Molecular understanding of GPR120 agonist binding using homology modelling and molecular dynamics

Suyash Pant, V. Ravichandiran*

National Institute of Pharmaceutical Education and Research, Kolkata

***Corresponding author's Email ID:** directorniperkolkata@gmail.com

Abstract

The toll of type-2 diabetes and associated complication are continues, efforts to identify possible targets are ongoing. Free fatty acid receptor 4 (FFAR4/GPR120) has been recently identified to be a promising therapeutic target for a group of metabolic associated disorders. For the prevention of type 2 diabetes, significant scientific and commercial interest has been developed around GPR120 and its role. Due to unavailability of a crystal structure, the interaction dynamics of GPR120 agonists were not yet determined till date. In the present study, we constructed the homology model for GPR120 and validated using available mutational data and molecular dynamics simulation and explored its binding modes with known small molecule agonists. So, sixteen propionic acid derivatives as GPR120 agonists were collected to elucidate there binding modes. Experiential and theoretical studies suggested that carboxylic group of ligands interact with Arg99, is an important interaction for GPR120 activation. However, earlier reports also suggest that this interaction is not stable during the molecular dynamics simulation, which contradicts the experimental observations. Evidently to refute this, we got a stable interaction of Arg99 with TUG891 and other recently reported 15 GPR120 agonists. In addition we have also observed that in 1 μs molecular dynamics simulation Arg183 present in ECL2 tends to come inside and interact with ligand. Molecular dynamics simulation study provides a list of key hotspot residues which plays a important role in ligand binding. The homology model and results provides could be further utilized as a powerful template to accelerate the research in this field.

Key Word: GPR120; Molecular Dynamics; Induced fit docking; FFAR4, TUG891

Introduction

Diabetes is contriving worldwide scourge and difficulties to established researchers, to curb this long-lasting condition. After China, more than 77 million in India suffering from type 2 diabetes which makes our country as a second highest rate of diabetic patients [1]. Dietary fats and oils could be one of the major involving factor for the metabolic health issues [2]. Elevated unsaturated fats level in plasma suggests that insulin resistance connected with the level of free fatty acids (FFAs) before the onset of hyperglycaemia for example, constant aggravation, loss of pancreatic β-cells, and so forth [3,4]. However not all fats are bad, choosing a suitable fat or oil can help in lowering the risk for metabolic disorder, stroke or other major problems. Some of these free fatty acids or their derivatives act as an endogenous ligand for several proteins including G-protein coupled receptors (GPCRs) [2].

GPCRs mediating the effect of about one third of FDA approved molecules. The family of FFAs receptors, belongs to rhodopsin family and share only 10% sequence homology (GPR120/FFAR4) with FFAR1. Unlike GPR40, GPR120 got extensive consideration as a potential target for type-2 diabetes mellitus (T2DM) and inflammation [5-8]. It has several impacts on insulin secretion, lowering inflammation, and fat digestion etc. These are highly expressed in the digestive system (lower gut, enteroendocrine L cells), mature adipocytes and macrophages with low expression in muscle and hepatocytes [6]. Upon actuation, it animates the emission of incretins such as, glucagon-like peptide-1, which thusly builds insulin discharge by invigorating β-cells in pancreases [3,7,9]. GPR120 also plays a major role in modifying physiological functions such as, energy homeostasis, immunological homeostasis and neuronal functions [6].

However, even with the expansion in auxiliary data, ongoing appraisal and benchmark read for GPCR (G-protein coupled receptor) structure displaying philosophies, which shows trouble in anticipating local ligand restricting adaptation. In this work, we have used experimental mutational data with *in-silico* computational approach to develop homology model for GPR120 [10]. Although homology models are frequently used in drug discovery projects where protein structure was not resolved experimentally, identifying a correct binding site is still a challenge. To overcome the problem we analyzed frequently used GPR120 agonist TUG891 [9,11] and performed induced-fit docking to get the biologically significant protein-ligand conformation. Later the generated structures were

correlated with experimental mutational data. Best suited complex was further taken for molecular dynamics simulation and utilized as a template for molecular docking.

Molecular docking and molecular dynamics simulation provides valuable information of conformational dynamics of aromatic cage of GPR120 binding site. This particular aromatic cage formed via Trp277, Phe115, Trp104 and Phe304 plays important role in agonist binding and GPR120 activation [9,10]. We have also observed that there is a significant conformational change in the ECL2 (extracellular loop 2). Conformational changes in ECL2 was considered to be important in ligand binding [12,13]. 1 μs molecular dynamics trajectory was further clustered into ten representative complexes for further enrichment calculation. The best enrichment complex was taken further for molecular dynamics of 200 ns with 15 selected agonists obtained from patents and publications. The modelled structure was well validated and was observed to show good correlation with experimental data. There are several ongoing clinical trials to access the effect of synthetic as well as natural compounds which can activate the GPR120 [14].

Materials and Methods

Homology modeling

Protein sequence for short isoform of human GPR120, was obtained from UniprotKB Database (identifier: Q5NUL3-2) [15]. A systematic blast search was done in NCBI [16] server with other homologous proteins to find the suitable template for modeling. In the best case the sequence identity was 27% and 40% similarity. The model was built using MODELLER9.23 [17] [18]. Ten thousand output structures were generated and ten lowest DOPE (Discrete Optimized Protein Energy) score structures was minimized using Modrefiner server. Ramachandran plot was used to check for potential anomalies. Quality of created model was evaluated utilizing different validation tools such as, Rampage, PROCHECK, ProSA-web [19,20] and ERRAT [21] refer figure S2. PROCHECK was utilized to check stereo chemical quality using the dihedral edge circulation of amino acids. ERRATA score to foresee model unwavering quality by breaking down non-fortified iota molecule interaction.

Sitemap analysis

Since there is no reported crystal structure for GPR120, we have utilized the theoretical methods to identify the putative binding site for GPR120 agonists. Schrodinger

sitemap tool [22-24] was utilized to determine the possible binding site in the developed homology model of GPR120. Sitemap analysis and correlation of sitemap identified sites with existing mutational data was carried out to validate the theoretical results [10,24]. The size of a binding site was approximated by the number site points within the putative site. The predicted site score and druggability score (DScore) having value greater than one was used (there was only one such predicted site) for molecular docking of considered compounds [23,24].

Ligand dataset

A set of sixteen GPR120 agonists reported in several patents were considered in this work to validate the homology model as well as to understand the GPR120 agonists binding modes [25,26]. A decoy set (800 molecules) for GPR120 agonists (16 molecules) [3,9] was generated from DUD-decoy [27,28]. Schrodinger decoy dataset was also used for docking based enrichment calculation on three different modes of glide docking (virtual screening, Standard precision and extra precision).

