Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry 15 (2007) 3054–3064
Bioorganic & Medicinal Chemistry

Modeling of human ghrelin receptor (hGHS-R1a) in its close state and validation by molecular docking

Alessandro Pedretti and Giulio Vistoli*

Istituto di Chimica Farmaceutica e Tossicologica ''Pietro Pratesi'', Facoltà di Farmacia, Università degli Studi di Milano, Viale Abruzzi 42, I-20131 Milano, Italy

Received 4 August 2006; revised 10 January 2007; accepted 31 January 2007
Available online 2 February 2007

Abstract—The objective of this study was to generate a reliable model of human ghrelin receptor (hGHS-R1a) in its close state by means of a hybrid fragmental approach in which the transmembrane bundle was modeled using the rhodopsin as the template to assure a marked closeness among the transmembrane helices, while the remaining segments (i.e., loops plus terminal domains) were modeled searching different templates to favor the local homologies. The reliability of this model was assessed docking both a tetrapeptide, which represents the ghrelin's active core, and a set of 50 peptidomimetic secretagogues taken from the literature. The analysis of obtained complexes unveils a set of stabilizing interactions with crucial hGHS-R1a residues in remarkable agreement with both mutational analyses and pharmacophore hypotheses. Also the significant correlation between docking scores and biological activities affords an encouraging validation for such hGHS-R1a model, suggesting that also the receptor in its close state (similarly to the hGHS-R1a in its open state which was modeled in our previous study, Pedretti A, Villa M, Pallavicini M, Valoti E, Vistoli G. *J. Med. Chem.* **2006**, *49*, p 3077.) may be involved in ligand binding and could find fertile applications in ligand design.

1. Introduction

The recently discovered hormone, ghrelin,[1] plays a key role in the growth hormone (GH) secretion and energy homeostasis due to its orexigenic and adipogenic effects.[2,3] Ghrelin is the first natural peptide having the hydroxy group of a serine residue (Ser3) acylated by *n*-octanoic acid and this posttranslational modification is essential for its bioactivity.[4] Ghrelin is produced in several tissues having both endocrine and paracrine effects,[5] and its concentration is influenced by acute and chronic changes in nutritional state.[6,7] Ghrelin affects a number of different systems mainly including GH, adrenocorticotropic hormone (ACTH), and prolactin release, feeding, gastric secretion and mobility, metabolism, cardiac performances, and cell proliferation.[8] Therefore, peptidomimetic GH secretagogues (GHSs)[9] find therapeutic applications in several pathological conditions, namely GH deficiency, negative energy balance status, obesity, diabetes mellitus and insulin resistance, Prader–Willi syndrome, cachexia, hypertension, and hyperthyroidism.[10,11]

The human ghrelin receptor (hGHS-R1a) belongs to a small family of receptors for peptide hormones and neuropeptides within class I of GPCRs.[12] One of the most important features of hGHS-R1a is its constitutive activity since it is able to change into an active conformation without the agonist, signaling with about 50% of maximal activity.[13] High constitutive activity has been demonstrated for many GPCRs, even if it is very difficult to verify whether such ligand-independent receptor signaling is of physiological relevance.[14] Conversely, the biological role of constitutive activity in hGHS-R1a was recently confirmed through the identification of a naturally occurring mutation, which selectively eliminates the constitutive activity without affecting the affinity, potency, and efficacy of the endogenous ghrelin.[15] The occurrence of such a mutation results in carriers with very short stature and obesity.

Mutagenesis studies on hGHS-R1a revealed the major role of Glu 124[16] in the third transmembrane helix (TM3), which presumably forms a salt bridge with ligand ammonium head. Other key residues are present in TM2 (Glu 99), TM3 (Gln120 and Ser 123), and TM5 (Met213).[17]

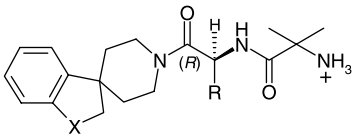
Keywords: Ghrelin; Human grelin receptor hGHS-R1a; Growth hormone secretagogues; GPCR; Homology modeling; Molecular docking.
* Corresponding author. Tel.: +39 02 50317545; fax: +39 02 50317565; e-mail: giulio.vistoli@unimi.it

0968-0896/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2007.01.057

Mutational analyses evidenced also that the constitutive activity is mainly due to an aromatic cluster formed by residues in TM5 (e.g., Phe220, Phe222, and Phe226) and TM6 (e.g., Phe279, Arg283, and Phe286) that, even without ligand, approach the inner face of TM2 and TM3, shifting into a constitutively active form.[18] In the light of these findings, the binding cavity of hGHS-R1a can be considered as formed by two pockets: a first polar cavity, lined by TM2, TM3, and a second aromatic cavity, lined by TM5 and TM6. Moreover, it is possible to consider two hGHS-R1a states: an open state, in which the pockets form two distinct binding sites, and a close state, in which the aromatic cluster approaches the polar pocket, and Glu124 (TM3) interacts with Arg283 (TM6).[13] Interestingly, mutagenesis experiments on Arg283 indicated that its mutation eliminates both agonist stimulation and constitutive activity, suggesting the implication of such a residue in the receptor activation and in the interaction with the agonists. This is in agreement with biophysical studies revealing that activation of many GPCRs involves an inward movement of TM6 and TM7 toward TM3.[19]

In our previous study,[20] we modeled the human ghrelin receptor in its open state and validated it by docking a heterogeneous set of ligands taken from the literature. The model was generated exploiting a fully fragmental approach, which (1) divides the sequence in structural fragments, (2) predicts their folding separately, and (3) assembles the obtained segments using the bovine rhodopsin structure (PDB Id: 1F88) as the template. Such a method allowed us to obtain a hGHS-R1a model whose transmembrane bundle was markedly wider than that of bovine rhodopsin even assuring an overall agreement with the experimental template. The results showed that in the open state the polar binding site is clearly involved in ligand recognition, but also the aromatic pocket can bind some of the considered ligands suggesting a cooperative effect of two binding sites.

Although in the hGHS-R1a close state the polar binding site is partially occupied by aromatic cluster, it remains to be clarified whether the close state is also involved in ligand binding and whether the analysis of such interactions can be useful to improve the prediction of ligand activity. Therefore, we undertook to generate a reliable model of ghrelin receptor in its close state using a hybrid approach, in which the folding of transmembrane segments was predicted using the rhodopsin structure as the template in order to assure a substantial closeness among the transmembrane helices, while the extramembrane segments (i.e., loops plus terminal domains) were separately predicted searching different templates to favor the local homologies (as later explained under Section 4). Similar to the previous study,[20] the goodness of the close model was verified by docking first the tetrapeptide, Gly-Ser-Ser(*n*-octanoyl)-Phe-NH2 (EC_{50} = 72 nM), which constitutes the 'active core' required for agonist potency at hGHS-R1a,[21] and then an enriched dataset which includes the same set of 35 GH secretagogues collected for the previous study (Charts 1–5) plus a new group of 15 GH secretagogues taken from the literature (Charts 6–9) in order to further validate the obtained hGHS-R1a close state model.[20]



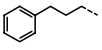
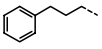
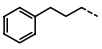
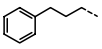
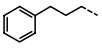
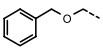
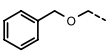
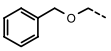
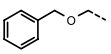
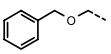
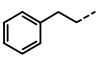
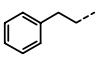
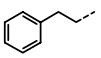
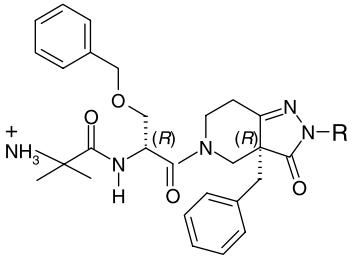
Cpd	R	X	EC_{50} (nM)
1		NSO_2CH_3	1.3
2		(*R*)CHOH	0.6
3		CH_2	14.0
4		S	5.6
5		SO_2	2.1
6		CH_2	10.0
7		S	4.8
8		SO	2.0
9		SO_2	1.1
10		O	2.7
11		CH2	17.0
12		S	10.0
13		SO	8.5
14		SO2	5.0
15		O	8.5
16		CH2	22.0
17		S	57.0
18		SO2	6.0

Chart 1. Spiroindane derivatives.

