An Exhaustive Exploration of the Semaglutide-GLP-1R Sequence Space towards the Design of Semaglutide Analogues with Elevated Binding Affinity to GLP-1R

Wei Li *

Posted Date: 6 May 2024

doi: 10.20944/preprints202405.0258.v1

Keywords: semaglutide; GLP-1R; semaglutide-GLP-1R sequence space; binding affinity

Article

An Exhaustive Exploration of the Semaglutide-GLP-1R Sequence Space towards the Design of Semaglutide Analogues with Elevated Binding Affinity to GLP-1R

Wei Li

Contrebola Institute of Computational Interstructural Biophysics, No. 88, Renaissance East Road, Nantong City 226000, Jiangsu Province, People's Republic of China; wli148@aucklanduni.ac.nz

Abstract: Semaglutide is a potent GLP-1 receptor agonist used in the treatment of type 2 diabetes mellitus due to its ability to regulate blood glucose levels and promote weight loss. On July 26, 2021, with a manually defined set of computational structural and biophysical analysis, a simple Val27-Arg28 exchange was for the first time introduced in the backbone of semaglutide to strengthen the semaglutide-GLP-1R binding affinity. In this article, a comprehensive structural and biophysical analysis approach is for the first time proposed towards an exhaustive exploration of the semaglutide-GLP-1R sequence space for the design of semaglutide analogues with elevated binding affinity to GLP-1R, thereby potentially enhancing therapeutic efficacy of structurally conceivable semaglutide analogues. Through structure biophysics-based rational design and computational modeling, this article puts forward a set of semaglutide analogues and calculated their binding affinities to GLP-1R, with one particular semaglutide analogue-GLP-1R structural model reaching a K_d of 3.0×10^{-8} M, while the K_d is 3.4×10^{-6} M for the binding of native semaglutide to GLP-1. Overall, the computationally designed semaglutide analogues here constitute a hopeful approach for developing GLP-1 receptor agonists with improved efficacy for the treatment of diabetes and weight management in future.

Keywords: semaglutide; GLP-1R; semaglutide-GLP-1R sequence space; binding affinity;

1. Introduction

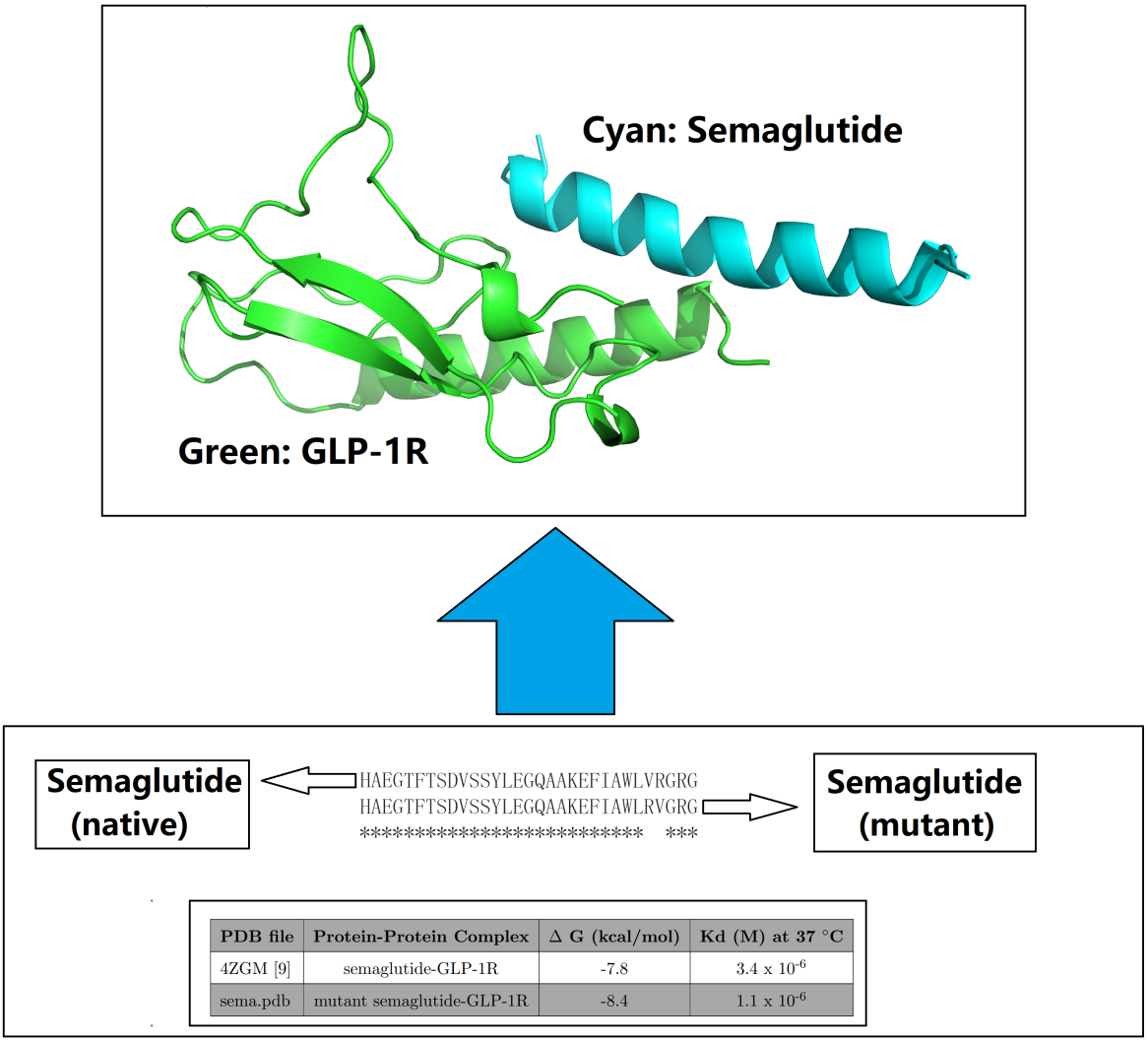
Semaglutide is a synthetic glucagon-like peptide-1 (GLP-1) receptor agonist that has garnered significant attention in the field of diabetes management [1–3]. Structurally, it is a synthetic peptide consisting of 39 amino acids. Semaglutide shares 94% sequence homology with natural human GLP-1 [2–9]. At the molecular level, semaglutide binds and activates the GLP-1 receptor, promoting insulin secretion and inhibiting glucagon release from pancreatic beta and alpha cells, respectively [10–12].

