Journal of Molecular Graphics and Modelling 28 (2010) 766–774

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Molecular docking and molecular dynamics simulation studies of GPR40 receptor–agonist interactions

Shao-Yong Lu [a], Yong-Jun Jiang [b,*], Jing Lv [a], Tian-Xing Wu [a], Qing-Sen Yu [a], Wei-Liang Zhu [c,**]

[a] Department of Chemistry, Zhejiang University, Hangzhou, Zhejiang 310027, China
[b] Key Laboratory for Molecular Design and Nutrition Engineering, Ningbo Institute of Technology, Zhejiang University, No. 1 Qianhunana Road, Ningbo 315104, China
[c] Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zuchongzhi Road, Shanghai 201203, China

ARTICLE INFO

Article history:
Received 5 November 2009
Accepted 1 February 2010
Available online 7 February 2010

Keywords:
GPR40
Molecular docking
Molecular dynamics simulation
AutoDock
GROMACS

ABSTRACT

In order to explore the agonistic activity of small-molecule agonists to GPR40, AutoDock and GROMACS software were used for docking and molecular dynamics studies. A molecular docking of eight structurally diverse agonists (six carboxylic acids (CAs) agonist, and two non-carboxylic acids (non-CAs) agonist) was performed and the differences in their binding modes were investigated. Moreover, a good linear relationship based on the predicted binding affinities (pK_i) determined by using AutoDock and experimental activity values (pEC50) was obtained. Then, the 10 ns molecular dynamics (MD) simulations of three obtained ligand–receptor complexes embedded into the phospholipid bilayer were carried out. The position fluctuations of the ligands located inside the transmembrane domain were explored, and the stable binding modes of the three studied agonists were determined. Furthermore, the residue-based decomposition of interaction energies in three systems identified several critical residues for ligand binding.

1. Introduction

Type 2 diabetes, accounting for 90–95% of all diabetes, is linked to obesity and is characterized by improper insulin secretion and insulin resistance [1]. Diabetes leads to several complications including cardiovascular disease, diabetic retinopathy, lipid disorders, and hypertension. It is estimated that there would be, by 2010, as many as 220 million people in the world projected to suffer from this debilitating disease [2]. Since glycemia control of type 2 diabetes often deteriorates in spite of aggressive treatment, there is today an urgent active search for novel therapy. An important strategy for treatment of type 2 diabetes is to stimulate insulin secretion.

The G-protein coupled receptor GPR40, recently named free fatty acid receptor 1 (FFAR1), is predominantly expressed in human and rodent pancreatic islets and is activated by physiological concentrations of medium- and long-chain free fatty acids (FFAs) found in plasma such as linoleic and palmitic acids [3]. Recently, three independent researches provide evidences that FFAs amplify the glucose-stimulated insulin secretion (GSIS) by the pancreatic β-cell through the activation of GPR40 [3–5]. Given the need for novel treatments for type 2 diabetes, GPR40 represents a potentially attractive target. Acute administration of FFAs stimulates insulin release. Conversely, chronic exposure to high levels of FFAs leads to the impairment of β-cell function and lipotoxicity [6]. Hence, it is unclear whether the agonists or antagonists of GPR40 could be applied to the treatment of type 2 diabetes. The development of both potent agonists and antagonists of GPR40 are therefore required so that the detailed mechanism of the receptor in glucose balance may be unraveled. However, more recent researches are in favor of agonist therapy [7–9].

A large number of synthetic agonists of GPR40 have been proposed during the past 5 years [10–15]. Generally, the structure of the carboxylic acids (CAs) is used as a starting point for the design of new agonists. Additionally, some non-carboxylic acids (non-CAs) were proposed as effective agonists [16,17]. Recently, Bharate et al. have published reviews on progress in the discovery and development of small-molecule modulators of GPR40 [18]. They divided small-molecule GPR40 agonists into seven chemical classes.

Like other G-protein coupled receptors belonging to the same cluster of family A GPCRs, GPR40 are membrane proteins. All of them share a common structural motif of seven putative α-helical transmembrane spanning regions connected by three extracellular and three intracellular hydrophilic loops, an extracellular N terminus and intracellular C-terminal tail [19]. Such macromolecules are not easily amenable to crystallization and, therefore, to

* Corresponding author. Tel.: +86 574 88229516; fax: +86 574 88229516.
** Corresponding author.
E-mail addresses: yjjiang@nit.zju.edu.cn (Y.-J. Jiang), wlzhu@mail.shcnc.ac.cn (W.-L. Zhu).

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.02.001

precise structure elucidation via X-ray diffraction. Thus, research focus has been on alternative techniques, such as, molecular modeling of GPCRs, either through homology modeling technique, or *de novo* model building, in order to derive reasonable 3D structures which can be used in structure-based drug design. Recently, Tikhonova et al. published the first structural model of the binding site of GPR40 complex with GW9508, which was obtained through a bidirectional, iterative approach that composed of molecular modeling and site-directed mutagenesis [20]. Twelve residues were identified within the GPR40 binding pocket.

Elucidation of ligand binding mechanisms is the necessary step to obtain more selective and potent drugs for this new potential target. Up to now, the correlation of agonist structure and its agonistic activity, the binding modes of non-CA agonists and GPR40, the binding energy, and the dynamics stability are still unknown. Therefore, in this paper, we made a molecular docking and molecular dynamics study to locate the binding site, get their dynamics information, and further identify the critical amino acid residues for ligand binding.

2. Methods and materials

2.1. Small-molecules preparation

We mainly selected eight small-molecule agonists which bear structure diversity including six classes of GPR40 agonists (class iii of GPR40 agonists is eliminated as they lack exact activity values). In order to get the most stable agonists conformations, the structure-optimizing calculation was carried out at the 6-31G** level by employing the Becke three-parameter Lee–Yang–Parr (B3LYP) hybrid density functional theory using the quantum chemistry software Gaussian 03 [21], and the structures with the lowest energy were selected for the following docking study. When docking, the Gasteiger–Hückel atomic charge was chosen for small-molecule agonists.

2.2. GPR40 model preparation

The homology model for GPR40 constructed based on the 2.65 Å resolution structure of rhodopsin (PDB code 1GZM) has been reported by Tikhonova et al. [20]. The binding site for GW9508 was experimentally verified and the coordinates of the validated GPR40 complex with GW9508 are available online [20]. Thus, we employ the Tikhonova's trusted model in the molecular docking study. The GW9508 was deleted and the GPR40 structure was then used in the docking experiments.

2.3. Molecular docking

Molecular docking of CA and non-CA agonists to the GPR40 model was carried out using the AutoDock3.0.5 software package [22]. All the torsion angles in the small-molecules were set free to perform flexible docking. Polar hydrogen was added by using the Hydrogen module in AutoDock Tools (ADT) for GPR40. After that, Kollman united atom partial charges were assigned for the receptor.