2. Results and discussion

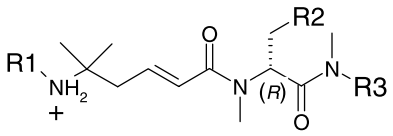
2.1. Analysis of GHS-R1a model in its close state

Figure 1 depicts the tube structure of the obtained hGHS-R1a model colored by segment, showing the typical folding of GPCRs with seven transmembrane



Cpd	R	EC_{50} (nM)
19	CH_3	3.0
20	CH_2CH_3	4.0
21	H	3.0

Chart 2. Pyrazolinone–piperidine derivatives.



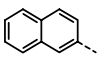
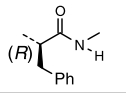
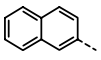
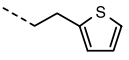
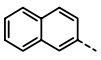
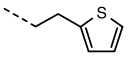
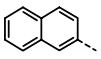
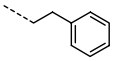
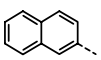
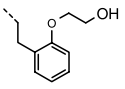
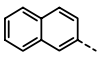
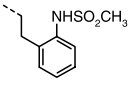
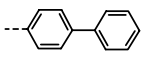
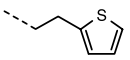
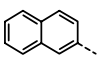
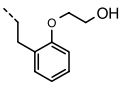
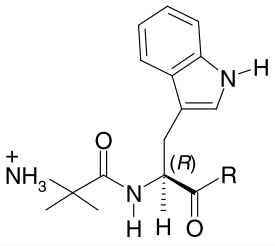
Cpd	R1	R2	R3	EC_{50} (nM)
22	H			3.0
23	H			7.5
24	CH_3			0.8
25	CH_3			3.0
26	H			5.0
27	H			18.0
28	H			5.0
29	CH_3			5.0

Chart 3. α,β-Unsaturated derivatives.



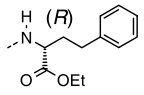
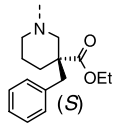
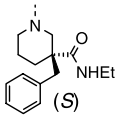
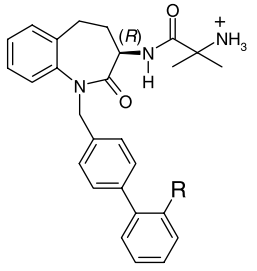
Cpd	R	EC_{50} (nM)
30		3.0
31		1.6
32		0.8

Chart 4. Indol derivatives.



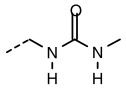
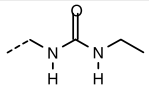
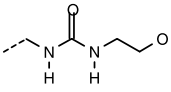
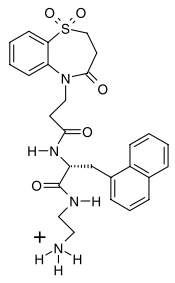
Cpd	R	EC_{50} (nM)
33		1.0
34		6.0
35		2.0

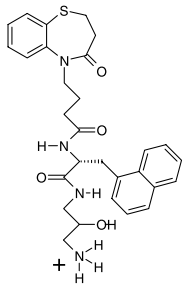
Chart 5. Benzolactam derivatives.

helices (TM1-7, the precise subdivision in segments is described in Table 3). The N-terminal domain (NT) mainly assumes a β-hairpin structure, involving residues between Met1 and Asp32. It is stabilized by a network of H bonds plus an electrostatic interaction between Asp32 and ammonium terminal group, which closes the hairpin motif.

The main differences between close and open states involve the transmembrane bundle. Indeed, while in the open state it shows a calyx shape with the extracellular side wider than the cytoplasmic side,[20] in the close state the TM segment assumes a cylindrical crown shape, in which extracellular and cytoplasmic sides have nearly the same wideness. The inter-helix distances, as compiled in Table 1, confirm that the transmembrane helices are organized with an anticlockwise connectivity, in

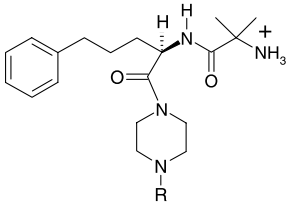


36 (EC_{50} = 10.0 nM)



37 (EC_{50} = 1.0 nM)

Chart 6. Benzothiazine derivatives.



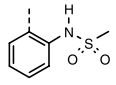
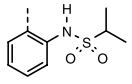
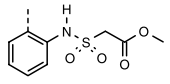
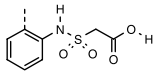
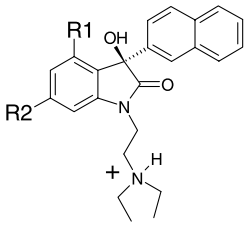
Cpd	R	EC_{50} (nM)
38		100.0
39		6.3
40		2.8
41		4.5
42		2.3

Chart 7. Phenylpiperazine derivatives.



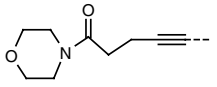
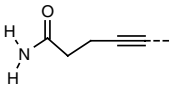
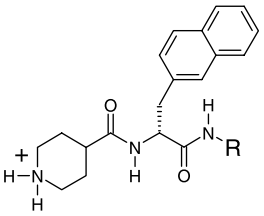
Cpd	R1	R2	EC_{50} (nM)
43	Br	H	77.0
44	CF_3		30.0
45	CF_3		30.0
46	CF_3	$CONH_2$	18.0

Chart 8. Oxindole derivatives.



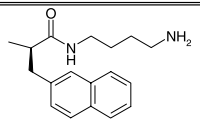
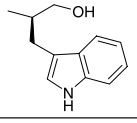
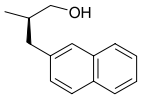
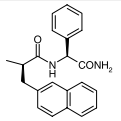
Cpd	R	EC_{50} (nM)
47		17.0
48		10.0
49		1.8
50		0.28

Chart 9. Naphthyl piperidine derivatives.

which each segment is closest to its numerically adjacent helices. The mean values show that TM3, which bears key residues in ligand interaction, takes the most central position in transmembrane segment, while TM5 takes the most peripheral one. The inter-helix mean distances also suggest that TM2, TM3, and TM7 compose the core of transmembrane segment, while TM1 and TM5 are in a more external position.