Ligprep module from Schrodinger 2021-1was utilized to prepare the selected ligand and decoy database at pH 7.4 (Epik was utilized to identify the ionization states of ionisable functional groups in the ligand and decoy dataset) [29] and using OPLS4 force field [30].

Docking

The receptor grid was generated by defining residues from mutational data (L94, R99, W100, L114, F115, M118, T119, G122, T195, I197, F211, N215, W277, I280, I281, I284, F303, F304, V307, F311, F88, I126, I201, W207, V212, F216, T310) and also using site points obtained from Sitemap prediction [10]. Initial docking study was carried out using Glide SP and XP on all ten selected models [31]. However, the enrichment figures were not encouraging and in most of the ligands carboxylic head was not making any interaction with Arg99. Experimental data suggests Arg99 is crucial for ligand binding and mutating it abolishes the activity hence, considered important in choosing docking poses. Conventional docking approached consider ligand to be flexible and protein as a rigid object, however a getting a relevant ligand conformation is difficult. To overcome the problem induced fit docking (IFD) was performed to generate an appropriate model using TUG891 as a reference molecule [31]. Later it was observed that IFD complexes improve the enrichment results and correlation with mutational data.

Molecular Dynamics Simulation

All MD simulation calculations were performed in Desmond package (Schrodinger2021-1) using OPLS4 force field [30,32]. The system was generated by embedding the GPR120-ligand complexes into pre-equilibrated POPC (Palmitoyl Oleyol Phosphatidyl Choline) phospholipid bilayer and solvated with TIP3P water model. The overall charge of the system was neutralized by adding chloride ions. System was minimized for 500 ps utilizing steepest decent and limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) algorithms [32]. The convergence threshold for energy as set to 1.0 kcal mol^{-1} $Å^{-1}$. A default relaxation protocol of Desmond was used to equilibrate the system followed by 1 µs of production run. The NPT thermodynamic ensemble were used with semi-isotropic coupling style of Martyna-Tobias-Klien method (barostat) along with the Nose-Hoover chain (thermostat) [32]. A constant pressure and constant temperature of 1atm and 298.15K was applied to the system. Hydrogen positions were constrained by the M-SHAKE algorithm. Frame were recorded after every 100 ps and energy's were recorded after every 1.2 ps.

RMSD based clustering

Clustering becomes a major cornerstone in every structure based drug discovery projects to identifying structurally similar group of protein conformations in an molecular dynamics simulation. Multiple ensemble conformational structures were clustered and thus 20 clusters were generated from throughout simulation of 1 µs. A representative structure (with smallest RMSD) from each cluster was selected for enrichment calculation to identify best possible conformation for further docking based enrichment study. The best snapshot structure was selected based on the correlation analysis with experimental information i.e., which are correlating more with mutational data.

Result and discussion

Homology model of GPR120

It is known that initial protein conformation makes a major impact on docking results, hence an active state template was considered for homology modeling. The reported crystal structure of kappa-opioid receptor (PDB ID: 6B73) [33] is in active state, therefore generated

structure shows similarity with active state class A GPCR. The percentage identity between GPR120 and template sequences were found to be 27% which was more challenging, Figure S1. Modeller tool was utilized to perform homology modeling and ten thousand different structures were generated [17,34]. The top 100 structures were accessed based upon their lowest DOPE score. The model was assessed for its quality parameters such as, Ramachandran plot, prosaweb etc summarized in the supplementary materials (see supplementary materials, Figure S2). Finally careful visual analysis was done to identify top ten homology models for further minimization with Modrefiner server [35]. After minimization Ramachandran plot for GPR120 homology model Figure S2 has indicated that 100% of residues were located in allowed regions while 0% of residues were located in disallowed regions.

After looking for all the parameters the model was prepared using schrodinger 2021-1 protein preparation wizard. All hydrogen atoms were added, partial charges were assigned, a disulfide bond between cys111 and cys194 was created and protonation state for protein residues were identified at pH 7.4 and finally protein was minimized using OPLS4 force field. A thorough analysis molecular details were done to identify any structural error in the modeled structure. The exploratory perception from mutational information was mapped over the anticipated sitemap data.

The binding site found through the sitemap and mutational information were superimposed and found to be well correlated, structure appeared in figure 1 and Table S1. Site map analyses shows that the majority of the region of identified binding site was covered by hydrophobic residues. Deep inside the binding site there is an aromatic cage formed by Trp107, Trp277, Phe115, Phe211 and Phe304. Further Arg99 was identified flanging inside the cavity which was considered to be most important for ligand binding. Point mutation studies shows also sheds the importance of Arg99 and aromatic cage in GPR120 ligand binding. However, mutating Arg99 shows a complete loss in activity and mutation in aromatic cage shows decrease in activity. Structural analysis reveals that Arg99 is important to hold the ligand within the binding pocket. Earlier studies shows that during the start of molecular dynamics simulation ligand contact with Arg99 and instead Trp104 and Trp299 was found to form an stable contact with ligand carboxylic group.