The empirical free energy function and Lamarckian genetic algorithm (LGA) were used for docking with the following settings: a maximum number of 25,000,000 energy evaluations, an initial population of 150 randomly placed individuals, a maximum number of 27,000 generations, a mutation rate of 0.02, a crossover rate of 0.80 and an elitism value (number of top individuals that automatically survive) of 1. For the local search, the so-called Solis and Wets algorithm was applied with a maximum of 300 iterations per search. Fifty independent docking runs were carried out for each ligand. Results were clustered according to the root-mean-square deviation (RMSD) criterion. The best docked conformations of small-molecule agonists were selected as initial active/binding conformations to evaluate potential correlations between experimental activities and predicted log K_i values.

2.4. Molecular dynamics simulations

Molecular dynamics (MD) simulations were performed using the GROMACS 3.3.3 package with the standard GROMOS96 force field [23,24]. The obtained complexes of GPR40 with three agonists (GW9508, compound **5** and compound **7**) were used for performing MD simulations. The model of the POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine) bilayer was used for simulation of the phospholipid environment around the receptor. The lipid parameters were taken from the literature [25–27]. The partial atomic charges of three agonists were calculated at DFT/B3LYP/6-31G** level using Gaussian 03 [21] package. Topology file and other force field parameters except the charges of ligands were generated using the PRODRG program [28]. Fig. 1, as an example, shows the GPR40-GW9508 complexes in the POPC/water systems. The three systems were neutralized by adding Cl^- counterions by replacing water molecules, respectively. The energy of these complexes was minimized using the steepest descent approach realized in the GROMACS package. Then, a 100 ps position restraining simulation was carried out to restrain the GPR40 by a 1000 kJ/mol Å^2 harmonic constraint to relieve close contacts before the actual simulation. Finally, three 10 ns MD simulations were performed at the NPT canonical ensemble and the periodic boundary conditions were used in all three dimensions. Phospholipids, water molecules, receptor, and ligand were coupled separately in a temperature bath at 300 K, with a coupling constant τ_t = 0.1 ps. The pressure coupling was set as independent in the *x* and *y* directions (semiisotropic coupling), with a constant pressure of 1 bar and a coupling constant τ_p of 1 ps [29]. The particle mesh Ewald (PME) method [30] for long-range electrostatics, a 14 Å cutoff for van der Walls interactions, a 12 Å cutoff for Coulomb interaction with updates every 10 steps, and the Lincs [31] algorithm for covalent bond constraints were used.

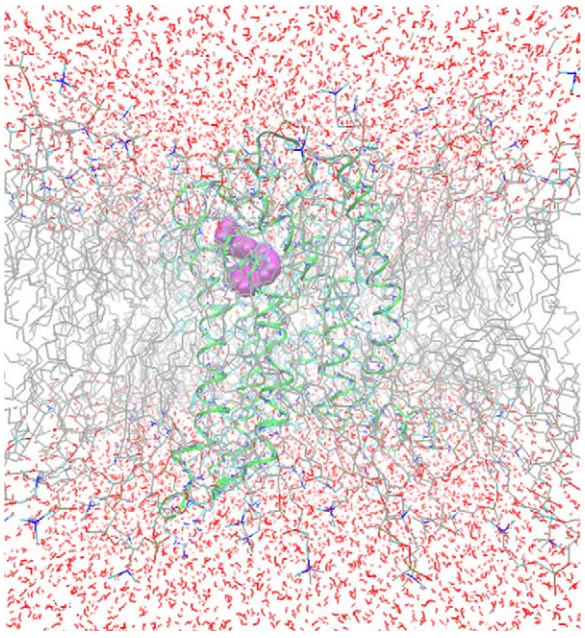

Fig. 1. Typical structure of GPR40 and GW9508 embedded in the hydrated lipid bilayer. The backbone of the receptor is represented in green, the POPC are in silver, water is in red and the agonist GW9508 is displayed as VDW in magenta.

3. Results and discussion

3.1. Molecular docking

In the many reported GPR40 agonists [10,11,13,14], para-substituted phenyl propionic acid scaffold has emerged as a common structure motif, and compounds having an aromatic ring and an acid group have shown good agonistic activity. Both characteristics can be found in CA agonists such as GW9508 and TUG424 (Fig. 2). For CA agonists, the aromatic ring and the two methylenes that connect the carboxylate head group and the phenyl provide the hydrophobic groups, while the carboxylate group provides the negative charge center. However, for non-CA agonists, the binding modes remain unknown. To study the differences between the binding modes of these agonists and to reveal the most essential amino acid residues involved in non-CAs ligand recognition, molecular docking was performed.