To put these results into perspective, Table 1 also reports the distance values as computed for the open state.[20] As a rule, the inter-helix distances calculated for the close state are shorter than those for the open state (excepting TM4) and the most relevant differences involve the distances between TM3 and TM5, TM3 and TM6, and TM5 and TM6. It confirms that the close state is characterized by an approach of TM5 and TM6 toward TM3. Conversely, TM4 shows a mean value greater in close state than in open state probably

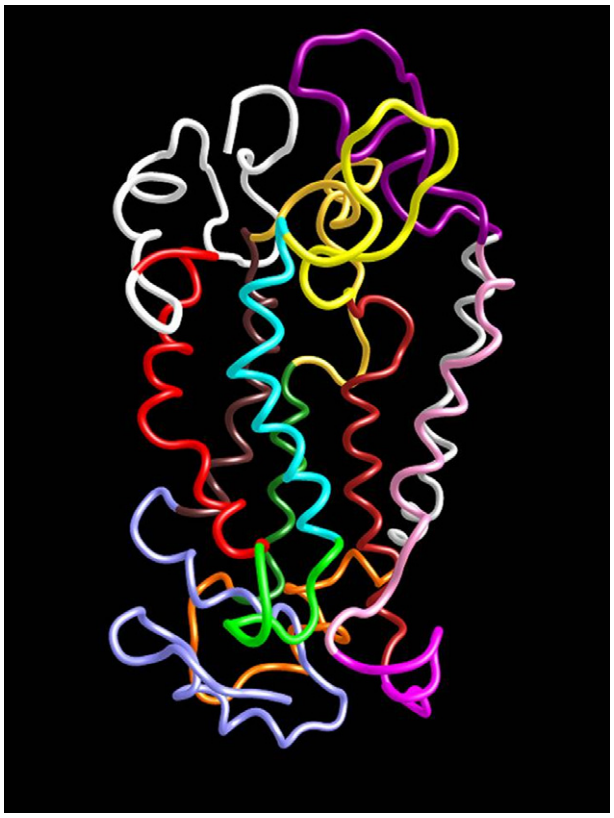


Figure 1. Tube structure of GHS-R1a model colored by segment. NT, white; TM1, red; CL1, green; TM2, azure; EL1, yellow; TM3, dark red; CL2, violet; TM4, pink; EL2, indigo; TM5, grey; CL3, orange; TM6, dark green; EL3, dark yellow; TM7, brown; and CT, blue.

Table 1. Matrix of inter-helix distances (the values, expressed in Å, are computed as described in Ref. 20)

helix	TM1	TM2	TM3	TM4	TM5	TM6	Mean
TM1							17.14
							17.52
TM2	8.87						13.10
	8.82						13.22
TM3	15.19	8.22					11.26
	16.03	8.25					12.48
TM4	22.92	14.13	10.78				15.35
	22.34	13.66	9.25				14.43
TM5	28.32	21.43	10.86	14.12			18.22
	28.95	22.05	14.04	14.19			19.37
TM6	17.60	15.39	11.05	17.69	13.85		14.16
	18.71	16.15	14.54	18.47	16.15		17.46
TM7	9.94	10.52	11.51	18.48	20.78	9.39	13.39
	10.27	10.44	12.78	18.53	20.76	9.28	13.68

In each cell, the first value concerns the close state and the second value the open state.

because the inward movement of TM5 and TM6 pushes to one side TM4 which goes away from other TM helices (especially from TM2 and TM3). The transmembrane segment is mainly stabilized by apolar contacts involving aromatic and aliphatic residues. The polar interactions between transmembrane helices are markedly less frequent, but they play pivotal roles to stabilize the transmembrane folding.

The crucial differences between open and close states concern the polar interactions which drive the approaching among TM3, TM5, and TM6. In particular, the contacts realized by Arg283 (TM6) with polar residues of TM3 (namely, Ser123, Glu124, Ser125, and Thr127) are a key factor in stabilizing the close state. Furthermore the aromatic residues of TM5 and TM6, which constitute the 'aromatic subpocket' in open state, also approach TM3 surrounding Glu124 and lining the binding site of close state.

The extracellular loops (EL1, EL2, and EL3) assume hairpin geometries vastly stabilized by polar contacts. The main difference between open and close states involves the EL2 segment, which in the open state partially penetrates the transmembrane segment, forming the upper side of polar pocket,[20] while in the close state it assumes a more folded conformation and its residues move away from transmembrane segment. The conformation of cytoplasmic loops (CL1, CL2, and CL3) in the close state is very similar to that of the open state assuming constrained turn geometries mainly stabilized by polar interactions.

Finally, the C-terminal domain (CT) assumes a helix–turn–helix motif, beginning with an eighth helix (Ser327–Pro342) perpendicular to seven helix and rich in positively charged residues (e.g., Lys328, Lys329, Arg331, and Arg336) which anchor the CT domain to phospholipidic heads.[22]

2.2. Docking results on GHS-R1a model: the tetrapeptide

Figure 2 shows the complex obtained between the ghrelin 'active core' [Gly-Ser-Ser(*n*-octanoyl)-Phe-NH2] and the binding site of hGHS-R1a in its close state. The docking results allow to reveal the main interactions stabilizing the complex. As confirmed by mutagenesis, the

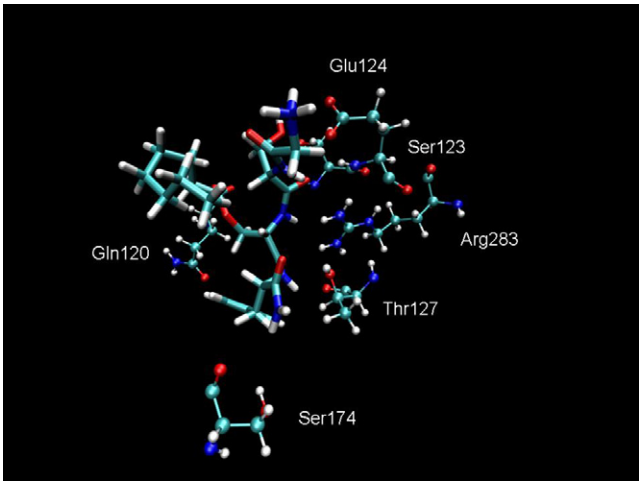


Figure 2. Representation of the complex realized by tetrapeptide [Gly-Ser-Ser(*n*-octanoyl)-Phe-NH_2] with hGHS-R1a. The main interactions involve: (1) the ligand ammonium head with Asp124, (2) unacylated serine with Ser123, (3) peptidic bonds with Arg283, and (4) terminal amide with Thr127. Notice that the apolar residues are not displayed for clarity.

terminal ammonium group plays a crucial role forming a salt bridge with Glu124 (TM3). Other polar interactions involve: (1) the tetrapeptide unacylated serine with Ser123 (TM3), (2) the ligand peptide bonds with Arg283 (TM6), (3) the ester function with Gln120 (TM3), and (4) the terminal amide with Thr127 (TM3) and Ser174 (TM4). The phenyl ring of tetrapeptide is involved in π–π interactions with Phe87 (TM2) and Phe119 (TM3), while the octanoyl chain realizes hydrophobic contacts with Leu181, Val182, Ile187 (TM4), Val216, and Ile219 (TM5). Gratifyingly, the binding role of many reported residues (e.g., Gln120, Ser123, Glu124, Val216, and Arg283) finds experimental proofs from mutational analyses.[23]

Taken globally, these results afford a preliminary validation of the hGHS-R1a close model unveiling that its binding site is clearly narrower than the corresponding polar pocket in the open state,[20] even showing a similar symmetric architecture which defines three regions with alternating polarity (one polar region in the middle and two lateral hydrophobic zones). In particular, the binding site of close state shows an amphiphilic structure with one side lined by polar residues (belonging to TM3 plus Arg283) and an opposite side lined by hydrophobic residues of TM4 and TM5. Moreover, the hydrophobic side can be divided in two portions since it shows a more superficial region richer in aliphatic residues which contacts the octanoyl chain and a deeper portion more abundant in aromatic residues which accommodates the phenyl ring of tetrapeptide.