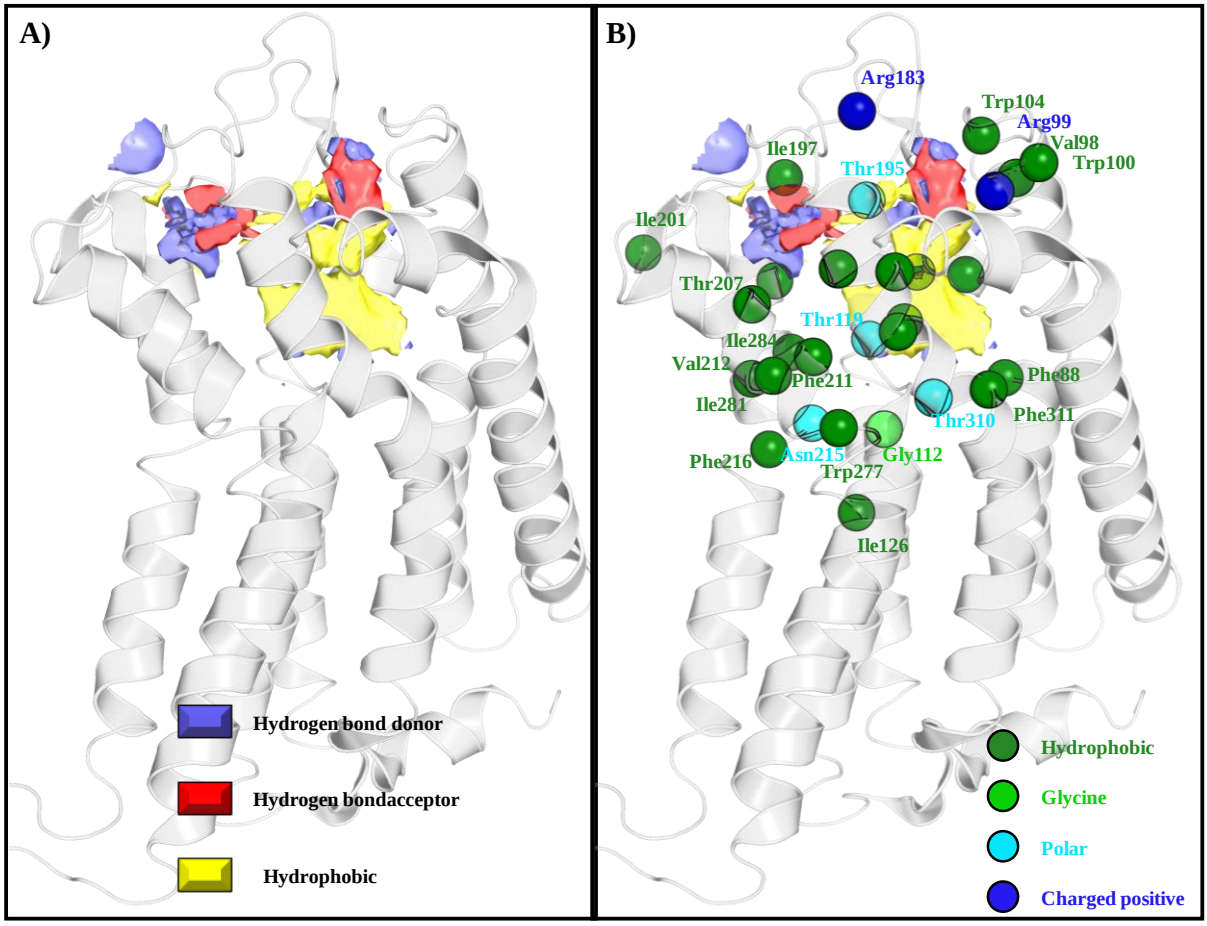

Figure 1. A) Theoretically identified putative binding site of GPR120 and B) Superimposed experimentally determined hot-spot residues (spheres) with identified putative binding site.

GPR120 agonists

After deorphanization of GPR120 in 2005 the first compound TUG891, derivative of phenyl propanoic acid was discovered. TUG891 was found to be selective for GPR120 over GPR40 (free fatty acid receptor 1) [36]. Free fatty acids like, linoleic acid and docosahexaenoic acids are the endogenous ligands for GPR120 [37]. Over the year several researchers are trying to mimic the structure of these free fatty acids to design synthetic GPR120 agonists. Figure 2 provides a detail structural diversity of propanoic acid derivatives identified as GPR120 agonists in recent time. GPR120 agonists posses a polar head, a carboxylic side chain connected with two carbon atom chain as a linker with an aromatic ring again connected with an aromatic or heteroaromatic ring system with an carbon linker which further extend its contact with an aliphatic tail, figure 2. SAR analysis of different series suggests that carboxylic acid group is crucial for the activity and making any changes in it diminishes the activity [38]. Presence of an C-O linkage is much favoured than an C-N linkage between the two rings in figure 2 [38]. All these information can provide an better

insight to design a better GPR120 agonist. Other fifteen molecules were also collected from the literature to get the better insight of GPR120 agonist binding [25,26] Figure 1.

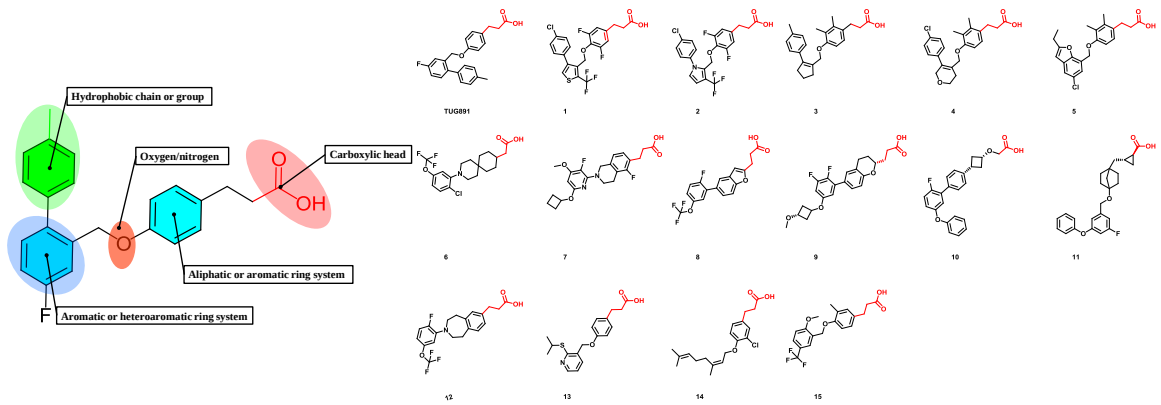

Figure 2. Basic pharmacophore skeleton of GPR120 agonist and selected molecules in this study.

Molecular docking

The refined homology model shares most of the common features of class A GPCRs and all the Ramachandran outliers were also fixed other than in the loop region. While using the glide module for molecular docking we found most of the time carboxylic group of ligand was not making any interaction with Arg99 or either carboxylic group was flipped inside the binding site. As conventional docking methods only considers ligand as flexible which reduces the chance of getting a correct binding pose of TUG891 and other molecules utilized in this study. Therefore, we have utilized induced fit docking (IFD) methodology for generating most suitable binding pose of considered ligand. IFD was done only for single ligand (TUG891) as it generate multiple number of poses and we have sufficient mutational information for TUG891, which help in selection of an ideal pose. Later this IFD pose was consider for glide docking of other ligands.

IFD calculation generated eighty protein-ligand complex, H-bond interaction with Arg99 and carboxylic group of ligand was considered as a pre filter for pose selection. Other interactions with aromatic residues like, Trp107, Trp277, Phe115, Phe211 and Phe304 were consider to choose the best binding pose. The best pose was further subjected for MD simulation. later the pose was also utilized for enrichment calculation utilizing the considered molecules in this study. Figure 3 provided the docking poses of all the sixteen compounds into GPR120 binding site.

Docking studies on all the representative structures from molecular dynamics trajectory shows that majority of time these molecules carboxylic group placed towards the Arg99. Red

arrow in figure 3, GPR120 binding site shows how all the ligands were ligand into the binding site, keeping carboxylic head towards the Arg99. All the considered compounds were able to keep at H-bond interaction between carboxylic head group and Arg99. Other interactions include pi-pi stacking between the aromatic residues and ligand aromatic rings. As suggested by site map analysis the binding site of GPR120 was very hydrophobic from inside, figure 3 provides the residues (green colour) which makes an suitable hydrophobic interaction with ligand. Looking into the docking scores from all the 10 clusters we found that a few times docking scores drops to minimum for some molecules, Table 1. this was due to flipping of molecule and placement of carboxylic group into the hydrophobic pocket, which cause penalty into the docking score.

