates the proper orientation required for GCGR allosteric inhibition. HOMO-LUMO orbital analysis described the interaction mechanism of ligand with the protein at quantum level. The presence of HOMO near hydrophobic H1 region and polar R1 region gives favorable interactions with amino acids Lys405, Leu399, Ser350, Arg346 having LUMO density. Similarly, the presence of LUMO near hydrophobic R2 and R3 region of the ligand gives favorable interactions with amino acids Leu329, Phe345, Ser350 having HOMO density. MM/GBSA calculation displayed that van der Waals and non-polar solvation energy terms contribute mostly for stabilizing the antagonist binding to GCGR. The binding energy of highly active **Compound 20** was found to be −63.475 kcal/mol. Further, stability of **Compound 20** at 5XEZ allosteric pocket was confirmed by 100 ns atomistic molecular dynamics simulation. MD simulation revealed that π-cation interaction of Lys349 and hydrogen bonding of Leu399 have crucial role in stabilizing **Compound 20** under motion. Water molecules near the hydrophilic part of the ligand found to have hydrogen bonding with the ligand, thereby stabilizing the protein-ligand complex effect. Lastly, ADME/tox calculation of top-scored compounds obtained from present study assured their safe administration in human body. The outcomes of the present study provide insightful information regarding the design of novel glucagon receptor antagonists to treat T2DM.

ORCID

Debashree Chakraborty https://orcid.org/0000-0002-0142-7941

REFERENCES

1. Jazayeri A, Doré AS, Lamb D, et al. Extra-helical binding site of a glucagon receptor antagonist. *Nature*. 2016;533(7602):274-277. https://doi.org/10.1038/nature17414.
2. Jiang G, Zhang BB. Glucagon and regulation of glucose metabolism. *Am J Physiol-Endocrinology and Metabol*. 2003;284(4):E671-E678. https://doi.org/10.1152/ajpendo.00492.2002.
3. Zhang H, Qiao A, Yang D, et al. Structure of the full-length glucagon class B G-protein-coupled receptor. *Nature*. 2017;546(7657):259-264. https://doi.org/10.1038/nature22363.
4. Mukund S, Shang Y, Clarke HJ, et al. Inhibitory mechanism of an allosteric antibody targeting the glucagon receptor. *J Biol Chem*. 2013;288(50):36168-36178. https://doi.org/10.1074/jbc.M113.496984.
5. Duffy JL, Kirk BA, Konteatis Z, et al. Discovery and investigation of a novel class of thiophene-derived antagonists of the human glucagon receptor. *Bioorg Med Chem Lett*. 2005;15(5):1401-1405. https://doi.org/10.1016/j.bmcl.2005.01.003.
6. Kurukulasuriya R, Sorensen BK, Link JT, et al. Biaryl amide glucagon receptor antagonists. *Bioorg Med Chem Lett*. 2004;14(9):2047-2050. https://doi.org/10.1016/j.bmcl.2004.02.056.
7. Shen D-M, Lin S, Parmee ER. A survey of small molecule glucagon receptor antagonists from recent patents (2006 – 2010). *Expert Opin Ther Pat*. 2011;21(8):1211-1240. https://doi.org/10.1517/13543776.2011.587001.
8. Cherkasov A, Muratov EN, Fourches D, et al. QSAR Modeling: where have you been? Where are you going to? *J Med Chem*. 2014;57(12):4977-5010. https://doi.org/10.1021/jm4004285.
9. Kang JS, Zhang AL, Faheem M, et al. Virtual screening and experimental testing of B1 Metallo-β-lactamase inhibitors. *J Chem Inf Model*. 2018;58(9):1902-1914. https://doi.org/10.1021/acs.jcim.8b00133.
10. Das BK, Pv P, Chakraborty D. Computational insights into factor affecting the potency of diaryl sulfone analogs as *Escherichia coli* dihydropteroate synthase inhibitors. *Comput Biol Chem*. 2019;78:37-52. https://doi.org/10.1016/j.compbiolchem.2018.11.005.
11. Malkhasian AYS, Howlin BJ. Docking and DFT studies on ligand binding to Quercetin 2,3-dioxygenase. *J Biomol Struct Dyn*. 2016;34(11):2453-2461. https://doi.org/10.1080/07391102.2015.1123190.
12. Rana MK, Sinha M, Panda S. Gas sensing behavior of metal-phthalocyanines: effects of electronic structure on sensitivity. *Chem Phys*. 2018;513:23-34. https://doi.org/10.1016/j.chemphys.2018.06.021.
13. Sarmah A, Roy RK. Understanding the preferential binding interaction of aqua-cisplatins with nucleobase guanine over adenine: a density functional reactivity theory based approach. *RSC Adv*. 2013;3(8):2822. https://doi.org/10.1039/c2ra23223e.
14. Filipski KJ, Bian J, Ebner DC, et al. A novel series of glucagon receptor antagonists with reduced molecular weight and lipophilicity. *Bioorg Med Chem Lett*. 2012;22(1):415-420. https://doi.org/10.1016/j.bmcl.2011.10.113.
15. Guzman-Perez A, Pfefferkorn JA, Lee ECY, et al. The design and synthesis of a potent glucagon receptor antagonist with favorable