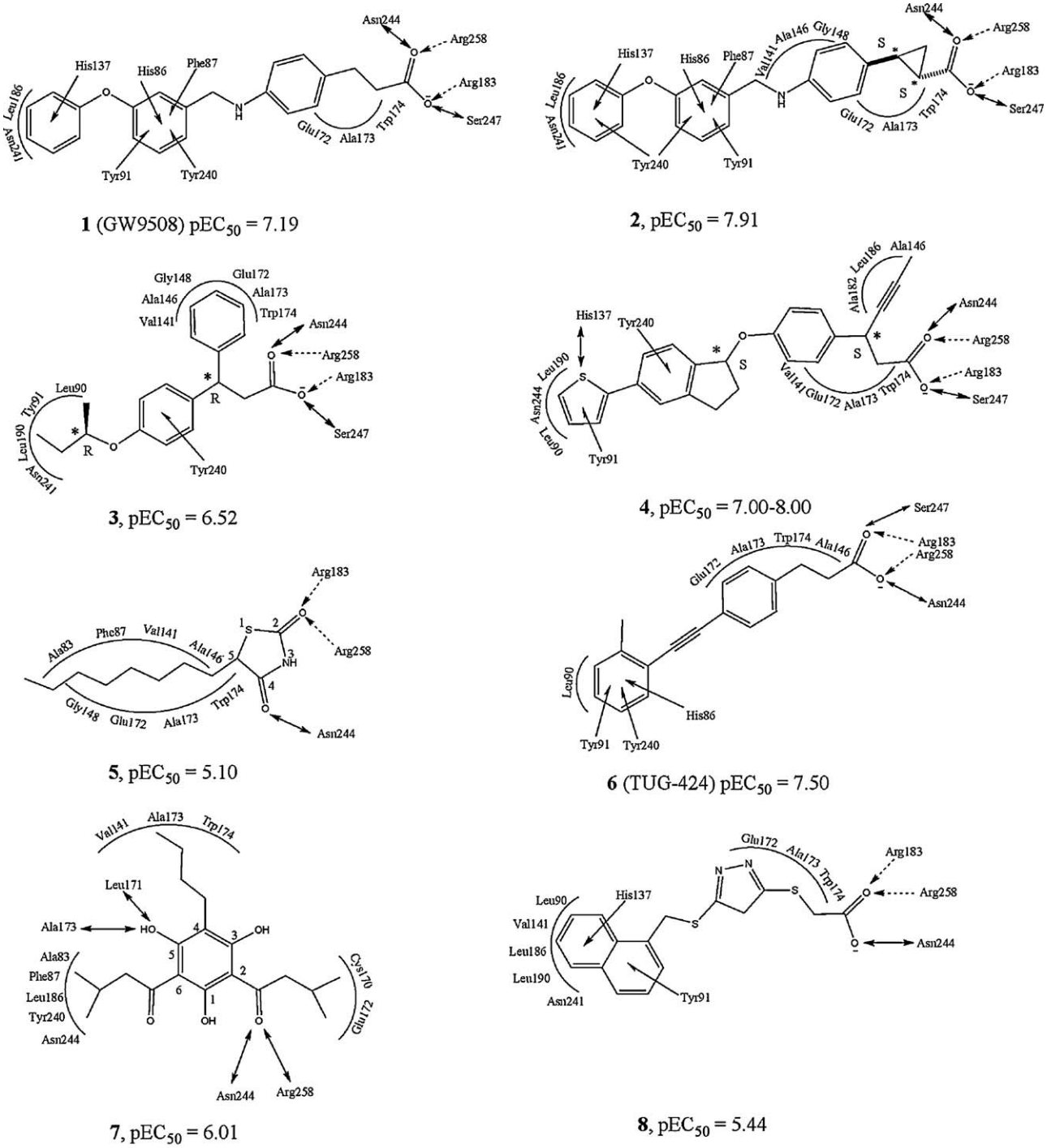

Fig. 2. Binding modes of GPR40 agonists.

Table 1
Binding and docking energies of agonists and GPR40 calculated by AutoDock.

Agonists	Binding energy (kcal/mol)	Docking energy (kcal/mol)	Inhibition constant (298.15 K), pKi	Experimental activity pEC_{50} (assay)[§]	Class[±]	Reference
1	−11.00	−13.65	8.06	7.19 (C)	i	[10]
2	−11.60	−12.97	8.50	7.91 (C)	i	[10]
3	−9.41	−11.09	6.90	6.52 (B)	ii	[12]
4	−11.55	−10.46	8.46	7.50[*] (D)	ii	[13]
5	−7.31	−9.72	5 36	5.10 (C)	iv	[17]
6	−10.00	−11.59	7.33	7.50 (A)	v	[14]
7	−9.54	−13.34	7.00	6.01 (A)	vi	[16]
8	−8.89	−9.95	6.52	5.44 (A)	vii	[15]

[*] The pEC_{50} value was mentioned as 7.00–8.00 in original reference. In this paper, we use the intermediate value 7.50 for the study.
[±] The categories were in accordance with the Bharates criteria [18].
[§] Assays and cell line used are mentioned here. A: FLIPR functional calcium-flux assay in CHO cell line; B: FDSS functional calcium-flux assay; C: Gal4 ELK1 luciferase reporter assay; D: cell-based aequorin assay.

The fifty docking conformations for each agonist were divided into groups according to a 1.0 Å RMSD criterion by using the Cluster module in ADT. Cluster conformation analysis are to compare the RMSD of the lowest energy conformations and their RMSD to one another, to group them into families of similar conformations or clusters. The reliability of the docked result depends on the similarity of its final docked conformation. The groups indicate that all the agonists mainly take one conformation. Besides RMSD cluster analysis, AutoDock also uses binding free energy evaluation to find the best binding mode. Energy items calculated by AutoDock are characterized by intermolecular energy (consist of van der Walls energy, hydrogen bonding energy, desolvation energy, and electrostatic energy), internal energy of ligand, and torsional free energy. The first two of these combined give the docking energy while the first and third terms build up the binding energy. During all these interactions, the electrostatic interaction between ligands and receptor is the most important, because in most cases it can decide the binding strength and the location of ligand, while the hydrophobic interaction of some certain groups can affect the agonistic activity to a larger extent [10,11]. The energy information is listed in Table 1, and the interaction modes of the agonists and GPR40 are depicted in Fig. 2, where only the amino acid residues located within 5 Å of the agonists are displayed.

A linear equation for the predicted binding affinities (pK_i) determined by using AutoDock and experimental activity values (pEC_{50}) was developed (Fig. 3). A linear correlation between pEC_{50} and pK_i was obtained from eight diverse agonists, yielding a good correlation coefficient ($R = 0.914$). Furthermore, it can be seen that the eight sample points are well scattered around with the fitting line, and there are no significant outliers. This equation based on the calculated pK_i can therefore be used for the prediction of the agonistic activity of other GPR40 agonists.

After docking the GW9508 back into GPR40, we superimposed the docked GW9508 and the reference GW9508 (Tikhonova's model) into the binding site of GPR40. As shown in Fig. 4, the RMSD value between the docked GW9508 and the reference GW9508 is 0.72 Å and the interactions between the docked GW9508 and GPR40 is similar to that of the original modes, suggesting that AutoDock is suitable for the docking of agonists into the binding site of GPR40. The obtained results of the molecular docking of the compound **1** (GW9508) and compound **2** are similar. Three hydrophilic residues, Arg183, Asn244, and Arg258, act as anchors for the carboxylate moiety of the two ligands, while Ser247 forms a hydrogen bond with the carboxyl group. The His137 and His86 form two cation–aromatic interactions with the terminal phenoxy ring and the benzylamino ring, respectively. A cluster of aromatic

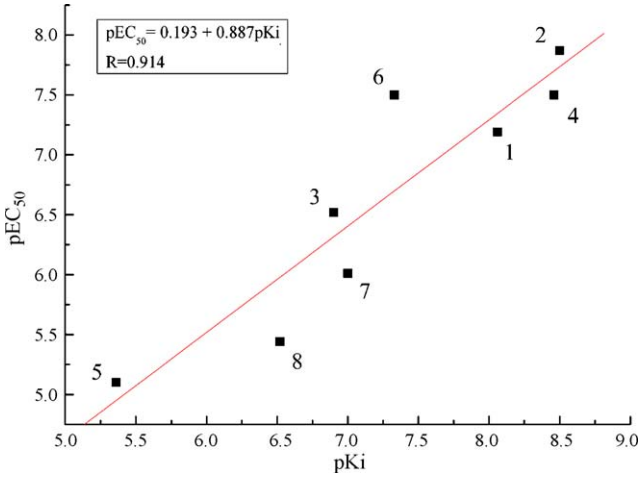

Fig. 3. Relationship between the experimental pEC_{50} and the calculated pK_i.