Docking poses from cluster 3 was chosen for molecular dynamics studies, as its enrichment score was fairly good and docking poses were well correlated with the existing data. Later on all these complexes were further taken for molecular dynamics simulation to better understand the stability of these interactions.

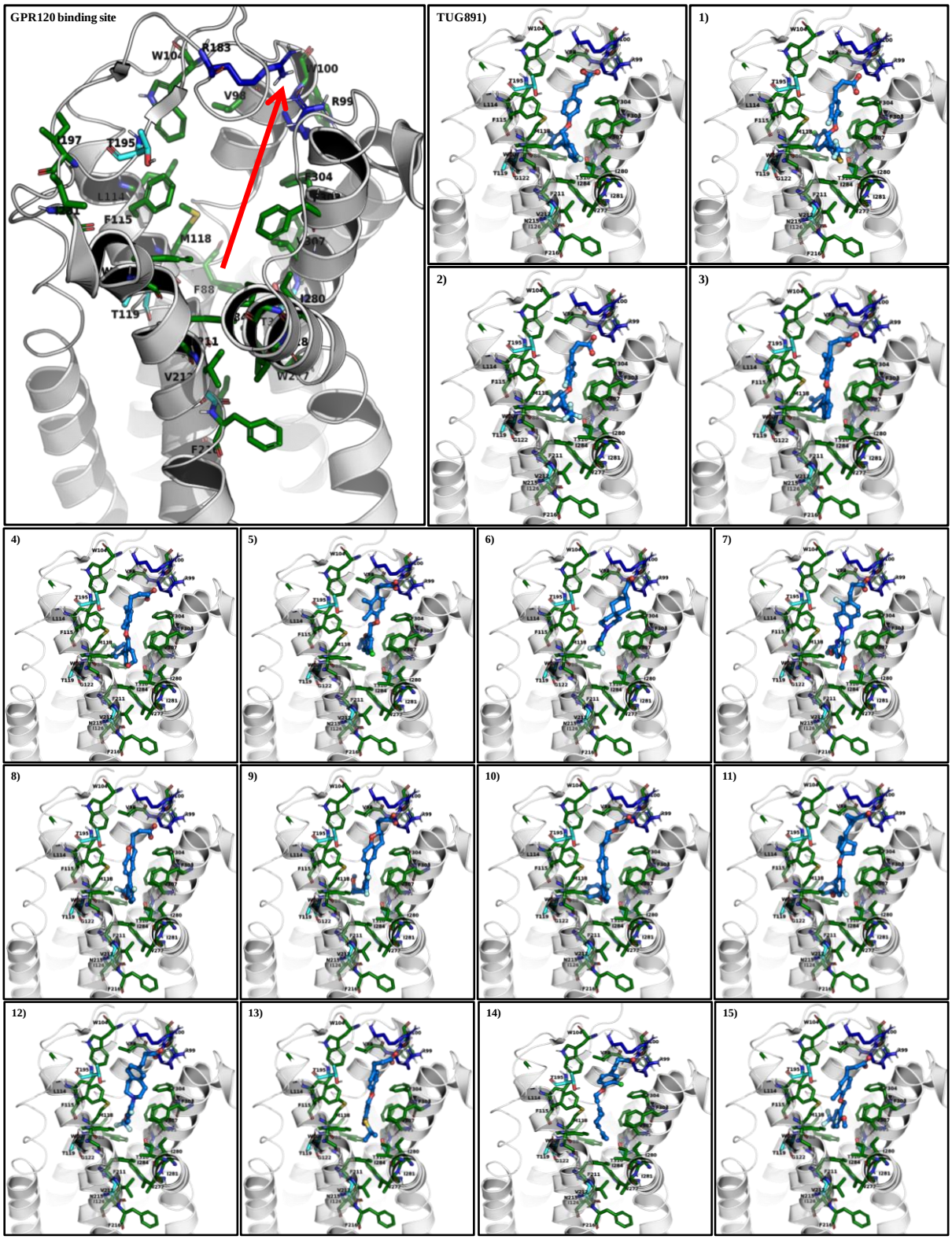

Figure 3. Molecular docking poses of all the compounds used in this study. residues were colour coded based on Hydrophobic (Green), polar (water blue), positive ionisable (dark blue) and glycine was shown in grey colour.

Table 1: Molecular docking results on ten identified clusters from 1 μs molecular dynamics of TUG891-GPR120 complex.

Compound	Activity data (nm)	DS1	DS2	DS 3	DS4	DS5	DS6	DS7	DS8	DS9	DS10
TUG891	6.93[a] and 7.19[b]	-11.73	-11.37	-11.08	-11.52	-10.29	-11.6	-11.64	-10.91	-12.4	-11.56
1	40[a] and 143[b]	-12.08	-11.13	-11.87	-9.62	-10.11	-11.1	-11.52	-11.87	-11.25	-12.1
2	80[a] and 69[b]	-12.02	-11.69	-11.79	-10.57	-8.42	-10.18	-11.1	-10.49	-11.7	-12.64
3	24[a] and 63[b]	-11.47	-10.81	-10.82	-10.79	-9.37	-11.1	-9.74	-10.23	-12.06	-11.9
4	110[a] and 243[b]	-11.26	-10.04	-10.94	-8.9	-9.84	-10.58	-10.84	-10.36	-11.97	-10.81
5	23[a] and 94[b]	-8.56	-10.58	-10.77	-11.3	-9.34	-9.53	-10.31	-9.26	-10.8	-9.96
6	98[c]	-9.44	-8.71	-6.29	-6.91	-6.68	-8.59	-7.94	-7.43	-8.99	-9.17
7	9[c]	-7.73	-8.84	-8.93	-5.04	-6.61	-8.03	-7.68	-7.32	-9.51	-8.79
8	35[c]	-10.54	-9.16	-10.48	-10.12	-8.06	-8.78	-7.83	-8.3	-10.89	-10.48
9	35[c]	-10.45	-10.05	-11.34	-9.56	-8.76	-8.65	-9.16	-7.44	-10.85	-11.03
10	170[d]	-11.56	-10.47	-11.93	-9.24	-10.24	-10.39	-10.5	-9.89	-11.78	-10.92
11	600[d]	-11.53	-11.13	-10.95	-10.67	-10.4	-7.03	-10.98	-6.6	-11.88	-10.47
12	7.6[b]	-10.15	-9.02	-9.51	-8.92	-7.11	-5.43	-9.12	-8.13	-8.81	-8.77
13	33[a]	-7.15	-6.67	-6.76	-6.77	-7.6	-8.2	-9.42	-6.52	-5.49	-7.06
14	570[a]	-10.35	-8.69	-8.67	-6.67	-8.23	-7.75	-7.77	-6.88	-8.72	-10.09
15	299[a]	-10.69	-9.97	-10.5	-9.78	-8.03	-8.46	-9.3	-9.8	-11.36	-9.87