Originally developed by Novo Nordisk, semaglutide was approved by regulatory agencies for the treatment of type 2 diabetes mellitus (T2DM) due to its potent glucose-lowering effects and additional benefits such as weight loss and cardiovascular risk reduction [13–16]. Semaglutide is available in both injectable and oral formulations, with the injectable form typically administered once weekly and the oral form taken once daily [17–19]. The therapeutic efficacy of semaglutide arises from its ability to activate GLP-1 receptors located on pancreatic beta cells, leading to increased insulin secretion in a glucose-dependent manner. Additionally, semaglutide slows gastric emptying, suppresses appetite, and promotes satiety, contributing to its effects on weight loss and glycemic control [20–22].

Ligand-receptor binding affinity is an essential parameter in computer-assisted drug discovery and structure-based drug design [23]. Thanks to the continued development of experimental structural biology and the half-a-century old Protein Data Bank (PDB) [24–28], a comprehensive structural biophysical analysis becomes possible [29,30] for specific ligand-receptor complex structures deposited in PDB, such that our understanding of the structural and biophysical basis of their interfacial stability

 doi:10.20944/preprints202405.0258.v1

is able to help us modify the binding affinity of certain drug target and its interacting partners [31–35]. Take semaglutide for instance. On July 26, 2021, with a manually defined set of computational structural and biophysical analysis, a simple Val27-Arg28 exchange was for the first time introduced in the backbone of semaglutide to strengthen the semaglutide-GLP-1R binding affinity [7,9,36].



PDB file	Protein-Protein Complex	Δ G (kcal/mol)	Kd (M) at 37 °C
4ZGM [9]	semaglutide-GLP-1R	-7.8	3.4×10^{-6}
sema.pdb	mutant semaglutide-GLP-1R	-8.4	1.1×10^{-6}

Figure 1. Strengthening semaglutide-GLP-1R binding affinity via a Val27-Arg28 exchange in the peptide backbone of semaglutide [9]. This figure was prepared with PyMol [37]

2. Motivation

The development of semaglutide analogues with increased GLP-1R binding affinity holds significant clinical relevance, offering the potential for enhanced glucose control, weight loss, and cardiovascular benefits in patients with type 2 diabetes and obesity [3,38,39]. By leveraging insights from structural biology and computational modeling and biophysics, this article seeks to design semaglutide derivatives that exhibit tighter interactions with the GLP-1R binding site, thereby improving receptor activation and downstream signaling pathways. These analogues may represent a new class of GLP-1R agonists with superior therapeutic efficacy and reduced dosing frequency, addressing current limitations in the management of metabolic disorders [40,41].

 doi:10.20944/preprints202405.0258.v1

3. Materials and Methods

As listed in Table 1, there is **one** structure (determined by Cryo-EM) of Semaglutide-bound Glucagon-Like Peptide-1 (GLP-1) Receptor in Complex with Gs protein (PDB ID: 7KI0 [42]) as of May 6, 2024.

Table 1. Experimentally determined semaglutide-related structures (released newest from oldest) in the Protein Data Bank (PDB [24]) as of May 6, 2024, QUERY code: QUERY: Polymer Entity Description = "Semaglutide".

PDB ID	Structure Title (release date from newest to oldest)
7KI0	Semaglutide-bound Glucagon-Like Peptide-1 (GLP-1) Receptor in Complex with Gs protein

However, with a QUERY code: QUERY: Full Text = "Semaglutide", a total of three experimental structures related to semaglutide were found in the Protein Data Bank (PDB [24]), as listed in Table 2.

Table 2. Experimentally determined semaglutide-related structures (released newest from oldest) in the Protein Data Bank (PDB [24]) as of May 6, 2024, QUERY code: QUERY: Full Text = "Semaglutide".

PDB ID	Structure Title (release date from newest to oldest)
7KI0	Semaglutide-bound Glucagon-Like Peptide-1 (GLP-1) Receptor in Complex with Gs protein
7KI1	Taspoglutide-bound Glucagon-Like Peptide-1 (GLP-1) Receptor in Complex with Gs Protein
4ZGM	Crystal structure of Semaglutide peptide backbone in complex with the GLP-1 receptor extracellular domain

Among the three, there is **one** structure (determined by X-ray diffraction) of the semaglutide peptide backbone in complex with the extracellular domain of GLP-1R (PDB ID: 4ZGM [39]). Briefly, the amino acid sequences of the two chains of semaglutide and GLP-1R (according to PDB entry 4ZGM [39]) are listed in italics in fasta format as below,

>4ZGM_1 | Chain A | Glucagon-like peptide 1 receptor | Homo sapiens (9606)

RPQGATVSLWETVQKWREYRRQCQRSLTEDPPPATDLFCNRTFDEYACWPDGEPGSFVNVSC PWYLPWASSVPQGHVYRFCTAEGLWLQKDNSSLPWRDLSECEESKRGERSSPEEQLLFLY

>4ZGM_2 | Chain B | Semaglutide peptide backbone; 8Aib,34R-GLP-1(7-37)-OH | Homo sapiens (9606)

HAEGTFTSDVSSYLEGQAAKEFIAWLVRGRG

In combination with the comprehensive structural and biophysical analysis [29], the key amino acid residues at the semaglutide-GLP-1R complex binding interface (PDB ID: 4ZGM) were examined carefully in PyMol [9,37,39], and the inter-residue distances were calculated by PyMol [37] to identify potential neighbouring residue pair(s) to modulate the structural stability of the semaglutide-GLP-1R complex structure, leading to the design of a set of semaglutide variants with enhanced binding affinity.