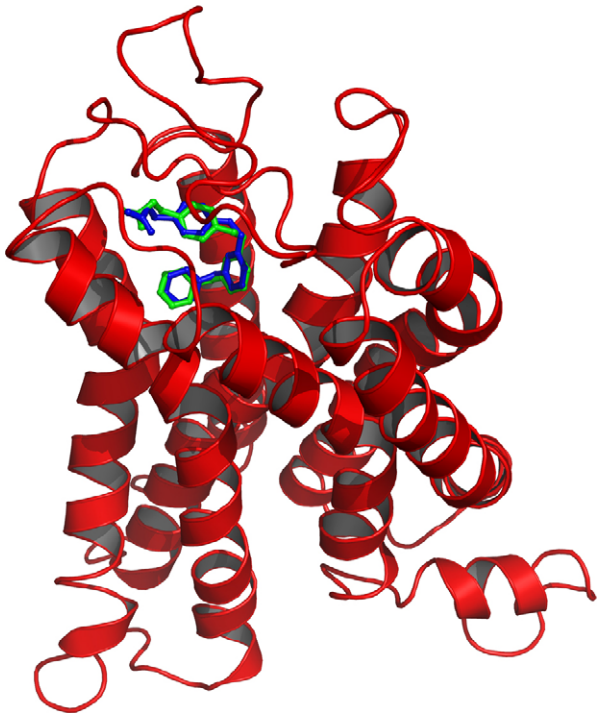

Fig. 4. The superimposition of the docked GW9508 (green) and the reference GW9508 (blue) into the homology model of GPR40. (For interpretation of the references to color in the citation of this figure, the reader is referred to the web version of the article.)

residues, Phe87, Tyr91, and Tyr240, interact with the benzylamino ring through π–π interaction. The terminal phenoxy ring is located inside the hydrophobic pocket formed by Leu186 and Asn241. The introduction of the (S,S)-cyclopropyl acid head group of compound **2** led to a significant improvement in potency over the ethyl-linked analogs [10]. The possible reason may be that the compound **2** is more rigid than GW9508, and is losing less torsional free energy compared with GW9508 during docking. These results are in good agreement with the available data on the site-directed mutagenesis obtained for GPR40 [20,32].

The molecular docking of compounds **3**, **4**, and **6** was performed in accordance with the results obtained for the GW9508. Three hydrophilic residues, Arg183, Asn244, and Arg258, act as anchors for the carboxylate moiety of ligands, while Ser247 forms a hydrogen bond with the carboxyl group. The proposed binding mode of compound **3** suggests that Tyr240 interacts with phenoxy ring through π–π interaction. The 3-phenyl group is surrounded by six hydrophobic residues. Additionally, the terminal 2-butoxy is located inside the hydrophobic pocket formed by Leu90, Tyr91, Leu190, and Asn241. The chain length of compound **3** is much shorter relative to GW9508, leading to significant decrease in the potency due to no cation–aromatic interaction formed between compound **3** and GPR40.

According to the model obtained for compound **4**, the arrangement of the carboxylate moiety of this ligand inside the binding site is also the same as for GW9508. The thienyl sulfur atom of the ligand is located at a distance of 3.25 Å from the protonated nitrogen atom of the His137 and seems to be involved in the hydrogen bonding with this residue. In addition, the thienyl group located inside the hydrophobic pocket formed by Leu90, Leu190, and Asn241, interacts with Tyr91 through π–π interaction. Ala146, Ala182, and Leu186 form a suitable pocket for the propymyl group of the ligand.

The data obtained for compound **6** indicate that the arrangement of the carboxylate moiety of the ligand is similar to that for the above-mentioned ligands. The terminal phenyl interacts with Tyr91and Tyr240 through π–π interactions, and with His86 through cation–aromatic interaction, respectively. In addition, the amino acid groups of Ala146, Glu172, Ala173, and Trp174 are involved in hydrophobic interactions with the phenylethyl of the ligand.

The molecular docking performed for compound **5** (thiazolidinedione), non-CA agonist, suggests the oxygen atom of its 4-carbonyl group is hydrogen-bonded to the Asn244 amino group. More importantly, 2-carbonyl forms two electrostatic interactions with the positively charged residues Arg183 and Arg258. Additionally, the *n*-octyl chain lies inside the hydrophobic pockets formed by the eight amino acid residues. The critical residues of the binding pocket of GPR40 for GW9508 identified by Sum et al. including His137, Arg183, Asn244, and Arg258, whereas for thiazolidinedione, the same amino acids, except His137, were also anticipated to be important [32]. The relevance of these residues to the binding and function at glitzones at GPR40 have been verified by Smith et al. [33], albeit variations in the changes of potency and efficacy of the two ligand classes in different functional end point assays, these were consistent with thiazolidinedione also binding at the orthosetric site.

The binding mode of compound **7**, non-CA agonist, with GPR40 is distinct to that for other compounds. The Asn244 is hydrogen-bonded to the 2-acyl oxygen of the ligand, while the 3-hydroxyl of the ligand is involved in the hydrogen bonding with the backbone of Leu171 and Ala173. In particular, the electrostatic interaction between the 2-acyl oxygen and the Arg258 make this diacylphloroglucionl compounds as a new class of GPR40 agonists. The diacyl groups are located inside two hydrophobic pockets formed by (i) Cys170 and Glu172 and (ii) Ala83, Phe87, Leu186, Tyr240, and Asn241, whereas the *n*-butyl chains lie around the Val141, Ala173, and Trp174.

The molecular docking performed for compound **8** suggests that the arrangement of the carboxylate moiety of the ligand is similar to that of CA agonists. The His137 and Tyr91 interact with the naphthyl of the compound **8** through cation–aromatic and π–π interactions, respectively. Leu90, Val141, Leu186, Leu190, and Asn241 form a suitable pocket for the naphthyl ring of the ligand. Substitution of the two-carbon linker with heteroatom sulfur decreased the agonist potency [11]. As a result, a two-carbon tether between the acid and phenyl units appears optimal for CA agonists.

The obtained results of the molecular docking of the small-molecule agonists allowed us to propose a general binding mode of these ligands and to determine residues involved in the ligand recognition. The obtained models demonstrate that Arg183, Asn244, and Arg258 anchor the carboxylate moiety of the CA agonists in the pocket; the His137 and His86 form cation–π with the ligands while Tyr91 and Tyr240 could be involved in the ligand binding due to π–π interactions with the ligands. For non-CA agonists, such as compounds **5** and **7**, the electronegative group is essential for interaction with the positively charged Arg183 or Arg258 through electrostatic interaction, making these non-CA agonists as novel GPR40 agonists. Most of these agonists are now in biological testing phase and a few (compound **1** and compound **3**) are at preclinical stage. The marked increase in activity of these synthetic agonists versus the endogenous fatty acid ligands (e.g., linoleic acid) will allow for a better understanding of the role of GPR40 in insulin secretion and as a potential treatment of type 2 diabetes.