#DS 1-10: Docking score from each consecutive cluster after 1 μs of simulation of TUG891-GPR120 complex

a: Calcium influx activity assay with CHO cells expressing hGPR120

b: BRET-based β-arrestin-2 recruitment assay

c: Inositol phosphate turnover assay

d: Measurement of phosphorylation of ERK

Molecular Dynamics

Molecular dynamics simulation of GPR120-TUG891 was done for 1 µs and with other molecules for 200 ns. RMSD and RMSF plots for TUG891 was within 3 Å which usually considered as an acceptable limit. However, if we see in the beginning of the simulation, 0-250 ns ligand conformation was changing which was shown in figure 4A. This shows that the complex was trying to attain a stable conformation. Later 400 ns the system was stabilized and minor fluctuations were observed. After 500 ns of simulation we observed that ECL2 of GPR120 was moving inside the binding site and Arg183 was making an stable interaction with the ligand carboxylic group. ECL2 was considered very important for ligand binding and unbinding for GPCRs and its length and secondary structure varies within the GPCR class [12,39]. ECL2 was usually considered as an gatekeeper and keeps the binding site open for ligand and closes it like a lid after ligand sits into the binding site [12,39]. Figure 4A and 4B shows the transition of ECL from outside to inside of binding site with respect to simulation time (blue to red). Later 200 ns simulation of TUG891 provided the similar results along with other 15 molecules. Apart from a movement of ECL2, we also observed there was slight change in TUG891 conformation. During the beginning of the simulation we observed higher RMSD change in ligand but after 400 ns it attain a stable conformation and stayed there till the end of the simulation. GPR120 favours more hydrophobic interactions then H-bond or other non-covalent interactions. Arg99, Arg183, Phe88 and Phe115 were having more than 50% interaction fraction during the 1 µs of MD simulation Figure 4C. According to the earlier studies on GPR120 homology models, Arg99 was not found to be stable and the interaction was lost in the start of the simulation. But in this case we found an strong and stable interaction of Arg99 with the ligand. Later we clustered the 1µs trajectory and took the intermediate structure to model the binding modes of other fifteen molecules.

Molecular dynamics simulation of TUG891 shows that how the ECL2 was moving from outside to inside, Figure 2. Figure 2 A shows how the conformational change of residues important in ligand binding happening during the simulation. Initially the Arg183 was away and later it started changing its conformation shown in blue to red colour. Ligand interaction fraction also shows how the Arg99 was stable during the simulation and making multiple interaction with COOH group of ligand, Figure 4C. As Arg183 was not making any interaction from the beginning, hence accounts for only 75% of interaction fraction, Figure 4C.

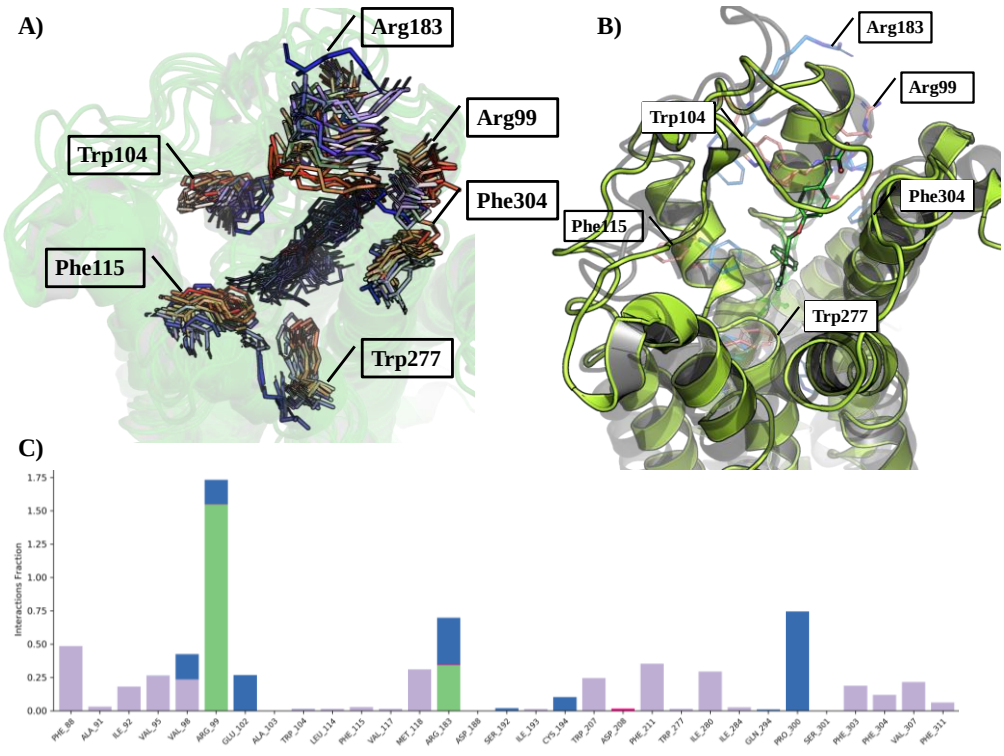

Figure 4. Molecular dynamics simulation data for TUG891 **A)** 1 µs simulation cluster superimposed into first pose (blue to red) **B)** Superimpose pose of first o ns (transparent black cartoon) to 1 µs (Green cartoon) and **C)** Interaction fraction of residues in contact with TUG891 throughout the 1 µs simulation grey (hydrophobic), Green (H-bond), Pink (ionic interaction) Blue (water bridge).

1 µs trajectory was clustered into ten representative complexes and utilized for docking based enrichment studies. The complex showing best enrichment was again utilized for 200 ns molecular dynamics simulation. In this regard TUG891 was again simulated for 200 ns to look for the previously generated result are consistent or not. However, the interactions identified in the 1 µs simulation were originally retained during the 200 ns simulation as well. The RMSD of the 1 µs and 200 ns of TUG891 complexes were shown in Figure 3, which shows that both the complexes were stable in the simulation. We have also shown the change in ligand orientation in 1 µs simulation in figure 3. RMSF plots for all the simulations were given in the supplementary file.