2.3. Docking results on GHS-R1a model: the ligand dataset

The docking results, obtained by tetrapeptide, find encouraging verifications when analyzing the complexes of peptidomimetic GHS ligands in dataset. Figure 3 shows the main interactions realized by the most active spiroindane derivative (**2**, EC_{50} = 0.6 nM, Chart 1).

In particular, the ligand's ammonium head realizes ionic interactions with Glu124 reinforced by H bonds with Ser123. The amidic functions are involved in H bonds with Gln120, Ser125, and Arg283. The indole moiety mimics the phenyl ring of tetrapeptide interacting with aromatic residues (e.g., Phe87, Phe119, Tyr128, and Phe172), while the spiroindane moiety mimics the octanoyl chain contacting both aliphatic (e.g., Leu114, Leu181, Val182, and Ile187) and aromatic (Phe121 and Phe179) residues. Due to the clear symmetry of the binding site, the spiroindane derivatives can bind the receptor assuming two possible binding modes, which preserve the same central polar interactions, but distinguish for the apolar contacts. In particular, in the first mode, as reported in Figure 3, the spiroindane moiety mimics the octanoyl chain, while in the second symmetric mode the spiroindane interacts in lieu of the indole (i.e. with Phe87, Phe119, Tyr128, and Phe172). When analyzing all computed poses for spiroindane derivatives, both binding modes are found in all ligands even if in most cases the best score poses involve the binding scheme as reported in Figure 3. A comparative

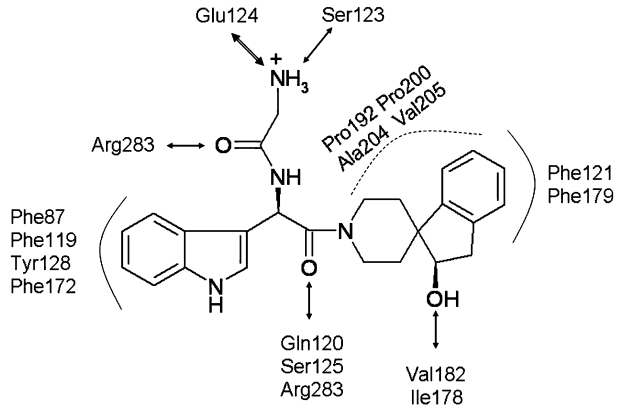


Figure 3. Bi-dimensional scheme showing the main interactions realized by **2** (EC_{50} = 0.6 nM) with GHS-R1a binding site. Dashed line, hydrophobic contacts; line, charge transfer, and π–π interactions; bold arrow, ionic interactions; arrow, H bond interactions.

analysis of these results suggests that the aromatic character of two ligand's hydrophobic portions drives the choice of the preferred binding mode. As an example, the aromatic character of indole in **2** is so prevalent to render the binding mode of Figure 3 markedly favored, while when the indole is replaced by arylalkyl moieties (as seen in ligands **6-18**), the modest difference between the two hydrophobic portions makes the binding modes equally likely.

The ability to form strong H bonds with Ser174 (TM4) is another key factor which rules the binding mode of considered ligands. A clear example is offered by indole derivative **32** (EC_{50} = 0.8 nM, Chart 4) which can be roughly considered as an open derivative of **2**. The pattern of involved residues, as depicted by Figure 4, is similar to that of **2** but the indole derivative preferentially assumes a binding mode which is symmetric in respect to that of **2** due to the strong H-bonds which the ethylamide function realizes with both Ser174 and Thr127. The indole moiety interacts with Phe121 and Phe179, while the phenyl ring contacts a set of aromatic residues (i.e., Phe87, Phe119, Tyr128, and Phe172).

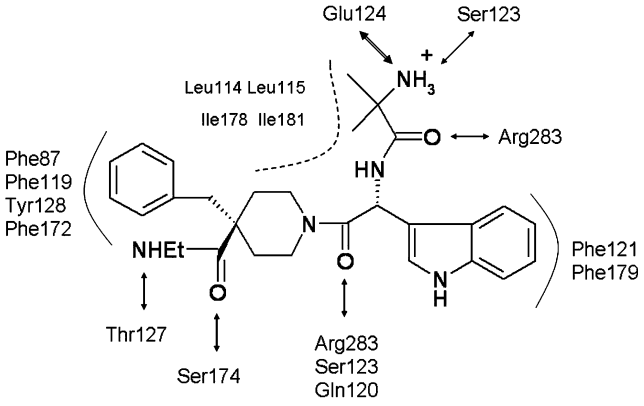


Figure 4. Bi-dimensional scheme showing the main interactions realized by **32** (EC_{50} = 0.8 nM) with GHS-R1a binding site. Dashed line, hydrophobic contacts; line, charge transfer and π–π interactions; bold arrow, ionic interactions; arrow, H bond interactions.

Interestingly, this ligand does not possess aliphatic groups, which can extensively mimic the octanoyl chain, even if some apolar contacts are realized by methyl groups vicinal to ammonium head and by piperidine ring. This seems to suggest that for both hydrophobic substructures (namely, the moieties which mimic the octanoyl chain and the tetrapeptide phenyl ring) the interactions with aromatic residues are more relevant than those with aliphatic ones, as indeed confirmed by the fact that aromatic groups, such as naphthyl, are preferable to aliphatic ones in short ghrelin peptidic derivatives.[4]

The beneficial role of aromatic moieties is further confirmed by the interaction pattern obtained with the most active compound (the naphthylpiridine derivative **50**, EC_{50} = 0.28 nM, Chart 9) as depicted in Figure 5. Indeed, such derivative allows to define three aromatic subpockets able to accomodate as many aromatic groups. Indeed, the napthyl moiety closer to the piridine ring interacts with both aromatic (i.e., Phe121 and Phe179) and aliphatic residues (i.e., Ile178, Leu181, and Val182), the farer napthyl moiety realizes π–π stacking with Phe119 and Phe172, and the phenyl ring contacts Phe87 and Phe119, suggesting that the last two aromatic subpockets are somewhat overlapped. The ammonium head realizes a set of polar interactions with Ser123, Glu124, Ser125, and Ser207. Again, the amidic functions are involved in H bonds with Gln120, Ser174, and Arg283.