In general, after homology structural modeling of semaglutide variants with Modeller [43], the binding affinity between semaglutide and GLP-1R was calculated using Prodigy [44,45]. Specifically, a total of ($s = g(28,4) = \frac{28!}{4!(24)!} \times 20^4$ [46]) set of semaglutide analogues were generated with in-house Python script with four site-specific missense mutations introduced into native semaglutide sequences as listed above. Afterwards, homology structural modeling was carried out using Modeller [43] with PDB entry 4ZGM [39] as the structural template. Finally, the binding affinity between semaglutide and GLP-1R was calculated using Prodigy [44,45] for native semaglutide (10000 times) and for $s = g(28,4) = \frac{28!}{4!(24)!} \times 20^4$ [46] semaglutide analogues (twenty times each).

 doi:10.20944/preprints202405.0258.v1

4. Results

First off, with the X-ray structure of the semaglutide peptide backbone in complex with the extracellular domain of GLP-1R (PDB ID: 4ZGM [39]) in place, Modeller [43] was employed to build 10000 structural models with 100% homology to PDB ID: 4ZGM [39], and the binding affinity between semaglutide and GLP-1R was calculated using Prodigy [44,45] for native semaglutide (10000 times). As shown in Figure 2, most of the K_d values are located between 2.5×10^{-6} M and 4.0×10^{-6} M, with an average at 3.278×10^{-6} M, which is rather close to the one K_d (3.4×10^{-6} M) as reported in [9].

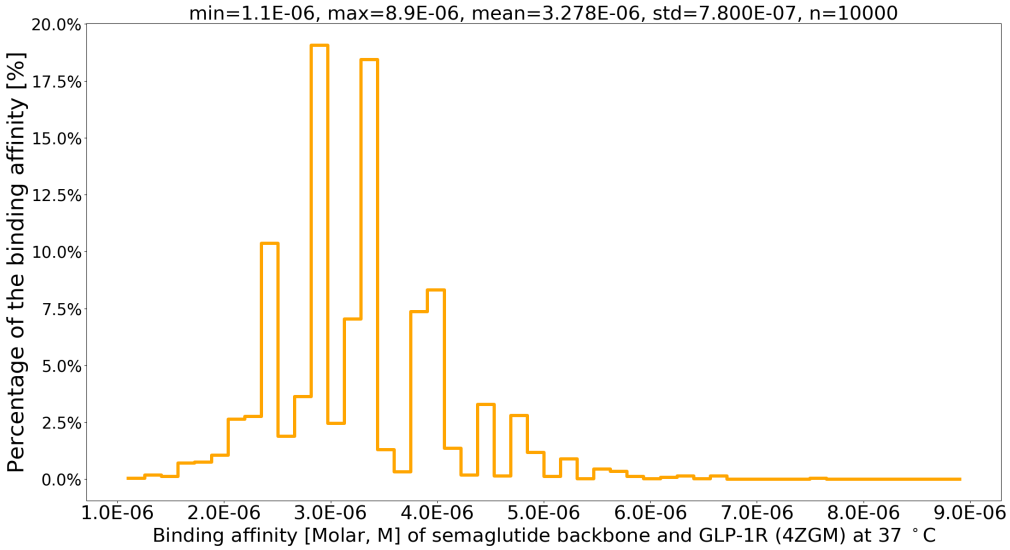


Figure 2. Distribution of the binding affinities between semaglutide (PDB ID: 4ZGM [39]) and GLP-1R as calculated by Prodigy [44,45].

Secondly, with the X-ray structure of the semaglutide peptide backbone in complex with the extracellular domain of GLP-1R (PDB ID: 4ZGM [39]) as the structural template, a total of $s = g(28,4) = \frac{28!}{4!(24)!} \times 20^4$ [46] semaglutide variants' sequence were generated, and plugged into Modeller [43] to build 20 structural models for each semaglutide analogue, and the binding affinity between semaglutide and GLP-1R was calculated using Prodigy [44,45]. In total, the binding affinities of 100 semaglutide analogues to GLP-1R are included in Table 3, including their minimum, maximum, average and standard deviation of the K_d values calculated using Prodigy [44,45] for the semaglutide analogues, each 20 times of homology structural modeling using Modeller [43] In supplementary file **supps.pdf**, Table 1 includes a total of 8915 semaglutide analogues, including their minimum, maximum, average and standard deviation of the K_d values calculated using Prodigy [44,45] for the semaglutide analogues, each 20 times of homology structural modeling using Modeller [43].

 doi:10.20944/preprints202405.0258.v1

Table 3. Computationally designed semaglutide analogues with elevated binding affinity to GLP-1R than native semaglutide. In this table, the binding affinity of semaglutide analogues to GLP-1R is calculated with Prodigy [44,45] at K_d (37 $^{\circ}C$) values, while **Muta1**, **Muta2**, **Muta3** and **Muta4** represent the four site-specific mutations introduced into the backbone of semaglutide, and **Min**, **Max**, **Mean** and **Std** represent the minimum, the maximum, the average and the standard deviation of the K_d values calculated using Prodigy [44,45] for the semaglutide analogues, each 20 times of homology structural modeling using Modeller [43].