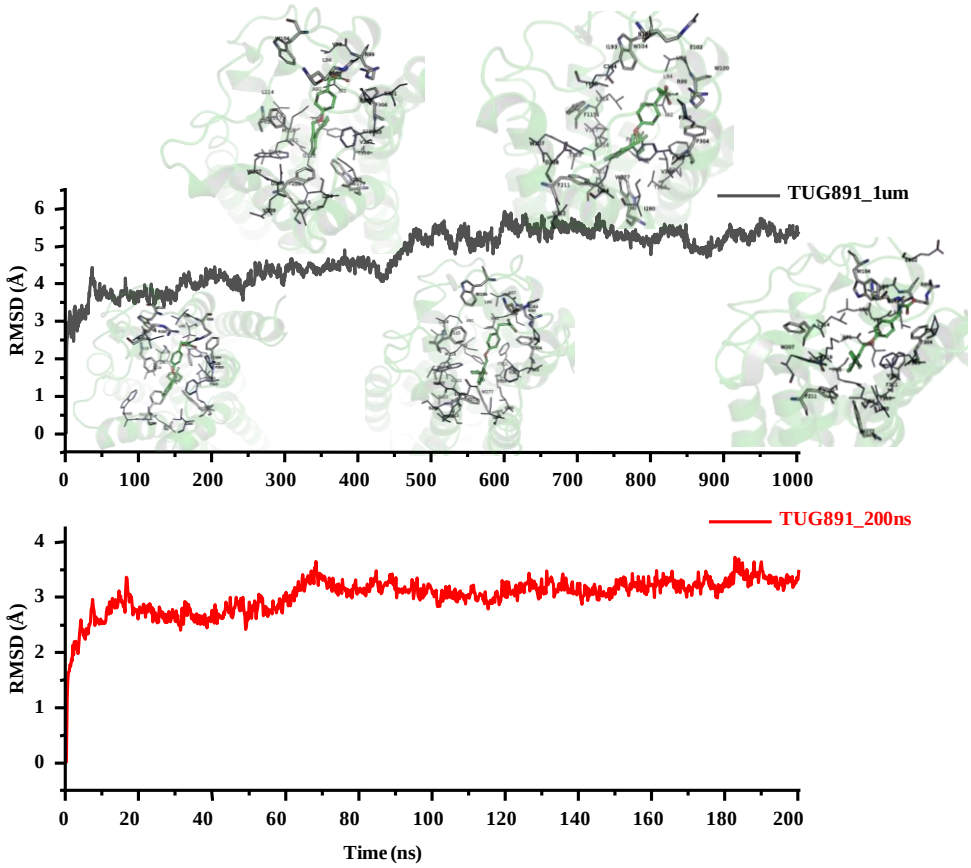

Figure 5. RMSD of TUG891 1 µs and later 200 ns from clustered complex.

Looking into the structure similarity, figure 1 the head group of all the compounds consists of propanoic acid and vary in the tail region. RMSD plot of all the compounds shows that the complexes were stable during the simulation, with minor fluctuations in few complexes. The RMSD and RMSF plot for all the fifteen complex along with a 200 ns simulation of TUG891 was shown in figure 6-9.

Compound 1-4 shows a decent structural similarity with the TUG891 and there binding modes were also very similar. The aromatic part in the tail region is well settled in the hydrophobic core of protein supported by RMSD and RMSF. However, compound 3 shows higher fluctuation in RMSD which was going to 5.6 Å, figure 6. Arg99 interaction was found throughout the simulation as shown in 2D interaction diagram, Figure 4 and protein ligand simulation contact map in Figure S3. Simulation contact map of protein-ligand shows, that aromatic residues like, Phe211 and Trp207 were stable along with few other hydrophobic residues. Ligand 3 was found to make good contact in the beginning of simulation with Trp207 till 100 ns. Later up to 200 ns only Arg99 and Arg183 interactions were retained, this could be a reason for higher RMSD for ligand 3 complex.

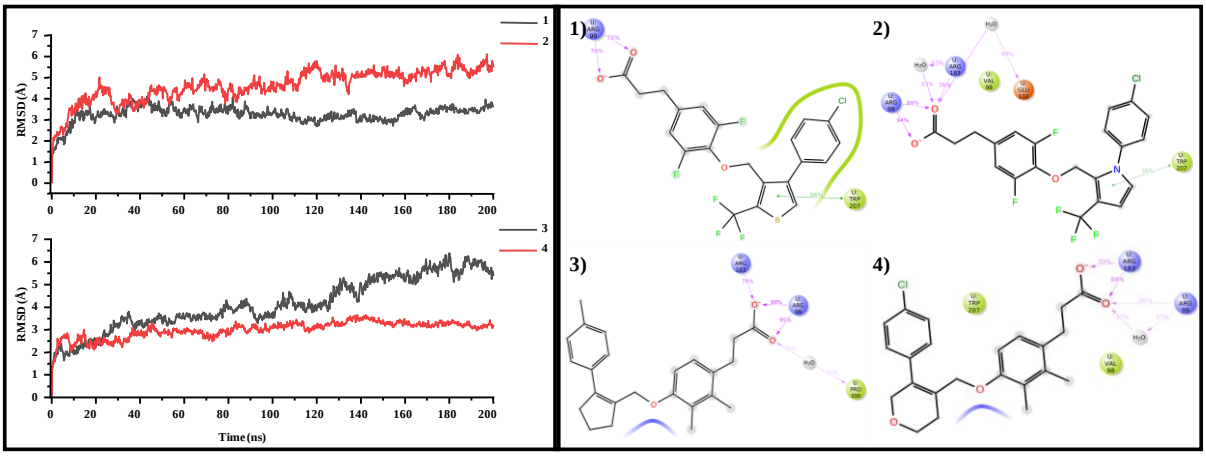

Figure 6. RMSD and ligand interaction diagram of compound 1-4.

Ligand 5, 7 and 8 hydrophobic tail region was firmly accommodated into the aromatic hydrophobic cluster as shown in figure 7. Due to the Spiro ring in ligand 6 was very rigid into the binding site and accept the carboxylic group, rest of the ligand was not observed to made any interaction with protein, shown in figure 7. Ligand 6 was only making stable interaction with Val98, Arg99 and Glu102, other interactions were not observed throughout the simulation, figure S3.