Finally, a quite different interaction pattern is presented by the benzolactam derivatives, as seen in Figure 6 for compound **33** (EC_{50} = 1.0 nM, Chart 5). Indeed, in these derivatives there is not any aromatic system able to interact with Phe121 and Phe 179 since the benzolactam moiety is too close to ammonium head and it can only form charge transfer interactions with Arg283. The biphenyl group realizes π–π stacking with Phe87, Phe119, Tyr128, and Phe172, and the urea function

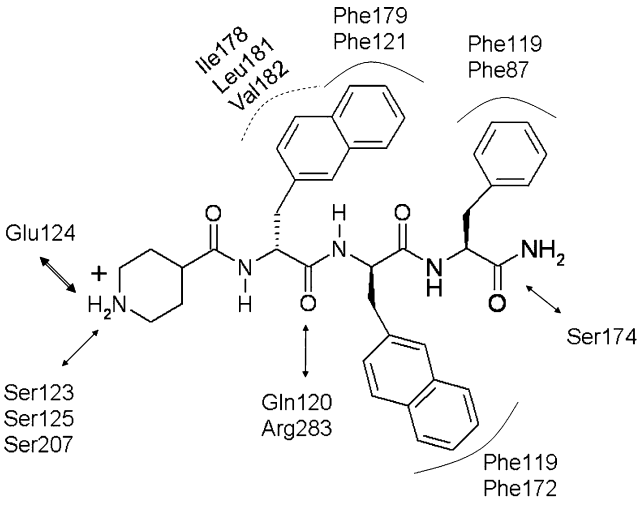


Figure 5. Bi-dimensional scheme showing the main interactions realized by **50** (EC_{50} = 0.28 nM) with GHS-R1a binding site. Dashed line, hydrophobic contacts; line, charge transfer and π–π interactions; bold arrow, ionic interactions; arrow, H bond interactions.

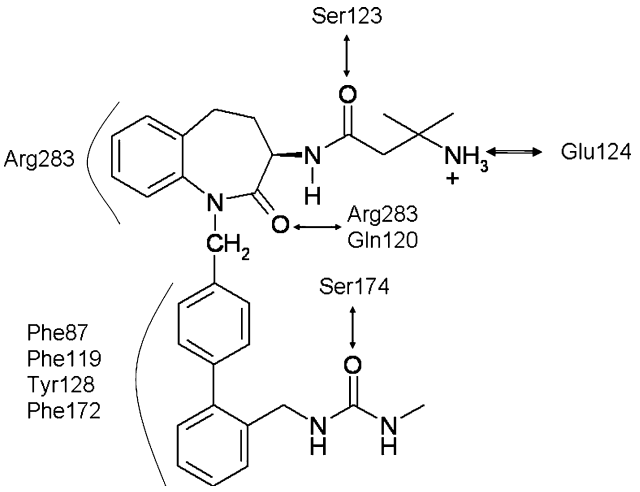


Figure 6. Bi-dimensional scheme showing the main interactions realized by **33** (EC_{50} = 1.0 nM) with GHS-R1a binding site. Dashed line, hydrophobic contacts; line, charge transfer and π–π interactions; bold arrow, ionic interactions; arrow, H bond interactions.

forms H bond with Ser174. Interestingly, the benzolactam derivatives showed an atypical binding mode also in the open state.[20]

The docking results on GHS peptidomimetics unveil a significant homogeneity among the involved residues and stabilizing interactions. Indeed, all considered ligands realize an ion-pair with Glu124 which seems to be mandatory for activity even if these interactions may be somewhat weakened by the closeness of Arg283, which in turn forms H bonds with more amidic functions. These interactions are reinforced by H bonds involving Gln120, Ser123, Ser125, Thr127, and Ser174.

The polar interactions act as a pivot around which the ligand's hydrophobic moieties engage apolar contacts with both aliphatic and aromatic residues. Even recognizing an upper region of the binding site richer in aromatic residues and a deeper region richer in aliphatic residues, the marked narrowness of binding pocket merges the two hydrophobic regions making their interaction capabilities so similar to justify the possibility of two symmetric binding modes.

2.4. Docking score predictivity

First, the 15 compounds added in dataset (namely, **36**–**50**) were docked also in the open state model[24] obtaining results in line with those described in previuos study (complexes not shown). Table 2 reports the best docking scores obtained by hGHS-R1a model in its close state and also the best scores obtained with hGHS-R1a model in its open state for polar subpocket.[25]

The plot in Figure 7 shows the remarkable correlation between docking scores as obtained in close model and biological activities (expressed as pEC_{50}). Eq. 1 confirms the remarkable predictive power of docking scores obtained by close state model as proven by significant statistical parameters (e.g., r^2 = 0.69).

Table 2. Best docking scores (in kcal/mol) obtained with close model and with open model for all ligands in dataset

Ligand	Activity pEC_{50}	Score close	Score open
1	8.87	−35.64	−39.01
2	9.22	−40.42	−36.71
3	7.85	−26.68	−25.4
4	8.25	−30.91	−30.08
5	8.68	−36.48	−33.46
6	8	−27.52	−26.67
7	8.32	−30.80	−28.72
8	8.6	−34.83	−32.96
9	8.96	−38.63	−39.13
10	8.57	−36.40	−37.48
11	7.78	−29.58	−29.74
12	8	−31.28	−32.55
13	8.07	−28.21	−33.49
14	8.3	−36.16	−36.74
15	8.07	−30.19	−33.81
16	7.66	−28.38	−25.33
17	7.24	−26.81	−22.83
18	8.22	−34.47	−33.34
19	8.52	−30.99	−31.3
20	8.4	−34.55	−33.6
21	8.52	−37.28	−36.77
22	8.52	−29.48	−37.01
23	8.12	−33.39	−31.8
24	9.1	−35.35	−36.4
25	8.52	−37.84	−33.54
26	8.3	−36.02	−29.02
27	7.74	−23.98	−24.38
28	8.3	−28.30	−30.66
29	8.3	−24.04	−28.01
30	8.52	−32.55	−35.57
31	8.8	−34.62	−32.32
32	9.1	−40.63	−35.95
33	9	−35.00	−39.11
34	8.22	−32.62	−36.21
35	8.7	−41.24	−44.79
36	8.22	−32.62	−36.21
37	8.70	−41.24	−44.79
38	7.85	−19.87	−19.04
39	8.20	−34.28	−35.28
40	8.55	−33.25	−40.01
41	8.35	−27.25	−37.78
42	8.64	−42.57	−40.63
43	7.11	−20.49	−15.02
44	7.52	−18.62	−19.60
45	7.52	−19.71	−21.99
46	7.74	−20.69	−24.47
47	7.77	−28.30	−27.59
48	8.00	−23.53	−29.73
49	8.74	−39.27	−40.82
50	9.55	−49.88	−38.00

$$pEC_{50} = -0.069(\pm 0.01)\ \text{Score close} + 6.11(\pm 0.22) \quad (1)$$
$$n = 50, r^2 = 0.69, q^2 = 0.65, s = 0.31, F = 108.17$$

Interestingly, the best scores obtained with open model afford a very similar and good correlation (as seen in Eq. 2) with bioactivity values, confirming the role of the polar subpocket in open state and also the soundness of such open model.

$$pEC_{50} = -0.072(\pm 0.01)\ \text{Score open} + 5.97(\pm 0.23) \quad (2)$$
$$n = 50, r^2 = 0.71, q^2 = 0.68, s = 0.30, F = 115.60$$

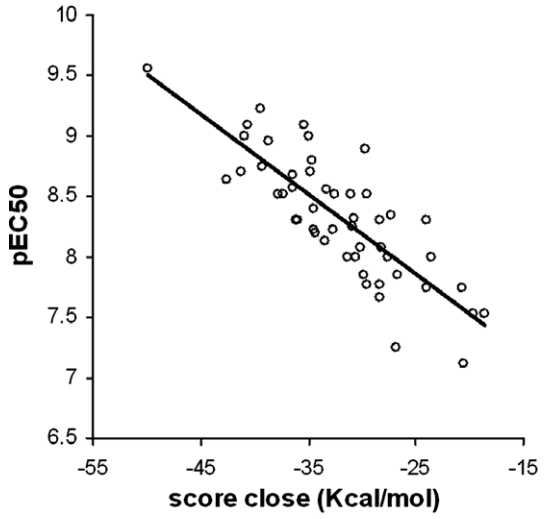


Figure 7. Correlation between biological activities (expressed as pEC_{50}) and docking scores as obtained in binding site of hGHS-R1a close model.