According to the docking results, three complexes of GPR40 with three structural diverse agonists (GW9508, compounds **5** and **7**) were selected as representatives for MD simulations. The aim of the MD simulations was to obtain more precise ligand–receptor models in the state close to natural conditions and to further explore the binding modes of the ligands.

3.2. Molecular dynamics simulations

In order to examine conformational variations of the GPR40 within the hydrated lipid environment, the root-mean-square deviation (RMSD) of the atomic positions with respect to the three starting structures were calculated. Fig. 5 shows the RMSD for Ca atoms of the protein as a function of the simulation time (10 ns of the final molecular dynamics). The obtained RMSD values are about 2.8 Å for each complex and relatively stable after 3 ns, indicating that the molecular systems were well behaved thereafter.

The RMSD of ligands to GPR40 were obtained based on the MD simulation of three systems to get information on position fluctuations. Fig. 6a, shows that the RMSD of GW9508 atoms rose to 1.5 Å after 3 ns and then leveled off after that. This indicates that, after an initial increase in the magnitude of ligand atoms fluctuation, the ligand reached an equilibrium state characterized by the RMSD profile. From Fig. 6b and c, it can be seen that

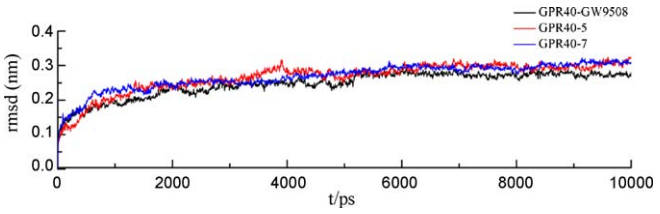

Fig. 5. The RMSD of the three studied complexes obtained during 10 ns of MD simulations.

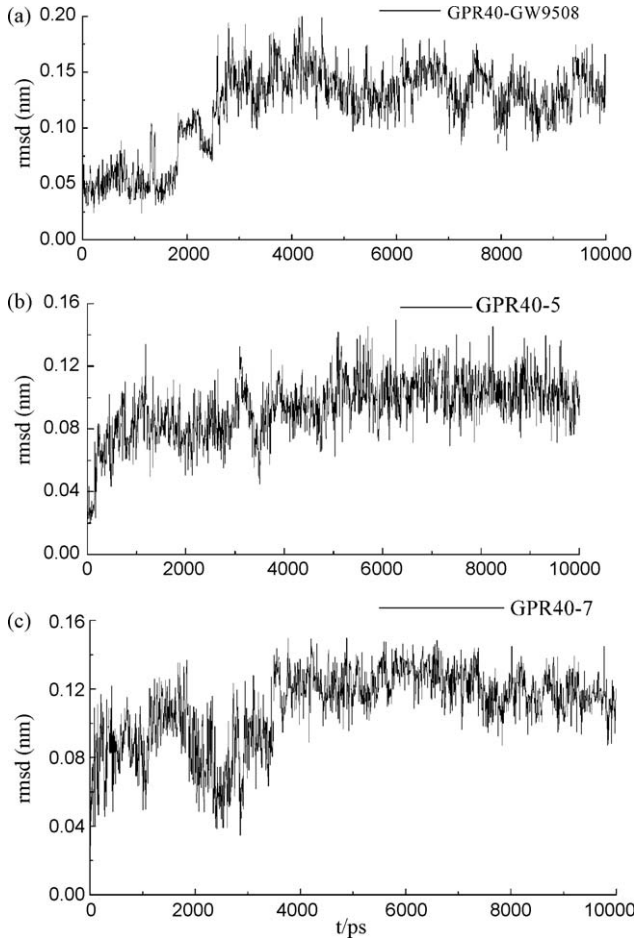

Fig. 6. RMSD of three ligands to GPR40 derived by molecular dynamics simulations. (a) GPR40–GW9508 system, (b) GPR40-**5** system, (c) GPR40-**7** system.

compounds **5** and **7** are also stable in relation to GPR40 after 3 ns judged by their RMSD values, respectively.

The analysis of the binding mode of the compound **5** obtained after MD simulation (Fig. 7), suggests that the thiazolidinedione plane rotated 45° compared with the initial docking structure, which is paralleling to the membrane plane. (The MD results of the binding mode of the ionized compound **5** are similar to that of neutral form of the compound **5**.) The original hydrogen bond between the Asn244 amino group and 4-carbonyl has vanished (4.41 Å). The 2-carbonyl forms an electrostatic interaction with the Arg183, while the 4-carboxyl forms another electrostatic interaction with the Arg258. The differences between binding modes of compound **5** after the molecular docking and after the MD simulation could be explained by the small size of the ligand and the big free space around it. Furthermore, the homology model of GPR40 is based on the inactive state of rhodopsin receptor, which may be more accurate template in the case of inverse agonists, and antagonists than in that of agonists [34]. Liapakis et al. have provided evidence for the existence of one or more intermediate conformational states linking the inactive receptor to the fully active receptor [35]. This is quite consistent with the idea that GPCRs undergo facile transitions between multiple conformational states under normal conditions [36], and that agonists preferentially bind and stabilize different receptor conformations [37,38]. Thus, it seems to be possible for compound **5** to take several reasonable binding modes. This binding mode after MD simulation was consistent with the recent experimental results that 2,4-dione function of the thiazolidinediones is coordinated to Arg183 and Arg258 in a similar manner of the free fatty acids and small-molecule agonists such as GW9508 and that this provides orientation of these ligands [33]. When an agonist interacts with Arg183 and Arg258 along with Asn244 function to coordinate the agonist at the head group, this interacting partners from receptor to the agonist allow the receptor to adopt an active conformation [39].

The results of the MD simulation obtained for compound **7** (Fig. 8) suggest that the hydrogen bonds between 3-hydroxyl of the ligand and the backbone of Leu171 and Ala173 have vanished, indicating that these hydrogen bonding interactions are weak. The distance between amino group of the Asn244 and the 2-acyl oxygen of the ligand (3.01 Å) seems to be enough for a hydrogen bond. Additionally, the positively charged residue Arg258 is involved in electrostatic interaction with the 2-acyl of the ligand. The diacyl groups and the *n*-butyl group of the compound **7** are located near in the same way as they were obtained after the molecular docking study.