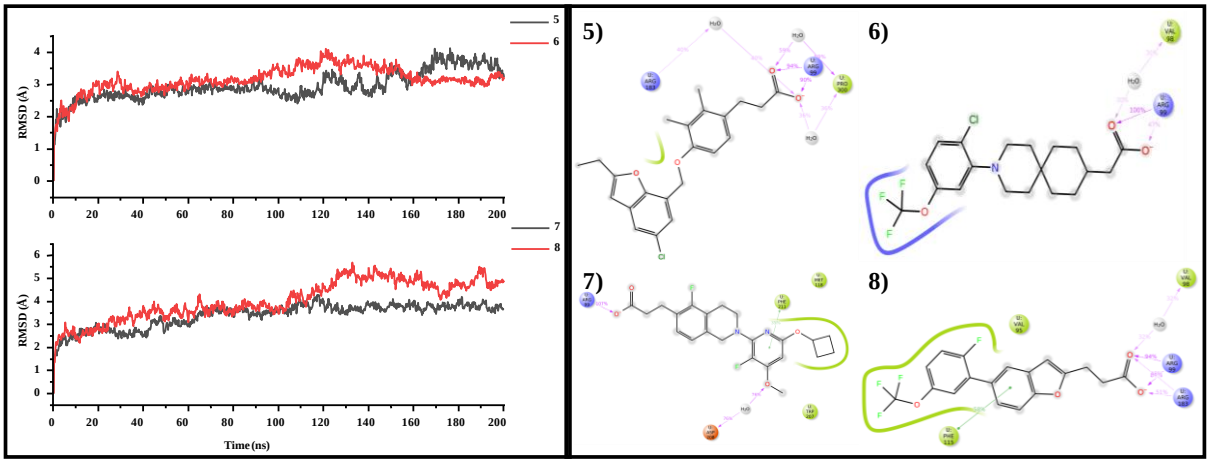

Figure 7. RMSD and ligand interaction diagram of compound 5-8.

Ligand 9-12 aliphatic tail part was well settled into the hydrophobic region of the protein figure 8. Figure 8 (9, 11 and 12) also shows that Arg183 (ECL2) was also making an stable contact with the carboxylic head portion of the ligand for more than 60 percent time. In case of ligand 10 Arg183 was interacting till 100 ns of simulation and later the interaction count was reduced, figure S3. The reason was after 100 ns ligand 10 changes its conformation and the two phenyl ring pulled the ligand little deeper inside the binding, which takes carboxylic head little far from Arg183.

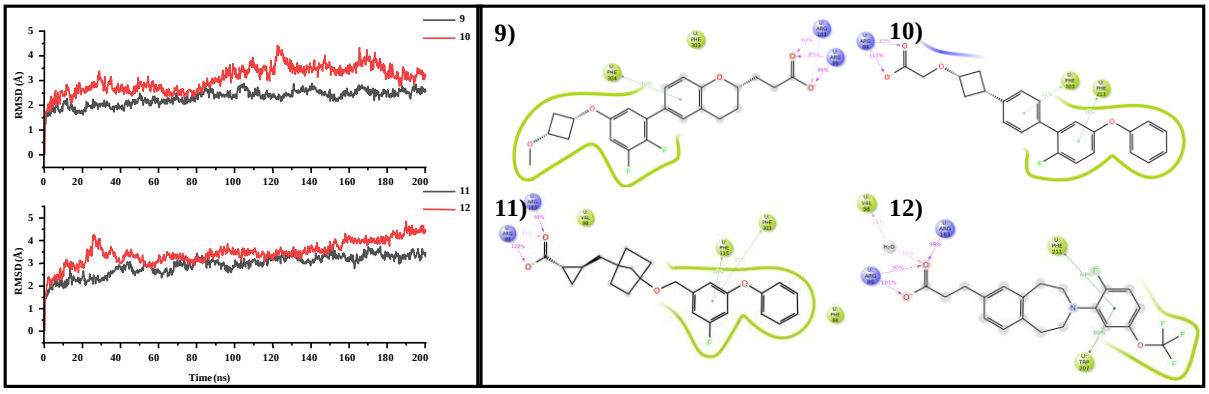

Figure 8. RMSD and ligand interaction diagram of compound 9-12.

Ligand 13-15 were also relatively stable as attain the equilibration after 100 ns of simulation run. Figure 9 shows that the aliphatic tail accommodated well inside the hydrophobic pocket of protein. Figure S3 shows that in the beginning of simulation ligand 14 aliphatic tail was not correctly placed in the binding site due to higher flexibility. But later 75 ns of simulation we observed large number of hydrophobic interaction between aromatic cage of the protein with aliphatic tail of ligand.

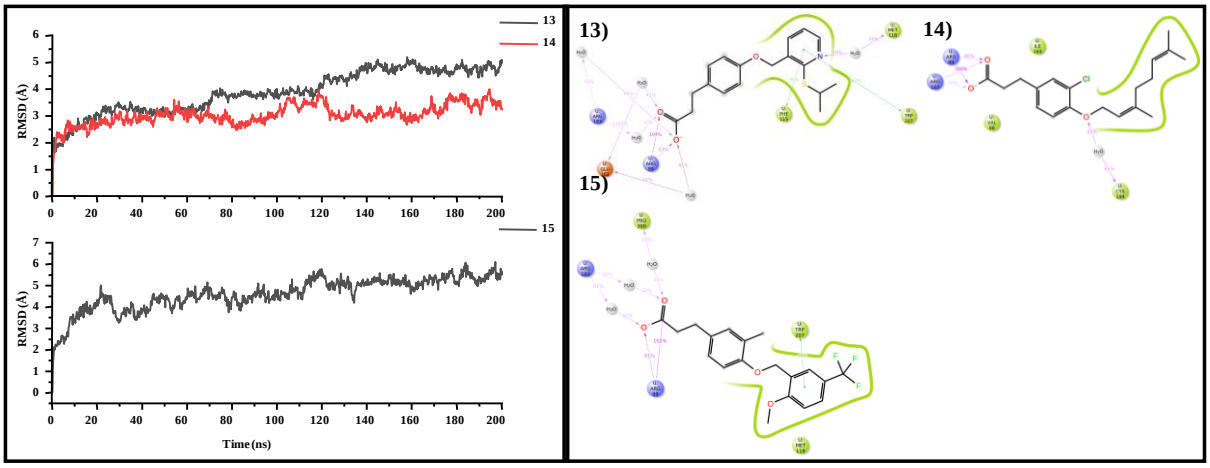

Figure 9. RMSD and ligand interaction diagram of compound 13-15.

All the considered ligands were found to interact with important residues reported in the literature. In addition we have also found the interaction of Arg183 with the ligand, which was a part of a loop. Figure S5 shows the final snapshot of all the ligands after 200 ns of simulation.