No.	Muta1	Muta2	Muta3	Muta4	Min	Max	Mean	Std
1	G13B_A	I20B_Q	L23B_Q	V24B_N	5.3E-08	2.2E-07	1.337E-07	4.778E-08
2	G13B_A	I20B_N	L23B_R	V24B_N	6.5E-08	2.4E-07	1.344E-07	4.996E-08
3	G13B_A	I20B_N	L23B_Q	V24B_T	6.6E-08	2.2E-07	1.376E-07	4.199E-08
4	G13B_A	I20B_T	L23B_Q	V24B_N	8.0E-08	3.1E-07	1.404E-07	5.478E-08
5	G13B_A	I20B_Q	L23B_Q	V24B_T	6.8E-08	2.0E-07	1.407E-07	3.779E-08
6	G13B_A	I20B_S	L23B_R	V24B_T	6.1E-08	2.5E-07	1.408E-07	5.527E-08
7	G13B_A	I20B_Q	L23B_R	V24B_N	3.0E-08	3.2E-07	1.461E-07	7.095E-08
8	G13B_A	I20B_T	L23B_R	V24B_N	8.3E-08	2.1E-07	1.467E-07	3.690E-08
9	G13B_A	I20B_N	L23B_R	V24B_Q	6.3E-08	2.9E-07	1.487E-07	5.848E-08
10	G13B_A	I20B_Q	L23B_R	V24B_Q	8.6E-08	2.5E-07	1.489E-07	5.170E-08
11	G13B_A	I20B_Q	L23B_Q	V24B_Q	6.3E-08	2.4E-07	1.505E-07	5.269E-08
12	G13B_A	I20B_S	L23B_R	V24B_N	4.4E-08	3.5E-07	1.520E-07	6.568E-08
13	G13B_A	I20B_T	L23B_R	V24B_T	9.4E-08	2.2E-07	1.545E-07	4.188E-08
14	G13B_A	I20B_N	L23B_Q	V24B_N	7.7E-08	2.2E-07	1.559E-07	4.164E-08
15	G13B_A	I20B_S	L23B_R	V24B_Q	7.7E-08	3.0E-07	1.571E-07	6.401E-08
16	G13B_A	I20B_S	F19B_Q	V24B_N	3.5E-08	2.8E-07	1.583E-07	6.648E-08
17	G13B_A	I20B_N	L23B_Q	V24B_Q	8.2E-08	2.9E-07	1.602E-07	5.879E-08
18	G13B_A	I20B_N	F19B_Q	V24B_N	5.0E-08	2.9E-07	1.634E-07	7.035E-08
19	G13B_A	I20B_T	F19B_Q	V24B_Q	9.7E-08	2.9E-07	1.653E-07	4.839E-08
20	G13B_A	I20B_N	L23B_R	V24B_T	8.0E-08	3.4E-07	1.662E-07	8.233E-08
21	G13B_A	I20B_S	L23B_Q	V24B_Q	7.8E-08	3.4E-07	1.682E-07	7.062E-08
22	G13B_A	I20B_Q	F19B_Q	V24B_T	3.6E-08	3.3E-07	1.686E-07	7.653E-08
23	G13B_A	I20B_S	L23B_Q	V24B_N	6.2E-08	3.0E-07	1.703E-07	6.652E-08
24	G13B_A	E18B_T	I20B_Q	V24B_Q	8.4E-08	2.7E-07	1.706E-07	5.558E-08
25	G13B_A	I20B_T	F19B_Q	V24B_N	6.3E-08	2.9E-07	1.711E-07	7.268E-08
26	Y10B_A	I20B_Q	L23B_Q	V24B_Q	7.3E-08	3.8E-07	1.725E-07	8.816E-08
27	G13B_A	I20B_S	F19B_Q	V24B_T	6.2E-08	3.2E-07	1.725E-07	6.884E-08
28	G13B_A	I20B_T	L23B_R	V24B_Q	5.5E-08	3.0E-07	1.725E-07	6.902E-08
29	G13B_A	I20B_T	F19B_Q	V24B_T	6.5E-08	3.3E-07	1.765E-07	7.216E-08
30	G13B_A	I20B_T	L23B_Q	V24B_T	9.2E-08	3.4E-07	1.765E-07	7.525E-08
31	G13B_A	I20B_Q	L23B_R	V24B_T	8.3E-08	3.7E-07	1.778E-07	7.613E-08
32	G13B_A	I20B_S	L23B_Q	V24B_T	7.8E-08	3.6E-07	1.785E-07	6.839E-08
33	G13B_A	I20B_T	L23B_Q	V24B_Q	9.1E-08	2.9E-07	1.791E-07	4.888E-08
34	G13B_A	I20B_Q	K17B_S	V24B_N	7.5E-08	2.9E-07	1.822E-07	7.042E-08
35	G13B_A	I20B_Q	K17B_S	V24B_T	8.2E-08	2.9E-07	1.826E-07	5.543E-08
36	G13B_A	I20B_Q	S9B_A	V24B_N	7.7E-08	3.5E-07	1.836E-07	6.655E-08
37	G13B_A	E18B_T	I20B_Q	V24B_N	9.5E-08	2.7E-07	1.842E-07	5.523E-08
38	Y10B_A	I20B_Q	L23B_R	V24B_N	8.1E-08	4.7E-07	1.851E-07	9.983E-08
39	G13B_A	I20B_Q	F19B_Q	V24B_N	6.0E-08	3.1E-07	1.869E-07	6.942E-08
40	G13B_A	I20B_Q	F19B_Q	V24B_Q	7.3E-08	3.4E-07	1.870E-07	7.308E-08
41	Y10B_A	I20B_N	L23B_R	V24B_Q	8.9E-08	3.7E-07	1.890E-07	7.203E-08
42	G13B_A	I20B_Q	S9B_A	V24B_T	9.1E-08	3.6E-07	1.895E-07	7.000E-08
43	G13B_A	I20B_Q	S9B_A	L23B_R	8.6E-08	4.6E-07	1.900E-07	8.860E-08
44	G13B_A	I20B_N	F19B_Q	V24B_Q	6.8E-08	3.4E-07	1.909E-07	6.762E-08
45	G13B_A	I20B_S	F19B_Q	V24B_Q	8.8E-08	3.5E-07	1.926E-07	7.739E-08
46	Y10B_A	I20B_Q	L23B_R	V24B_T	9.9E-08	3.5E-07	1.934E-07	7.962E-08
47	Y10B_A	I20B_Q	L23B_R	V24B_Q	6.6E-08	3.3E-07	1.951E-07	7.414E-08
48	G13B_A	I20B_Q	K17B_S	V24B_Q	8.3E-08	3.6E-07	1.962E-07	6.561E-08
49	Y10B_A	I20B_T	L23B_R	V24B_T	1.1E-07	3.4E-07	1.970E-07	6.689E-08
50	G13B_A	I20B_T	K17B_S	V24B_N	1.1E-07	3.0E-07	1.975E-07	5.524E-08