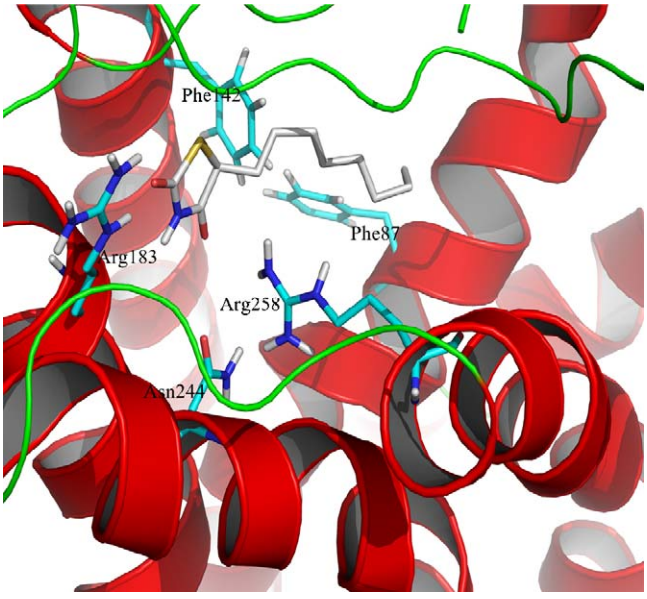

Fig. 7. Binding mode of compound **5** obtained after MD simulations.

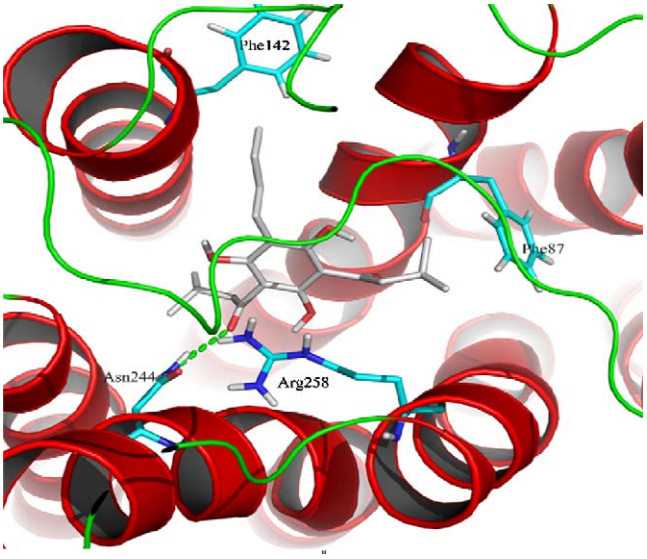

Fig. 8. Binding mode of compound **7** obtained after MD simulations.

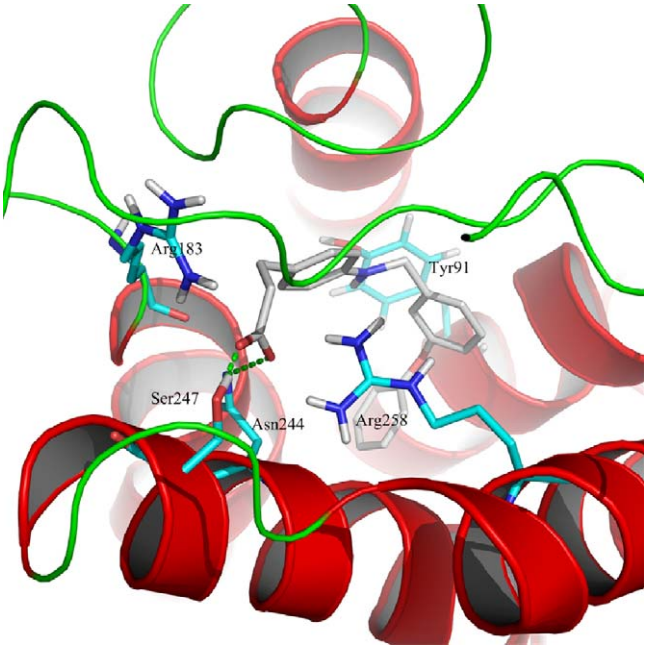

Fig. 9. Binding mode of compound **1** (GW9508) obtained after MD simulations.

The binding mode of the GW9508 established after MD simulation (Fig. 9) is nearly the same as that obtained after the molecular docking. The Arg183, Asn244, and Arg258 act as anchors for the carboxylate moiety of the ligand. The residue Ser247 forms a hydrogen bond with carboxyl oxygen of the ligand. The His137 and His86 are involved in cation–aromatic interaction with the terminal phenoxy ring and the benzylamino ring, respectively. Additionally, the phenyl rings of the Phe87, Tyr91, and Tyr240 are involved in ligand recognition through π–π interactions. The results of molecular docking and molecular dynamics simulations of GW9508 to GPR40 are summarized and compared with available experimental data of site-directed mutagenesis in Table 2.

The obtained results of molecular docking and molecular dynamics simulations of three protein–ligand systems confirmed the existence of a suitable ligand binding site located inside the transmembrane domain. For CA agonists, such as GW9508, the carboxylate moiety of the ligand directly interacts with the positively charged residues Arg183 and Arg258. On the other hand, for non-CA agonists, such as compound **5** and compound **7**, the electronegative group of these ligands (i.e., carbonyl group) is involved in the interaction with the Arg183 or Arg258. In contrast to the CA agonists, this electrostatic interaction is weaker. In general, the agonistic activity of the CA agonists is much higher than non-CA agonists.

3.3. Qualitative energy analysis

To get a detailed view of the effects of individual residues on binding affinity of the three simulated systems, a per-residue decomposition of the total energy was performed to evaluate the energetic influences of critical residues on the binding. We selected 10 conformations in the last 1 ns MD trajectories from the three systems, respectively. Energy of interaction between individual residues and ligands was calculated with the Calculate Interaction

Table 2
Comparison of molecular dynamics results with experimental data [20,32].

Residue	Mutagenesis data	
Y91[a](3.37)	Y91A: substantial reduction potency	Y91F: marginal changes potency
H86[b](3.32)	H86A: partially restored potency	H86F: marginal changes potency
H137[b](4.56)	H137A: substantial reduction potency	HI 37F: partially restored potency
R183[c](5.39)	R183A: no response	
Y240[a](6.51)	Y240A: substantial reduction potency	Y240F: marginal changes potency
N244[d](6.51)	N244A: substantial reduction potency	
S247[d](6.58)		
R258[c](7.35)	R258A: no response	R258K: substantial reduction potency

[a] π–π interactions.
[b] Cation–π interactions.
[c] Electrostatic interactions.
[d] Hydrogen bonding.

Table 3
Average energy of interaction between individual residues and ligands for the three receptor–ligand systems[a].

Ammo acid residues	Energy of interaction with compound **5**	Energy of interaction with compound **7**	Energy of interaction with GW9508
Ala83	−1.09	−3.62	−0.24
His86	−0.65		−1.3l
Phe87	−0.96	−1.33	−3.60
Tyr91		−0.36	−4.02
His137	−0.07	−0.48	−3.72
Val141	−0.20	−2.86	−1.98
Ala146	−0.78	−0.50	
Glyl48	−0.87		
Argl83	−9.80	−1.04	−18.56
Leu186	−1.54	−0.21	−4.5.
Tyr240		−2.04	−2.50
Asn244	−0.61	−4.74	−12.09
Ser247	−0.04		−12.37
Arg258	−4.22	−11.15	−30.22

[a] Energies are in kcal/mol.