Since the cross-correlation between the scores obtained in close state and those computed for polar pocket in open state is modest ($r^2 = 0.55$), a third equation was calculated inserting both docking scores. The obtained Eq. 3 shows an interesting improvement of the statistical parameters ($r^2 = 0.78$ vs. 0.69) suggesting that a good ligand would have to successfully interact with both hGHS-R1a states.

$$pEC_{50} = -0.037(\pm 0.01)\ \text{Score\_close} - 0.042\ (\pm 0.01)\ \text{Score\_open} + 5.77\ (\pm 0.20) \quad (3)$$
$$n = 50, r^2 = 0.78, q^2 = 0.76, s = 0.26, F = 84.75$$

Although the statistical parameters could be fortuitously overvalued by some approximations of docking simulations (i.e., rigidity of protein structure, overestimation of polar contacts, and suitability of score function), the obtained results seem to afford an encouraging validation of both hGHS-R1a homology models confirming the reliability of fragmental prediction approaches to derive improved full-length models of GPCRs.

3. Conclusions

The homogeneity of docking complexes, the agreement with mutational analyses, and the relevant correlations between docking scores and biological activities confirm the reliability of the here proposed hGHS-R1a model. What is more, these results seem to suggest that also the close state is involved in ligand binding and the docking analyses with such a model can find fertile applications to predict and rationalize the ligand activity.

Since the docking calculations are static simulations, it is impossible to establish whether the ligand can recognize only the open state and the interactions with polar pocket in open state and with close state represent two steps of the same binding process or whether a ligand can

independently recognize both the open and close states. However, the obtained results suggest that a ligand can successfully interact with both states (as also proven by significant statistical parameters of Eq. 3) and the docking results on close state can be used to further confirm those obtained with open state (or vice versa).

Finally, the reliability of the described hGHS-R1a model confirms the soundness of the fragmental prediction which can be exploited by such a hybrid approach to explore the local homologies only in selected segments (i.e., the loops and the terminal domains) modulating the compliance with the experimental template in order to generate different states for the same protein.

4. Computational details

4.1. Construction of GHS-R1a starting model

The primary sequence of the human GHS-R1a receptor was retrieved from Swiss-Prot database[26] (entry code Q92847, GHSR_HUMAN). As mentioned in Introduction, hGHS-R1a starting model was generated using a hybrid fragmental strategy that involves: (1) the fragmentation of amino acidic sequence in 15 segments (namely, 7 TM segments, six loops, and two terminal segments, as reported in Table 3), (2) the modeling of the transmembrane segments using constantly the rhodopsin structure as the template (PDB Id: 1f88), (3) the separate modeling of other segments (namely, loops plus terminal domains) searching different templates to favor local homologies (as already done in the previous study[20]), and (4) the assembly of obtained fragments using rhodopsin structure as the template.

In particular, the segments were predicted using Fugue approach, an on-line 3D structure prediction software, which generates realistic models recognizing distant homologies by sequence–structure comparison.[27] The subdivision of amino acidic sequence was carried out using TMPRED,[26] which defines the length and position of the seven transmembrane segments, as reported in Table 3. For transmembrane helices Fugue was constrained to use the rhodopsin structure as the template, the use of which is clearly justified by vast homology percentage between bovine rhodopsin and hGHSR-1a (i.e., 58.4%, as computed using pairwise alignment). Conversely, the extramembrane fragments were predicted letting the software free to choose the best template from its protein database. For each segment Fugue was able to generate several realistic models and the best structure has been chosen considering what result better fulfilled the following major conditions: (a) the predicted secondary structure from the sequence alignment; (b) the lack of not predicted gaps; (c) the prediction score (*ZSCORE*) calculated by Fugue program; (d) the helix conformation of seven transmembrane segments with characteristic slight bend of helices containing proline and glycine residues; (e) the global 'U' shape for the loops in which the two ends are close enough to join to adjacent TM segments, (f) the disulfide bridge between Cys126 (TM3) and Cys304 (EL2) that tethers the structure of almost all GPCRs,[28] and (g) the presence of an eighth helix in the starting portion of the CT segment, placed at straight angle in respect of TM7 helix.[22]

Table 3 reports the templates used by Fugue to generate the best model for each fragment: apart from transmembrane helices which were forcedly modeled using rhodopsin, only one fragment (i.e., LC1) was modeled using the rhodopsin structure as the template, while other segments were built using different templates. It confirms the usefulness to explore local homologies to generate more reliable models.

Finally, the assembly of predicted fragments was performed superimposing the backbone of a fragment with that of the correspondent segment in rhodopsin structure and manually connecting the adjacent segments using VEGA software.[29] In particular, the superimposition involved the Cα atoms of transmembrane helices only, as the loop arrangements are clearly defined by

Table 3. Definition of hGHS-R1a segments and description of templates used in fragmental prediction (in parentheses, the resolutions of template crystal structures are reported)

Segment	Position	Length	Template PDB code	Template description
N–T	Met1–Leu42	42	1cee (NMR)	Cdc42 with the GTPase Binding Domain of Wasp
TM1	Leu43–Val68	26	1f88 (2.8 Å)	Rhodopsin*
CL1	Ser69–Asn79	11	1f88	Rhodopsin
TM2	Leu80–Val101	22	1f88	Rhodopsin*
EL1	Arg102–Lys117	16	1gzt (1.3 Å)	Pseudomonas aeruginosa lectin II with fucose
TM3	Leu118–Val139	22	1f88	Rhodopsin*
CL2	Glu140–Arg159	20	1tgg (2.0 Å)	Rh3 designed right-handed coiled coil trimer
TM4	Val160–Val182	23	1f88	Rhodopsin*
EL2	Gly183–Thr211	29	1i25 (NMR)	Huwentoxin-II
TM5	Val212–Ile235	24	1f88	Rhodopsin*
CL3	Gly236–Lys263	28	1kd0 (1.9 Å)	Methylaspartate ammonia-lyase
TM6	Met264–Phe286	23	1f88	Rhodopsin*
EL3	Ser287–Asn305	19	1ej6 (3.6 Å)	Reovirus core
TM7	Leu306–Met326	21	1f88	Rhodopsin*
C-T	Ser327–Thr366	40	1i26 (NMR)	Ptu-1, toxin from the assassin bugs *Peirates turpis*

The '*' indicates the segments for which the rhodopine was forcedly used as the template.

the position of TMs, and their conformation was further relaxed by successive MD simulations (while the transmembrane bundle remains constrained during the molecular dynamics, as later described in next section). The soundness of assembled receptor in its close state was assessed by checking the distance between TM3 and TM6, and in particular monitoring the distance between Asp124 and Arg283 but also the position of aromatic residues which surround Arg283 was carefully verified.