 doi:10.20944/preprints202405.0258.v1

Table 3. *Cont.*

51	Y10B_A	I20B_N	L23B_Q	V24B_N	8.3E-08	3.4E-07	1.976E-07	6.715E-08
52	Y10B_A	I20B_N	L23B_R	V24B_T	1.1E-07	3.5E-07	1.985E-07	6.784E-08
53	G13B_A	I20B_Q	S9B_A	V24B_Q	9.0E-08	3.0E-07	1.985E-07	6.548E-08
54	G13B_A	I20B_S	F19B_Q	L23B_R	1.2E-07	3.1E-07	1.990E-07	5.077E-08
55	G13B_A	I20B_N	K17B_S	V24B_N	1.2E-07	4.4E-07	2.015E-07	7.264E-08
56	G13B_A	I20B_S	F19B_Q	L23B_Q	7.7E-08	4.1E-07	2.023E-07	9.647E-08
57	G13B_A	I20B_S	L23B_R	K17B_S	1.1E-07	2.8E-07	2.030E-07	4.736E-08
58	Y10B_A	I20B_T	L23B_R	V24B_N	8.9E-08	4.4E-07	2.034E-07	7.708E-08
59	G13B_A	I20B_N	K17B_S	V24B_Q	1.0E-07	5.2E-07	2.035E-07	8.707E-08
60	Y10B_A	I20B_S	L23B_R	V24B_T	8.6E-08	3.5E-07	2.038E-07	7.595E-08
61	G13B_A	I20B_N	S9B_A	L23B_R	1.2E-07	3.8E-07	2.050E-07	8.300E-08
62	G13B_A	I20B_N	F19B_Q	V24B_T	8.4E-08	4.1E-07	2.067E-07	7.351E-08
63	G13B_A	I20B_N	L23B_R	K17B_S	1.1E-07	3.8E-07	2.070E-07	8.228E-08
64	G13B_A	E18B_T	I20B_Q	L23B_Q	1.1E-07	3.4E-07	2.080E-07	7.317E-08
65	G13B_A	I20B_N	L23B_Q	K17B_S	1.1E-07	3.6E-07	2.080E-07	7.557E-08
66	G13B_A	I20B_T	F19B_Q	L23B_R	6.7E-08	3.5E-07	2.097E-07	9.136E-08
67	G13B_A	I20B_S	K17B_S	V24B_N	8.6E-08	3.6E-07	2.098E-07	7.261E-08
68	Y10B_A	I20B_N	F19B_Q	V24B_Q	7.6E-08	3.9E-07	2.098E-07	8.400E-08
69	Y10B_A	I20B_N	L23B_R	V24B_N	1.1E-07	3.4E-07	2.100E-07	6.696E-08
70	G13B_A	I20B_T	S9B_A	L23B_R	8.9E-08	3.8E-07	2.104E-07	7.158E-08
71	Y10B_A	I20B_T	L23B_R	V24B_Q	1.2E-07	3.2E-07	2.105E-07	5.698E-08
72	G13B_A	I20B_Q	L23B_R	K17B_S	8.2E-08	7.9E-07	2.120E-07	1.610E-07
73	G13B_A	I20B_T	K17B_S	V24B_Q	9.2E-08	4.2E-07	2.121E-07	7.875E-08
74	G13B_A	I20B_N	E18B_T	V24B_N	1.3E-07	3.8E-07	2.125E-07	7.181E-08
75	G13B_A	I20B_T	L23B_R	K17B_S	9.5E-08	4.5E-07	2.128E-07	9.408E-08
76	G13B_A	E18B_T	I20B_S	V24B_T	1.0E-07	3.5E-07	2.135E-07	7.081E-08
77	G13B_A	I20B_Q	F19B_W	V24B_Q	8.9E-08	3.9E-07	2.149E-07	8.920E-08
78	Y10B_A	I20B_Q	L23B_Q	V24B_N	6.7E-08	4.5E-07	2.161E-07	1.064E-07
79	Y10B_A	I20B_N	L23B_Q	V24B_Q	9.7E-08	3.3E-07	2.168E-07	5.626E-08
80	G13B_A	I20B_N	E18B_T	L23B_R	1.1E-07	4.0E-07	2.175E-07	9.118E-08
81	G13B_A	I20B_Q	V24B_Q	E12B_A	1.0E-07	3.9E-07	2.175E-07	8.175E-08
82	Y10B_A	I20B_S	L23B_Q	V24B_Q	8.7E-08	3.4E-07	2.179E-07	7.437E-08
83	G13B_A	I20B_N	F19B_Q	L23B_R	8.9E-08	4.4E-07	2.179E-07	9.986E-08
84	Y10B_A	I20B_Q	K17B_S	V24B_T	1.1E-07	4.0E-07	2.195E-07	7.970E-08
85	G13B_A	I20B_S	S9B_A	V24B_Q	9.9E-08	3.2E-07	2.199E-07	5.489E-08
86	G13B_A	I20B_Q	L23B_Q	K17B_S	7.5E-08	3.6E-07	2.208E-07	7.533E-08
87	Y10B_A	I20B_T	L23B_Q	V24B_N	1.0E-07	4.7E-07	2.210E-07	9.037E-08
88	Y10B_A	I20B_Q	F19B_Q	V24B_N	6.8E-08	5.7E-07	2.218E-07	1.294E-07
89	G13B_A	I20B_S	F19B_W	V24B_N	1.3E-07	3.7E-07	2.225E-07	7.376E-08
90	Y10B_A	I20B_Q	F19B_Q	V24B_Q	9.9E-08	4.3E-07	2.244E-07	8.656E-08
91	G13B_A	E18B_T	I20B_Q	L23B_R	8.6E-08	3.9E-07	2.248E-07	9.776E-08
92	G13B_A	I20B_T	S9B_A	V24B_Q	1.4E-07	3.7E-07	2.250E-07	6.653E-08
93	Y10B_A	I20B_S	L23B_R	V24B_N	1.4E-07	3.5E-07	2.260E-07	6.344E-08
94	G13B_A	I20B_T	F19B_Q	L23B_Q	6.0E-08	4.6E-07	2.263E-07	1.111E-07
95	G13B_A	I20B_N	E18B_T	V24B_Q	1.0E-07	4.0E-07	2.265E-07	7.700E-08
96	Y10B_A	I20B_Q	F19B_Q	V24B_T	8.5E-08	3.2E-07	2.267E-07	6.989E-08
97	Y10B_A	I20B_T	L23B_Q	V24B_Q	9.6E-08	4.5E-07	2.268E-07	9.045E-08
98	G13B_A	I20B_N	E18B_T	V24B_T	1.2E-07	4.4E-07	2.270E-07	8.572E-08
99	G13B_A	E18B_T	I20B_T	L23B_Q	1.0E-07	4.9E-07	2.270E-07	1.064E-07
100	Y10B_A	I20B_N	F19B_Q	V24B_N	9.0E-08	4.4E-07	2.277E-07	9.537E-08