physicochemical and pharmacokinetic properties as a candidate for the treatment of type 2 diabetes mellitus. *Bioorg Med Chem Lett*. 2013;23(10):3051-3058. https://doi.org/10.1016/j.bmcl.2013.03.014.

16. Lee ECY, Tu M, Stevens BD, et al. Identification of a novel conformationally constrained glucagon receptor antagonist. *Bioorg Med Chem Lett*. 2014;24(3):839-844. https://doi.org/10.1016/j.bmcl.2013.12.090.

17. Stewart JJP. Optimization of parameters for semiempirical methods I. method. *J Comp Chem*. 1989;10(2):209-220. https://doi.org/10.1002/jcc.540100208.

18. Ditchfield R, Hehre WJ, Pople JA. Self-consistent molecular-orbital methods. IX an extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J Chem Phys*. 1971;54(2):724-728. https://doi.org/10.1063/1.1674902.

19. Lee C, Yang W, Parr RG. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys Rev B*. 1988;37(2):785-789. https://doi.org/10.1103/PhysRevB.37.785.

20. Frisch, M., Trucks, G., Schlegel, H., Scuseria, G., Robb, M., Cheeseman, J., ... Fox, D.. *Gaussian 09, Revision B.01*. Wallingford, CT: Gaussian, Inc; 2009.

21. Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J Comp Aided Mol Des*. 2006;20(10–11):647-671. https://doi.org/10.1007/s10822-006-9087-6.

22. Rácz A, Bajusz D, Héberger K. Consistency of QSAR models: correct split of training and test sets, ranking of models and performance parameters. *SAR QSAR Environ Res*. 2015;26(7–9):683-700. https://doi.org/10.1080/1062936X.2015.1084647.

23. Golbraikh A, Shen M, Xiao Z, Xiao Y-D, Lee K-H, Tropsha A. Rational selection of training and test sets for the development of validated QSAR models. *J Comput Aided Mol Des*. 2003;17(2–4):241-253.

24. Polański J, Gieleciak R, Bąk A. The comparative molecular surface analysis (COMSA) – A nongrid 3D QSAR method by a coupled neural network and PLS system: predicting pK_a values of benzoic and Alkanoic acids. *J Chem Inf Comput Sci*. 2002;42(2):184-191. https://doi.org/10.1021/ci010031t.

25. Jacobson MP, Pincus DL, Rapp CS, et al. A hierarchical approach to all-atom protein loop prediction. *Proteins*. 2004;55(2):351-367. https://doi.org/10.1002/prot.10613.

26. Morris GM, Huey R, Lindstrom W, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem*. 2009;30(16):2785-2791. https://doi.org/10.1002/jcc.21256.

27. Gasteiger J, Marsili M. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron*. 1980;36(22):3219-3228. https://doi.org/10.1016/0040-4020(80)80168-2.

28. Morris GM, Goodsell DS, Halliday RS, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem*. 1998;19(14):1639-1662.

29. Kramer B, Rarey M, Lengauer T. Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. *Proteins: Structure, Function, and Genetics*. 1999;37(2):228-241. https://doi.org/10.1002/(SICI)1097-0134(19991101)37:2<228::AID-PROT8>3.0.CO;2-8.

30. Blinder SM. Basic concepts of self-consistent-field theory. *Am J Phys*. 1965;33(6):431-443. https://doi.org/10.1119/1.1971665.

31. Becke AD. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A*. 1988;38(6):3098-3100. https://doi.org/10.1103/PhysRevA.38.3098.

32. Becke AD. Density-functional thermochemistry. III the role of exact exchange. *J Chem Phys*. 1993;98(7):5648-5652. https://doi.org/10.1063/1.464913.

33. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol des*. 2013;27(3):221-234. https://doi.org/10.1007/s10822-013-9644-8.

34. Hou T, Wang J, Li Y, Wang W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model*. 2011;51(1):69-82. https://doi.org/10.1021/ci100275a.