Energy protocol encoded in Discovery Studio 2.1 [40] for the three receptor–ligand complexes, which is allowing to calculate the non-bonded interactions (i.e., the van der Waals term and the electrostatic term) between two sets of atoms in a specific structure or trajectory. The complexes were typed with CHARMm force field and the dielectric model was assigned to Implicit Distance-Dependent Dielectrics. All other parameters used were kept at their default settings. The residue-based decomposition of interaction energies in three systems identified several critical residues of GPR40. Table 3 lists the average energy contributions of these key residues of interest for three systems. As can be seen, major favorable energy contributions originate predominantly from Ala83, Phe87, Ala146, Gly148, Arg183, Leu186, Arg258 of GPR40-**5** system, Ala83, Phe87, Val141, Tyr240, Asn244, Arg258 of GPR40-**7** system, and His86, Phe87, Tyr91, His137, Arg183, Leu186, Tyr240, Asn244, Ser247, Arg258 of GPR40–GW9508 system. Moreover, the electrostatic interaction between ligands and the positively charged residues Arg183, Arg258 are the most favorable dominant contributions to the binding of three ligands to GPR40. Also, the energies of interaction combined contributions were greater for GW9508 than compound **5** and compound **7**. These key residues are well coincided with the molecular docking studies and MD simulations results, respectively. Thus, these critical residues may provide guidance for the rational design to discover more potent GPR40 agonists.

4. Conclusions

In the present study, the molecular docking of the eight diverse agonists (six CA agonists and two non-CA agonists) was performed, which revealed the main differences of binding modes of these agonists and the critical amino acid residues involved in the recognition of the agonists. Also, a linear equation based on the predicted binding affinities pK_i and the experimental pEC_{50} value was obtained. The good correlation coefficient indicated that our model can be used for the prediction of the agonistic activity of other GPR40 agonists.

The molecular dynamics simulations of the complexes of GPR40 with three different agonists embedded into the phospholipid bilayer were performed. The ligands were stable inside the transmembrane domain according to their RMS deviations during the simulations. Our results suggested that agonists can exactly bind to the active pocket of GPR40 to display agonistic activity. The binding mode after MD simulations may alter with the original molecular docking structure. Furthermore, the residue-based decomposition of interaction energies in three systems identified several critical residues for ligand binding. The obtained binding modes of the compound **5** and GW9508 are in good agreement with known site-directed mutagenesis data.

Acknowledgements

This work was supported by National High Technology Research and Development Program of China (863 Programs, 2007AA02Z301and 2006AA02Z336), Natural Science Foundation of China (No. 20803063).

References

[1] R.A. DeFronzo, R.C. Bonadonna, E. Ferrannini, Pathogenesis of NIDDM, Diabetes Care 15 (1992) 318–368.
[2] A.F. Amos, D.J. McCarty, P. Zimmet, The rising global burden of diabetes and its complications: estimates and projections to the year 2010, Diabet. Med. 14 (1997) S7–S85.
[3] Y. Itoh, Y. Kawamata, M. Harada, M. Kobayashi, R. Fujii, S. Fukusumi, K. Ogi, M. Hosoya, Y. Tanaka, H. Uejima, H. Tanaka, M. Maruyama, R. Satoh, S. Okubo, H. Kizawa, H. Komatsu, F. Matsumura, Y. Noguchi, T. Shinohara, S. Hinuma, Y. Fujisawa, M. Fujino, Free fatty acids regulate insulin secretion from pancreatic beta cells through GPR40, Nature 422 (2003) 173–176.
[4] C.P. Briscoe, M. Tadayyon, J.L. Andrews, W.G. Benson, J.K. Chambers, M.M. Eilert, C. Ellis, N.A. Elshourbagy, A.S. Goetz, D.T. Minnick, P.R. Murdock, H.R. Sauls Jr., U. Shabon, L.D. Spinage, J.C. Strum, P.G. Szekeres, K.B. Tan, J.M. Way, D.M. Ignar, S. Wilson, A.I. Muir, The orphan G protein-coupled receptor GPR40 is activated by medium and long chain fatty acids, J. Biol. Chem. 278 (2003) 11303–11311.
[5] K. Kotarsky, N.E. Nilsson, E. Flodgren, B. Olde, C. Owman, A human cell surface receptor activated by free fatty acids and thiazolidinedione drugs, Biochem. Biophys. Res. Commun. 301 (2003) 406–410.
[6] P. Steneberg, N. Rubins, R. Bartoov-Shifman, M.D. Walker, H. Edlund, The FFA receptor GPR40 links hyperinsulinemia, hepatic steatosis, and impaired glucose homeostasis in mouse, Cell Metab. 1 (2005) 245–258.
[7] C.P. Tan, Y. Feng, Y.P. Zhou, G.J. Eiermann, A. Petrov, C. Zhou, S. Lin, G. Salituro, P. Meinke, R. Mosley, T.E. Akiyama, M. Einstein, S. Kumar, J.P. Berger, S.G. Mills, N.A. Thornberry, L. Yang, A.D. Howard, Selective small-molecule agonists of G protein-coupled receptor 40 promote glucose-dependent insulin secretion and reduce blood glucose in mice, Diabetes 57 (2008) 2211–2219.
[8] M. Kebede, T. Alquier, M.G. Latour, M. Semache, C. Tremblay, V. Poitout, The fatty-acids receptor GPR40 plays a role in insulin secretion in vivo after high-fat feeding, Diabetes 57 (2008) 2432–2437.
[9] R. Brownlie, R.M. Mayers, J.A. Pierce, A.E. Marley, D.M. Smith, The long-chain fatty acids receptor, GPR40, and glucolipotoxicity: investigations using GPR40-knock-out mice, Biochem. Soc. Trans. 36 (2008) 950–954.
[10] D.M. Garrido, D.F. Corbett, K.A. Dwornik, A.S. Goetz, T.R. Littleton, S.C. McKeown, W.Y. Mills, T.L. Smalley, Smalley T.L.Jr., C.P. Briscoe, A.J. Peat, Synthesis and activity of small molecule GPR40 agonists, Bioorg. Med. Chem. Lett. 16 (2006) 1840–1845.
[11] S.C. McKeown, D.F. Corbett, A.S. Goetz, T.R. Littleton, E. Bigham, C.P. Briscoe, A.J. Peat, S.P. Watson, D.M.B. Hickey, Solid phase synthesis and SAR of small molecule agonists for the GPR40 receptor, Bioorg. Med. Chem. Lett. 17 (2007) 1584–1589.
[12] F. Song, S. Lu, J. Gunnet, J.Z. Xu, Y. Liang, C. Baumann, J. Lenhard, W.V. Murray, K.T. Demarest, G.H. Kuo, Synthesis and biological evaluation of 3-aryl-3-(4-phenoxy)-propionic acid as a novel series of G protein-coupled receptor 40 agonists, J. Med. Chem. 50 (2007) 2807–2817.
[13] M. Akerman, S. Brown, J.B. Houze, M.A. Zhihua, J.C. Medina, W. Qiu, M.J. Schmott, Y. Wang, L. Zhu, Conformationally constrained 3-(4-hydroxy-phenyl)-substituted-propanoic acids useful for treating metabolic disorders, Amgen Inc. WO2007033002, 2007.
[14] E. Christiansen, C. Urban, N. Merten, K. Liebscher, K.K. Karlsen, A. Hamacher, A.D. Spinrath, A.D. Bond, C. Drewke, S. Ullrich, M.U. Kassack, E. Kostenis, T. Ulven, Discovery of potent and selective agonists for the free fatty acids receptor 1 (FFA/GPR40), a potential target for the treatment of type II diabetes, J. Med. Chem. 51 (2008) 7061–7064.
[15] I.G. Tikhonova, C.S. Sum, S. Neumann, S. Engel, B.M. Raaka, S. Costanzi, M.C. Gershengorn, Discovery of novel agonists and antagonists of the free fatty acid receptor 1 (FFAR1) using virtual screening, J. Med. Chem. 51 (2008) 625–633.
[16] S.B. Bharate, A. Rodge, R.K. Joshi, J. Kaur, S. Srinivasan, S.S. Kumar, A. Kulkarni-Almeida, S. Balachandran, A. Balakrishnan, R.A. Vishwakarma, Discovery of diacylphloroglucinols as a new class of GPR40 (FFAR1) agonists, Bioorg. Med. Chem. Lett. 18 (2008) 6357–6361.
[17] C. Owman, B. Olde, D. Roeme, O. Sterner, Therapeutic modulators of GPR40, Heptahelix AB. WO2007049050, 2007.
[18] S.B. Bharate, K.V.S. Nemmani, R.A. Vishwakarma, Progress in the discovery and development of small-molecule modulators of G-protein-coupled receptor 40 (GPR40/FFA1/FFAR1): an emerging target for type 2 diabetes, Expert. Opin. Ther. Patents 19 (2009) 237–264.
[19] J. Bockaert, J.P. Pin, Molecular tinkering of G protein-coupled receptors: an evolutionary success, EMBO J. 18 (1999) 1723–1729.
[20] I.G. Tikhonova, C.S. Sum, S. Neumann, C.J. Thomas, B.M. Raaka, S. Costanzi, M.C. Gershengorn, Bidirectional, iterative approach to the structural delineation of the functional "chemoprint" in GPR40 for agonist recognition, J. Med. Chem. 50 (2007) 2981–2989.
[21] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B.G. Johnson, W. Chen, M.W. Wong, J.L. Andres, M. Head-Gordon, E.S. Replogle, J.A. Pople, Gaussian 03, Gaussian, Wallingford, CT, 2003.
[22] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, J. Comput. Chem. 19 (1998) 1639–1662.
[23] H.J.C. Berendsen, V.D. Spoel, R.V. Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Comput. Phys. Commun. 95 (1995) 43–56.
[24] E. Lindahl, B. Hess, V.D. Spoel, A package for molecular simulation and trajectory analysis, J. Mol. Mod. 7 (2001) 306–317.
[25] S.J. Marrink, O. Berger, P. Tieleman, F. Jahnig, Adhesion forces of lipids in a phospholipid membrane studied by molecular dynamics simulations, Biophys. J. 74 (1998) 931–943.