Among the 100 semaglutide analogues included in Table 3, one particular semaglutide analogue stood out, named here as semaglutideX, where the semaglutideX-GLP-1R structural model is reaching a K_d value of 3.0×10^{-8} M, while the K_d is 3.4×10^{-6} M for the binding of native semaglutide to GLP-1 [9]. The amino acid sequence of semaglutideX is listed in italics in fasta format as below,

>semaglutideX (supplementary file **semx.pdb**)

HAEGTFTSDVSSYLEAQAAKEFQAWRNRGRG

 doi:10.20944/preprints202405.0258.v1

For a close comparison, the amino acid sequence of semaglutide (PDB ID: 4ZGM [39]) is listed in italics in fasta format as below,

>4ZGM_2 | Chain B | Semaglutide peptide backbone; 8Aib,34R-GLP-1(7-37)-OH | Homo sapiens (9606)

HAEGTFTSDVSSYLEGQAAKEFIAWLVRGRG

and the amino acid sequence of semaglutide with a Val27-Arg28 exchange [9] is listed in italics in fasta format as below,

>Val27-Arg28 exchange

HAEGTFTSDVSSYLEGQAAKEFIAWLRVGRG

Table 4. The binding affinities of semaglutide, semaglutide with a Val27-Arg28 exchange [9] and semaglutideX to GLP-1R calculated by Prodigy [44,45]. In this table, 4ZGM represents the experimental structure (determined by X-ray diffraction) of the semaglutide peptide backbone in complex with the extracellular domain of GLP-1R (PDB ID: 4ZGM), mutant semaglutide represents the B27Arg-B28Val mutant of semaglutide, whose structural model is described in the supplementary file **semx.pdb**, and semaglutideX represents a semaglutide variant with four site-specific missense mutations, i.e., G13B_A I20B_Q L23B_R V24B_N.

PDB file	Protein-Protein Complex	Δ G (kcal/mol)	Kd (M) at 37 °C	Fold to 4ZGM
4ZGM [39]	semaglutide-GLP-1R [39]	-7.8	3.4×10^{-6}	1
sema.pdb [9]	Val27-Arg28 exchange [9]	-8.4	1.1×10^{-6}	3.09
semx.pdb	G13B_A I20B_Q L23B_R V24B_N	-10.7	3.0×10^{-8}	113.33

With the X-ray structure of the semaglutide peptide backbone in complex with the extracellular domain of GLP-1R (PDB ID: 4ZGM [39]) in place, Modeller [43] was employed to build 20 structural models for each one of the 8915 semaglutide variants (supplementary file **semx.pdb**), and the binding affinity between semaglutide variants and GLP-1R were calculated using Prodigy [44,45] for 20 times. As shown in Figure 3, most of the K_d values are located between 1.0×10^{-7} M and 2.0×10^{-7} M, with an average at 1.437×10^{-7} M, which is at least one order of degree lower that the K_d (3.4×10^{-6} M) as reported in [9].