Conclusion

Conventional therapeutic options available to alleviate the insulin level in type 2 diabetes increases risk of hypoglycaemia. Recently studied incretin mimetic were found to show better glycemic control and reduces the risk of hypoglycaemia. GPR120 has been a keen interest for an attractive therapeutic target for metabolic syndromes, including diabetes. However, lack of crystal structure hinders the better understanding of protein-ligand binding modes and its key interactions. Here, we modelled homology model surrounding interaction with Arg99 which was earlier confirmed using several point mutation studies. Apart from Arg99 several other important residues identified by point mutations were also analyzed in this study. The most studied GPR120 agonist, TUG891 was considered to improve the homology model using induced fit docking followed by 1 μs of molecular dynamics. In earlier studies on GPCRs it was found that after ligand binding ECL2 changes its conformation and closes the binding pocket. Our results also shows similar observations, how the ECL2 comes inside and closes the binding site of GPR120 after 500 ns of MD simulation. Earlier studies have failed to establish an correlation with experimental and theoretical findings. However, in our study the Arg99 was stable and the interaction was consistent throughout the other 15 ligand taken in the study. To confirm this interaction reproducibility fifteen other GPR120 agonists were also taken, where the interaction was intact throughout the 200 ns of simulation. Other residues like, Val98, Trp207, Phe211, Arg183, Phe303, Ile193 and Phe211 were also found to be important for ligand binding. Overall results gives an better insight of GPR120 agonist binding site and shows a better agreement of computational models with experimental findings. Developed homology model could be further utilized for virtual screening and molecular docking study to design potent and novel GPR120 agonists.

Acknowledgments

SP is thankful to NIPER, Kolkata, IIT-Guwahati, NIPER-Guwahati and the Department of Pharmaceuticals, Ministry of Chemicals and Fertilizers, Government of India, for providing necessary resources and fellowship. We acknowledge the contributions of the late Dr. Hemant Kumar Srivastava to this article and want to dedicate this work to his memory.

References

1. WHO (2020)
2. Menendez JA, Mehmi I, Atlas E, Colomer R, Lupu R (2004) Int J Oncol 24(3):591
3. Kimura I, Ichimura A, Ohue-Kitano R, Igarashi M (2020) Physiol Rev 100(1):171
4. Clifton P (2019) Nutrients 11(7)
5. Karakula-Juchnowicz H, Rog J, Juchnowicz D, Morylowska-Topolska J (2017) Postepy Hig Med Dosw (Online) 71(0):942
6. Ichimura A, Hara T, Hirasawa A (2014) Front Endocrinol (Lausanne) 5:111
7. Lay AC (2021) Endocrinology 162(10)
8. Bianchini G, Nigro C, Sirico A, Novelli R, Prevenzano I, Miele C, Beguinot F, Aramini A (2021) Biomed Pharmacother 139:111613
9. Carullo G, Mazzotta S, Vega-Holm M, Iglesias-Guerra F, Vega-Perez JM, Aiello F, Brizzi A (2021) J Med Chem 64(8):4312
10. Hudson BD, Shimpukade B, Mackenzie AE, Butcher AJ, Pediani JD, Christiansen E, Heathcote H, Tobin AB, Ulven T, Milligan G (2013) Mol Pharmacol 84(5):710
11. Schilperoort M, van Dam AD, Hoeke G, Shabalina IG, Okolo A, Hanyaloglu AC, Dib LH, Mol IM, Caengprasath N, Chan YW, Damak S, Miller AR, Coskun T, Shimpukade B, Ulven T, Kooijman S, Rensen PC, Christian M (2018) EMBO Mol Med 10(3)
12. Wheatley M, Wootten D, Conner MT, Simms J, Kendrick R, Logan RT, Poyner DR, Barwell J (2012) Br J Pharmacol 165(6):1688
13. Unal H, Jagannathan R, Bhatnagar A, Tirupula K, Desnoyer R, Karnik SS (2013) J Biol Chem 288(1):540
14. Son SE, Kim NJ, Im DS (2021) Biomol Ther (Seoul) 29(1):22
15. UniProt C (2021) Nucleic Acids Res 49(D1):D480
16. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL (2008) Nucleic Acids Res 36(Web Server issue):W5
17. Webb B, Sali A (2016) Curr Protoc Bioinformatics 54:5 6 1
18. Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A (2000) Annu Rev Biophys Biomol Struct 29:291
19. Wiederstein M, Sippl MJ (2007) Nucleic Acids Res 35(Web Server issue):W407
20. Sippl MJ (1993) Proteins 17(4):355
21. Colovos C, Yeates TO (1993) Protein Sci 2(9):1511
22. Schrödinger L, New York, NY (2021)
23. Halgren TA (2009) J Chem Inf Model 49(2):377
24. Halgren T (2007) Chem Biol Drug Des 69(2):146
25. Halder S, Kumar S, Sharma R (2013) Expert Opin Ther Pat 23(12):1581
26. Zhang X, Macielag MJ (2020) Expert Opin Ther Pat 30(10):729
27. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK (2012) J Med Chem 55(14):6582
28. Huang N, Shoichet BK, Irwin JJ (2006) J Med Chem 49(23):6789
29. Greenwood JR, Calkins D, Sullivan AP, Shelley JC (2010) J Comput Aided Mol Des 24(6-7):591
30. Harder E, Damm W, Maple J, Wu C, Reboul M, Xiang JY, Wang L, Lupyan D, Dahlgren MK, Knight JL, Kaus JW, Cerutti DS, Krilov G, Jorgensen WL, Abel R, Friesner RA (2016) J Chem Theory Comput 12(1):281
31. Schrödinger L, New York (2021)
32. Proceedings of the 2006 ACM/IEEE conference on Supercomputing. Tampa, Florida: Association for Computing Machinery, 2006
33. Che T, Majumdar S, Zaidi SA, Ondachi P, McCorvy JD, Wang S, Mosier PD, Uprety R, Vardy E, Krumm BE, Han GW, Lee MY, Pardon E, Steyaert J, Huang XP, Strachan RT, Tribo AR, Pasternak GW, Carroll FI, Stevens RC, Cherezov V, Katritch V, Wacker D, Roth BL (2018) Cell 172(1-2):55
34. Fiser A, Do RK, Sali A (2000) Protein Sci 9(9):1753
35. Xu D, Zhang Y (2011) Biophys J 101(10):2525

36. Suzuki T, Igari S, Hirasawa A, Hata M, Ishiguro M, Fujieda H, Itoh Y, Hirano T, Nakagawa H, Ogura M, Makishima M, Tsujimoto G, Miyata N (2008) J Med Chem 51(23):7640
37. Shimpukade B, Hudson BD, Hovgaard CK, Milligan G, Ulven T (2012) J Med Chem 55(9):4511
38. Xu F, Zhao Y, Zhou H, Li C, Zhang X, Hou T, Qu L, Wei L, Wang J, Liu Y, Liang X (2020) Bioorg Med Chem Lett 30(24):127650
39. Wifling D, Bernhardt G, Dove S, Buschauer A (2015) PLoS One 10(1):e0117185

Graphical abstract