[26] D.P. Tieleman, H.J. Berendsen, M.S. Sansom, An alamethicin channel in a lipid bilayer: molecular dynamics simulations, Biophys. J. 76 (1999) 1757–1769.
[27] D.P. Tieleman, S.J. Marrink, H.J.C. Berendsen, A computer perspective of membrane: molecular dynamics studies of lipid bilayers systems, Biochim. Biophys. Acta 1331 (1997) 235–270.
[28] A.W. Schüttelkopf, D.M.F. van Aalten, PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, Acta Cryst. D60 (2004) 1355–1363.
[29] H.J.C. Berendsen, J.P.M. Postma, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J Chem. Phys. 81 (1984) 3684–3690.
[30] T. Darden, D. York, L. Pedersen, Particle mesh Ewald-an Nlog(N) method for Ewald sums in large system, J Chem. Phys. 98 (1993) 10089–10092.
[31] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraajie, LINCS: a linear constraint solver molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[32] C.S. Sum, I.G. Tikhonova, S. Neumann, S. Engel, B.M. Raaka, S. Costanzi, M.C. Gershengorn, Identification of residues important for agonist recognition and activation in GPR40, J. Biol. Chem. 282 (2007) 29248–29255.
[33] N.J. Smith, L.A. Stoddart, N.M. Devine, L. Jenkins, G. Milligan, The action and mode of binding of thiazolidinedione ligands at free fatty acid receptor 1, J. Biol. Chem. 284 (2009) 17527–17539.
[34] K.E. Furse, T.P. Lybrand, Three-Dimensional models for ß-adrenergic receptor complexes with agonists and antagonists, J. Med. Chem. 46 (2003) 4450–4462.
[35] G. Liapakis, W.C. Chan, M. Papadokostaki, J.A. Javitch, Synergistic contributions of the functional groups of epinephrine to its affinity and efficacy at the ß2 adrenergic receptor, Mol. Pharmacol. 65 (2004) 1181–1190.
[36] T. Kenakin, Agonist-specific receptor conformations, Trends Pharm. Sci. 18 (1997) 416–417.
[37] G. Peleg, P. Ghanouni, B.K. Kobilka, R.N. Zare, Single-molecule spectroscopy of the beta(2) adrenergic receptor: observation of conformational substates in a membrane protein, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 8469–8474.
[38] Z. Salamon, V.J. Hrubv, G. Tollin, S. Cowell, Binding of agonists, antagonists, and inverse agonists to the human delta-opioid receptor produces distinctly different conformational states distinguishable by Plasmon-waveguide resonance spectroscopy, J. Pept. Res. 60 (2002) 322–328.
[39] C.S. Sum, I.G. Tikhonova, S. Costanzi, M.C. Gershengorn, Two arginine-glutamate ionic locks near the extracellular surface of FFAR1 gate receptor activation, J. Biol. Chem. 284 (2009) 3529–3536.
[40] Accelrys, Inc. Discovery Studio 2.1, 2008, 10188 Telesis Court, Suite 100 San Diego, CA 92121-4779, U.S.A.