35. Sirin S, Kumar R, Martinez C, et al. A computational approach to enzyme design: predicting ω-aminotransferase catalytic activity using docking and MM-GBSA scoring. *J Chem Inf Model*. 2014;54(8):2334-2346. https://doi.org/10.1021/ci5002185.

36. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc*. 1996;118(45):11225-11236. https://doi.org/10.1021/ja9621760.

37. Li J, Abel R, Zhu K, Cao Y, Zhao S, Friesner RA. The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins*. 2011;79(10):2794-2812. https://doi.org/10.1002/prot.23106.

38. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983;79(2):926-935. https://doi.org/10.1063/1.445869.

39. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *J Chem Phys*. 1995;103(19):8577-8593. https://doi.org/10.1063/1.470117.

40. Martyna GJ, Klein ML, Tuckerman M. Nosé–hoover chains: the canonical ensemble via continuous dynamics. *J Chem Phys*. 1992;97(4):2635-2643. https://doi.org/10.1063/1.463940.

41. Tuckerman ME, Berne BJ, Rossi A. Molecular dynamics algorithm for multiple time scales: systems with disparate masses. *J Chem Phys*. 1991;94(2):1465-1469. https://doi.org/10.1063/1.460004.

42. Tuckerman M, Berne BJ, Martyna GJ. Reversible multiple time scale molecular dynamics. *J Chem Phys*. 1992;97(3):1990-2001. https://doi.org/10.1063/1.463137.

43. Martyna GJ, Tobias DJ, Klein ML. Constant pressure molecular dynamics algorithms. *J Chem Phys*. 1994;101(5):4177-4189. https://doi.org/10.1063/1.467468.

44. Kennedy T. Managing the drug discovery/development interface. *Drug Discov Today*. 1997;2(10):436-444. https://doi.org/10.1016/S1359-6446(97)01099-4.

45. Ioakimidis L, Thoukydidis L, Mirza A, Naeem S, Reynisson J. Benchmarking the reliability of QikProp. Correlation between experimental and predicted values. *QSAR & Combinatorial Science*. 2008;27(4):445-456. https://doi.org/10.1002/qsar.200730051.

46. Jorgensen WL, Duffy EM. Prediction of drug solubility from structure. *Adv Drug Deliv Rev*. 2002;54(3):355-366. https://doi.org/10.1016/S0169-409X(02)00008-X.

47. Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol*. 2004;1(4):337-341. https://doi.org/10.1016/j.ddtec.2004.11.007.

48. Golbraikh A, Tropsha A. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. *J Comput Aided Mol des*. 2002;16(5–6):357-369.

49. Dalvi VH, Rossky PJ. Molecular origins of fluorocarbon hydrophobicity. *Proc Natl Acad Sci*. 2010;107(31):13603-13607. https://doi.org/10.1073/pnas.0915169107.

50. Pang X, Zhou L, Zhang M, et al. Two rules on the protein-ligand interaction. *Open Conf Proceed J*. 2012;3(1):70-80. https://doi.org/10.2174/2210289201203010070.

51. Correa-Basurto J, Ramos-Morales FR, Matus MH, et al. Docking and DFT studies to explore the topoisomerase II ATP pocket

employing 3-substituted 2,6-Piperazindiones for drug design. *Mol Simul*. 2012;38(13):1072-1084. https://doi.org/10.1080/08927022.2012.690877.

52. Kollman PA, Massova I, Reyes C, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res*. 2000;33(12):889-897. https://doi.org/10.1021/ar000033j.

53. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Structure, Function, and Bioinformatics*. 2004;55(2):383-394. https://doi.org/10.1002/prot.20033.

54. Amidon GL, Yalkowsky SH, Anik ST, Valvani SC. Solubility of nonelectrolytes in polar solvents. V estimation of the solubility of aliphatic monofunctional compounds in water using a molecular surface area approach. *J Phys Chem*. 1975;79(21):2239-2246. https://doi.org/10.1021/j100588a008.

55. Wang J, Hou T, Xu X. Recent advances in free energy calculations with a combination of molecular mechanics and continuum models.

Curr Comp Aided-Drug Des. 2006;2(3):287-306. https://doi.org/10.2174/157340906778226454.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Venugopal PP, Das BK, Soorya E, Chakraborty D. Effect of hydrophobic and hydrogen bonding interactions on the potency of ß-alanine analogs of G-protein coupled glucagon receptor inhibitors. *Proteins*. 2019;1–18. https://doi.org/10.1002/prot.25807