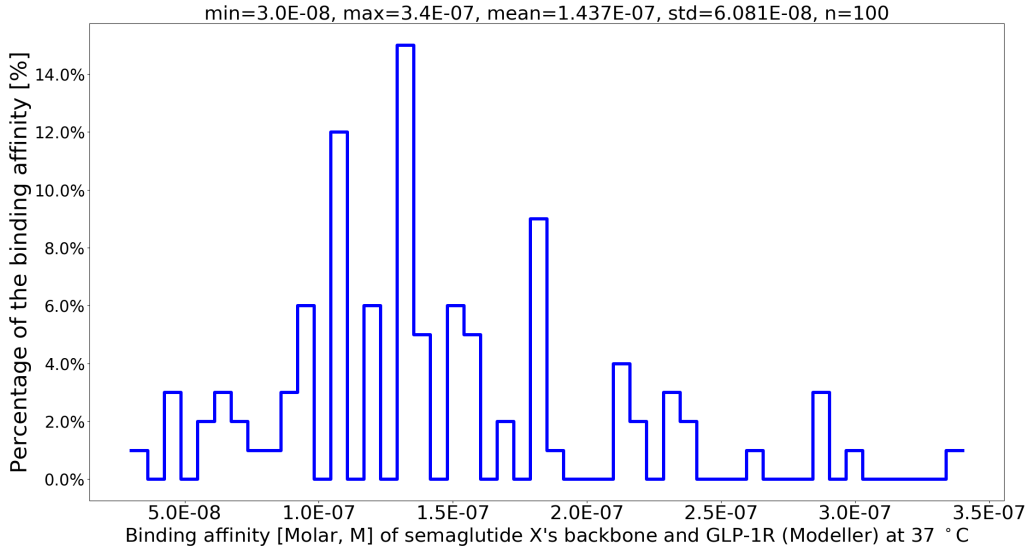


Figure 3. Distribution of the binding affinities between semaglutideX (supplementary file **semx.pdb**) and GLP-1R as calculated by Prodigy [44,45].

 doi:10.20944/preprints202405.0258.v1

5. Conclusion

To sum up, this article reports a set of computationally designed semaglutide analogues with elevated binding affinity to the GLP-1 receptor compared to native semaglutide and also to a previous semaglutide with a Val27-Arg28 exchange [9]. Overall, these analogues hold promise for improving the treatment of type 2 diabetes and obesity by enhancing receptor activation and downstream signaling cascades. Future preclinical and clinical studies are needed to evaluate the pharmacological properties and therapeutic potential of these novel semaglutide derivatives. Moreover, continued optimization of GLP-1R agonists through structure-based drug design approaches could lead to further advancements in the field, ultimately benefiting patients with diabetes.

6. Discussion

Pharmacologically, there are reasons to design semaglutide analogues with elevated binding affinity to the glucagon-like peptide-1 receptor (GLP-1R) compared to semaglutide itself. However, designing semaglutide analogues with heightened binding affinity to the glucagon-like peptide-1 receptor (GLP-1R) presents both opportunities and challenges in the field of metabolic therapeutics. On the positive side, analogues with enhanced receptor affinity hold the potential to improve therapeutic outcomes by maximizing receptor activation and downstream signaling pathways, leading to more robust glucose control, weight loss, and potentially cardioprotective effects [47]. Additionally, these analogues may offer the advantage of reduced dosing frequency, enhancing patient compliance and convenience [48]. However, there are also drawbacks to consider, including the risk of off-target effects [49] and increased receptor desensitization with prolonged exposure to highly potent semaglutide analogues. Furthermore, the development of analogues with elevated binding affinity necessitates careful optimization to maintain selectivity and minimize adverse reactions [50,51]. Balancing these factors is crucial for realizing the full therapeutic potential of semaglutide analogues with enhanced GLP-1R binding affinity while mitigating potential risks, which is where a truly general intermolecular binding affinity calculator [46,52,53] will be useful to accomodate off-target effects and drug-drug interactions [49–51].

Finally, while this study represents a series of semaglutide analogues with elevated binding affinity to the GLP-1 receptor compared to native semaglutide and also to a previous semaglutide with a Val27-Arg28 exchange [9], the entire process of the design, along with the structural biophysics-based strategy for the molecular design [31–34,54], is essentially also a process of the construction of a semaglutide-GLP-1R based mini general intermolecular binding affinity calculator (GIBAC) [46,52, 53] based on the structure of semaglutide peptide backbone in complex with the GLP-1 receptor extracellular domain determined by X-ray diffraction [39]. With a series of semaglutide analogues with elevated binding affinity to the GLP-1R available and this semaglutide-GLP-1R based mini GIBAC in place, here, this articles calls again for the construction of a truly general intermolecular binding affinity calculator to be listed on the agenda of the entire community of drug discovery and design [46,52,53].

7. Acknowledgment

The author is grateful to the communities of structural biology, biophysics, medicinal and computational chemistry and algorithm design, for the continued accumulation of knowledge and data for drug discovery & design, and for the continued development of tools (hardware, software and algorithm) for drug discovery & design.