Rotamer libraries were applied to insert side chains and hydrogen atoms were added using VEGA. According physiological pH, Arg, Lys, Glu, and Asp residues were preserved ionized, while His residues were considered neutral by default. After a careful visual scrutiny of obtained structure to avoid unphysical conditions, the hGHS-R1a model underwent an initial minimization until RMS = 1 to discard high-energy interactions, followed by a local minimization until RMS = 0.05, where all atoms were kept fixed except for atoms included within a 7.5 Å sphere around the manually connected bonds (at the fragment ends). Finally, the model was optimized by a final minimization made up by two phases: first a minimization without constraints until RMS = 0.1 and then a second minimization with backbone fixed until RMS = 0.01 to preserve the predicted structure. In these phases and in the following steps the model goodness was assessed using Procheck[30] and Verify3D.[31]

4.2. Model equilibration

In order to gain a better relaxation and a more correct arrangement of the whole hGHR-1a model, a molecular dynamics equilibration was performed in vacuo. The simulations were carried out in 3 phases: (1) heating from 0 to 300 K over 3000 iterations (3 ps, i.e. 1 K/10 iterations), (2) starting equilibration of 2500 ps, where the transmembrane backbone was kept fixed, and (3) equilibration of 7500 ps, in which the transmembrane backbone was harmonically restrained with decreasing harmonic force constants. More in detail, harmonic force constant value was equal to 1 (1000 kJ mol^{-1} nm^{-2}) at the beginning of simulation and then was divided into two every 1.5 ns (then 5 MD were performed with harmonic force constant equal to 1, 0.5, 0.25, 0.12, and 0.06). Globally, the MD simulation lasted 10 ns and helices were correctly preserved also with harmonic force constant equal to 0.06. The last frame was used for the docking calculations after a final minimization until RMS = 0.01 (with harmonic force constant equal to 0.06).

The MD simulations had the following general characteristics: constant temperature at 300 ± 10 K by means of Langevin's algorithm; Lennard-Jones (L-J) interactions were calculated with a cut-off of 10 Å and the pair list was updated every 20 iterations; Newton's equation was integrated, using r-RESPA method, every 4 fs for long-range electrostatic forces, 2 fs for short-range nonbonded forces, and 1 fs for bonded forces; a frame was stored every 5 ps, yielding 2000 frames. All calculations were carried out in a dual Athlon PC. The package Namd2.51[32] was used with the force-field CHARMm v22 and Gasteiger's atomic charges. All minimizations in model construction and equilibration were performed using the conjugated gradients algorithm.

4.3. GHS dataset

A dataset with 50 heterogeneous peptidomimetic GH secretagogues was compiled from the literature, considering GHR-1a agonists for which the biological activity was evaluated through in vitro assay of GH release from rat pituitary cells.[33] This kind of biological data was mainly chosen since a vast majority of pharmacological data reported in the literature for GHSs are obtained using this assay that seems the unique well-standardized method, yielding fully comparable results.

For each considered series, the most active compounds only were included in dataset (i.e., with $EC_{50} \leqslant 100$ nM). According to their structure, the considered GHSs can be classified in nine classes: spiroindanes[34] (**1**–**18** including MK-0677, Chart 1), pyrazolinone-piperidines[35] (**19**–**21** including capromorelin, Chart 2), α,β-unsaturated derivatives[36] (**22**–**29**, Chart 3), indoles[37] (**30**–**32** Chart 4), benzolactams[38] (**33**–**35**, Chart 5), benzothiazepines[39] (**36**–**37**, Chart 6), phenylpiperazines[40] (**38**–**42**, Chart 7), oxindoles[41] (**43**–**46**, Chart 8), and naphthylpiperidines[42] (**47**–**50**, Chart 9). The first five charts include the ligands already used in the previuos study[20] (**1**–**35**), while the last four charts report the new compounds (**36**–**50**) here included to improve the model validation. The biological activities within the dataset ranged from 0.28 nM (**50**) to 100 nM (**38**). As suggested by pharmacophore hypothesis,[39] all selected GHSs possess an ionizable amine group, at least one H bond acceptor function, and at least two aromatic moieties.

The GHSs were built preserving the stereochemistry reported in the literature (as indicated in Charts 1–9), and when more stereoisomers are described in the literature only the most active one was considered in docking analyses. The compounds were simulated in their protonated form as it is involved in receptor recognition. After a preliminary energy minimization to discard high-energy intramolecular interactions, the overall geometry and the atomic charges were optimized using *MOPAC6.0* (keywords: 'AM1', 'PRECISE', 'GEOOK', 'MMOK'). The tetrapeptide Gly-Ser-Ser(*n*-octanoyl)-Phe-NH2 was built using VEGA and underwent the same minimization procedure as peptidomimetic GHSs.

4.4. Docking analyses

The FlexX program was used to dock the compounds to the binding site of GHS-R1a in its close state. FlexX is a fast-automated docking program that considers ligand conformational flexibility by an incremental fragment placing technique.[43] In this study, the docking analyses involved the full-length hGHS-R1a model considering all residues enclosed within 15.0 Å radius sphere

centered on Arg283 (TM6) so that the ligands can interact with the binding site generated by the approach of TM6 and TM7 toward TM2 and TM3. For each molecule, 30 docking solutions (poses) were computed and scored.

References and notes

1. Kojima, M.; Hosoda, H.; Date, Y.; Nakazato, M.; Matsuo, H.; Kangawa, K. *Nature* **1999**, *402*, 656.
2. Korbonits, M.; Goldstone, A. P.; Gueorguiev, M.; Grossman, A. B. *Front. Neuroendocrinol.* **2004**, *25*, 27.
3. Ghigo, E.; Broglio, F.; Arvat, E.; Maccario, M.; Papotti, M.; Cuccioli, G. *Clin. Endocrinol.* **2005**, *62*, 1.
4. Matsumoto, M.; Hosoda, H.; Morizumi, N.; Hayashi, Y. *Biochem. Biophys. Res. Commun.* **2001**, *287*, 142.
5. Kojima, M.; Kangawa, K. *Physiol. Rev.* **2005**, *85*, 495.
6. Otto, B.; Spranger, J.; Benoit, S. C.; Clegg, D. J.; Tschop, M. H. *Br. J. Nutr.* **2005**, *93*, 765.
7. Ueno, H.; Yamaguchi, H.; Kangawa, K.; Nakazato, M. *Regulatory Peptides* **2005**, *126*, 11.
8. van der Lely, A. J.; Tschop, M.; Heiman, M. L.; Ghigo, E. *Endocr. Rev.* **2004**, *25*, 426.
9. Smith, R. G. *Endocr. Rev.* **2005**, *26*, 346.
10. Cummings, D. E.; Foster-Schubert, K. E.; Overduin, J. *Curr. Drug Targets* **2005**, *6*, 153.
11. Smith, R. G.; Sun, Y.; Betancourt, L.; Asnicar, M. *Best Pract. Res. Clin. Endocrinol. Metab.* **2004**, *18*, 333.
12. Smith, R. G.; Leonard, R.; Bailey, A. R. T.; Palyha, O.; Feighner, S.; Tan, C.; McKee, K. K.; Pong, S.-S.; Griffin, P.; Howard, A. *Endocrine* **2001**, *14*, 9.
13. Holst, B.; Cygankiewicz, A.; Jensen, T. H.; Ankersen, M.; Schwartz, T. W. *Mol. Endocrinol.* **2003**, *17*, 2201.
14. Seifert, R.; Wenzel-Seifert, K. *Naunyn Schmiedebergs Arch. Pharmacol.* **2002**, *366*, 381.
15. Pantel, J.; Legendre, M.; Cabrol, S.; Hilal, L.; Hajaji, Y.; Morisset, S.; Nivot, S.; Vie-Luton, M. P.; Grouselle, D.; de Kerdanet, M.; Kadiri, A.; Epelbaum, J.; Le Bouc, Y.; Amselem, S. *J. Clin. Invest.* **2006**, *116*, 760.
16. In this paper the residues are progressively numbered from 1 to 366, irrespectively from the corresponding segment.
17. Feighner, S. D.; Howard, A. D.; Prendergast, K.; Palyha, O. C.; Hreniuk, D. L.; Nargund, R.; Underwood, D.; Tata, J. R.; Dean, D. C.; Tan, C. P.; McKee, K. K.; Woods, J. W.; Patchett, A. A.; Smith, R. G.; Van der Ploeg, L. H. T. *Mol. Endocrinol.* **1998**, *12*, 137.
18. Holst, B.; Holliday, N. D.; Bach, A.; Elling, C. E.; Cox, H. M.; Schwartz, T. W. *J. Biol. Chem.* **2004**, *279*, 53806.
19. Elling, C. E.; Thirstrup, K.; Holst, B.; Schwartz, T. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 12322.
20. Pedretti, A.; Villa, M.; Pallavicini, M.; Valoti, E.; Vistoli, G. *J. Med. Chem.* **2006**, *49*, 3077.
21. Bednarek, M. A.; Feighner, S. D.; Pong, S. S.; McKee, K. K.; Hreniuk, D. L.; Silva, M. V.; Warren, V. A.; Howard, A. D.; Van Der Ploeg, L. H.; Heck, J. V. *J. Med. Chem.* **2000**, *43*, 4370.
22. Sakmar, T. P.; Menon, S. T.; Marin, E. P.; Awad, E. S. *Annu. Rev. Biophys. Biomol. Struct.* **2002**, *31*, 443.
23. Holst, B.; Lang, M.; Brandt, E.; Bach, A.; Howard, A.; Frimurer, T. M.; Beck-Sickinger, A.; Schwartz, T. W. *Mol. Pharmacol.* **2006**, Jun 23; [Epub ahead of print].
24. The new derivatives were docked in open state using the same computational procedure adopted in previous study, focusing the analyses in a 20.0-Å radius sphere centered on Ile178.
25. The role of apolar subpocket in open state of GHS-R1a was here ignored for cleanness.
26. <http://www.ch.embnet.org/software/TMPRED_form.html>.
27. Shi, J.; Blundell, T. L.; Mizuguchi, K. *J. Mol. Biol.* **2001**, *310*, 243.
28. Karnik, S. S.; Gogonea, C.; Patil, S.; Saad, Y.; Takezako, T. *Trends Endocrinol. Metab.* **2003**, *14*, 431.
29. Pedretti, A.; Villa, L.; Vistoli, G. *J. Mol. Graph.* **2002**, *21*, 47.
30. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *26*, 283.
31. Bowie, J. U.; Luthy, R.; Eisenberg, D. *Science* **1991**, *253*, 164.
32. Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comp. Phys.* **1999**, *151*, 283.
33. Heiman, M. L.; Nekola, M. V.; Murphy, W. A.; Lance, V. A.; Coy, D. H. *Endocrinology* **1985**, *116*, 410.
34. Chen, M. H.; Pollard, P. P.; Patchett, A. A.; Cheng, K.; Wei, L.; Chan, W. W.; Butler, B.; Jacks, T. M.; Smith, R. G. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1261.
35. Carpino, P. A.; Lefker, B. A.; Toler, S. M.; Pan, L. C.; Hadcock, J. R.; Cook, E. R.; DiBrino, J. N.; Campeta, A. M.; DeNinno, S. L.; Chidsey-Frink, K. L.; Hada, W. A.; Inthavongsay, J.; Mangano, F. M.; Mullins, M. A.; Nickerson, D. F.; Ng, O.; Pirie, C. M.; Ragan, J. A.; Rose, C. R.; Tess, D. A.; Wright, A. S.; Yu, L.; Zawistoski, M. P.; DaSilva-Jardine, P. A.; Wilson, T. C.; Thompson, D. D. *Bioorg. Med. Chem.* **2003**, *11*, 581.
36. Peschke, B.; Ankersen, M.; Hansen, T. K.; Hansen, B. S.; Lau, J.; Nielsen, K. K.; Raun, K. *Eur. J. Med. Chem.* **2000**, *35*, 599.
37. Yang, L.; Morriello, G.; Patchett, A. A.; Leung, K.; Jacks, T.; Cheng, K.; Schleim, K. D.; Feeney, W.; Chan, W. W.; Chiu, S. H.; Smith, R. G. *J. Med. Chem.* **1998**, *41*, 2439.
38. DeVita, R. J.; Bochis, R.; Frontier, A. J.; Kotliar, A.; Fisher, M. H.; Schoen, W. R.; Wyvratt, M. J.; Cheng, K.; Chan, W. W.; Butler, B.; Jacks, T. M.; Hickey, G. J.; Schleim, K. D.; Leung, K.; Chen, Z.; Chiu, S. L.; Feeney, W. P.; Cunningham, P. K.; Smith, R. G. *J. Med. Chem.* **1998**, *41*, 1716.
39. Huang, P.; Loew, G. H.; Funamizu, H.; Mimura, M.; Ishiyama, N.; Hayashida, M.; Okuno, T.; Shimada, O.; Okuyama, A.; Ikegami, S.; Nakano, J.; Inoguchi, K. *J. Med. Chem.* **2001**, *44*, 4082.
40. Barakat, K. J.; Cheng, K.; Chan, W.-S.; Butler, B.; Jacks, T. M.; Schleim, K. D.; Hora, D. F., Jr.; Hickey, G. J.; Smith, R. G.; Patchett, A. A.; Nurgund, R. P. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1431.
41. Tokunaga, T.; Hume, W. E.; Umezome, T.; Okazaki, K.; Ueki, Y.; Kumagai, K.; Hourai, S.; Nagamine, J.; Seki, H.; Taiji, M.; Noguchi, H.; Nagata, R. *J. Med. Chem.* **2001**, *44*, 4641.
42. McDowell, R. S.; Elias, K. A.; Stanley, M. S.; Burdick, D. J.; Burnier, J. P.; Chan, K. S.; Fairbrother, W. J.; Hammonds, R. G.; Ingle, G. S.; Jacobsen, N. E.; Mortensen, D. L.; Rawson, T. E.; Won, W. B.; Clark, R. G.; Somers, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 11165.
43. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, *261*, 470.