8. Ethical Statement

No ethical approval is required.

 doi:10.20944/preprints202405.0258.v1

9. Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the author used OpenAI's ChatGPT in order to improve the readability of the manuscript, and to make it as concise and short as possible. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Author Contributions: Conceptualization, W.L.; methodology, W.L.; software, W.L.; validation, W.L.; formal analysis, W.L.; investigation, W.L.; resources, W.L.; data duration, W.L.; writing–original draft preparation, W.L.; writing–review and editing, W.L.; visualization, W.L.; supervision, W.L.; project administration, W.L.; funding acquisition, not applicable.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

1. Marijic, J.; Neelankavil, J.P. Semaglutide: A New Medical Swiss Army Knife? *Journal of Cardiothoracic and Vascular Anesthesia* **2024**, *38*, 871–873. doi:10.1053/j.jvca.2023.12.032.
2. Cowart, K. Oral Semaglutide: First-in-Class Oral GLP-1 Receptor Agonist for the Treatment of Type 2 Diabetes Mellitus. *Annals of Pharmacotherapy* **2019**, *54*, 478–485.
3. Knudsen, L.B.; Lau, J. The Discovery and Development of Liraglutide and Semaglutide. *Frontiers in Endocrinology* **2019**, *10*, 1–32.
4. Yang, J.; Anishchenko, I.; Park, H.; Peng, Z.; Ovchinnikov, S.; Baker, D. Improved protein structure prediction using predicted interresidue orientations. *Proceedings of the National Academy of Sciences* **2020**, *117*, 1496–1503.
5. Han, J.; Fu, J.; Yang, Q.; Zhou, F.; Chen, X.; Li, C.; Yin, J. Rational design and biological evaluation of gemfibrozil modified Xenopus GLP-1 derivatives as long-acting hypoglycemic agents. *European Journal of Medicinal Chemistry* **2020**, *198*, 112389.
6. Aschenbrenner, D.S. New Drug for Type 2 Diabetes. *AJN, American Journal of Nursing* **2020**, *120*, 25.
7. Ahrén, B.; Atkin, S.L.; Charpentier, G.; Warren, M.L.; Wilding, J.P.H.; Birch, S.; Holst, A.G.; Leiter, L.A. Semaglutide induces weight loss in subjects with type 2 diabetes regardless of baseline BMI or gastrointestinal adverse events in the SUSTAIN 1 to 5 trials. *Diabetes, Obesity and Metabolism* **2018**, *20*, 2210–2219.
8. Lee, Y.S.; Jun, H.S. Anti-diabetic actions of glucagon-like peptide-1 on pancreatic beta-cells. *Metabolism* **2014**, *63*, 9–19.
9. Li, W. Strengthening Semaglutide-GLP-1R Binding Affinity via a Val27-Arg28 Exchange in the Peptide Backbone of Semaglutide: A Computational Structural Approach. *Journal of Computational Biophysics and Chemistry* **2021**, *20*, 495–499.
10. Wang, W.; Volkow, N.D.; Berger, N.A.; Davis, P.B.; Kaelber, D.C.; Xu, R. Association of semaglutide with risk of suicidal ideation in a real-world cohort. *Nature Medicine* **2024**, *30*, 168–176. doi:10.1038/s41591-023-02672-2.
11. Li, W. Designing Insulin Analogues with Lower Binding Affinity to Insulin Receptor than That of Insulin Icodec **2024**. doi:10.20944/preprints202404.1922.v1.
12. Li, W. Delving Deep into the Structural Aspects of the BPro28-BLys29 Exchange in Insulin Lispro: A Structural Biophysical Lesson **2020**.
13. Kosiborod, M.N.; Petrie, M.C.; Borlaug, B.A.; Butler, J.; Davies, M.J.; Hovingh, G.K.; Kitzman, D.W.; Møller, D.V.; Treppendahl, M.B.; Verma, S.; Jensen, T.J.; Liisberg, K.; Lindegaard, M.L.; Abhayaratna, W.; Ahmed, F.Z.; Ben-Gal, T.; Chopra, V.; Ezekowitz, J.A.; Fu, M.; Ito, H.; Lelonek, M.; Melenovský, V.; Merkely, B.; Núñez, J.; Perna, E.; Schou, M.; Senni, M.; Sharma, K.; van der Meer, P.; Von Lewinski, D.; Wolf, D.; Shah, S.J. Semaglutide in Patients with Obesity-Related Heart Failure and Type 2 Diabetes. *New England Journal of Medicine* **2024**, *390*, 1394–1407. doi:10.1056/nejmoa2313917.
14. Wilding, J.P.H. Semaglutide in weight management: Author's reply. *The Lancet* **2019**, *394*, 1226–1227.

15. Bucheit, J.D.; Pamulapati, L.G.; Carter, N.; Malloy, K.; Dixon, D.L.; Sisson, E.M. Oral Semaglutide: A Review of the First Oral Glucagon-Like Peptide 1 Receptor Agonist. *Diabetes Technology & Therapeutics* **2020**, *22*, 10–18.
16. Kanters, S.; Wilkinson, L.; Vrazic, H.; Sharma, R.; Lopes, S.; Popoff, E.; Druyts, E. Comparative efficacy of once-weekly semaglutide versus SGLT-2 inhibitors in patients inadequately controlled with one to two oral antidiabetic drugs: a systematic literature review and network meta-analysis. *BMJ Open* **2019**, *9*, e023458.
17. Kosiborod, M.N.; Verma, S.; Borlaug, B.A.; Butler, J.; Davies, M.J.; Jon Jensen, T.; Rasmussen, S.; Erlang Marstrand, P.; Petrie, M.C.; Shah, S.J.; Ito, H.; Schou, M.; Melenovský, V.; Abhayaratna, W.; Kitzman, D.W. Effects of Semaglutide on Symptoms, Function, and Quality of Life in Patients With Heart Failure With Preserved Ejection Fraction and Obesity: A Prespecified Analysis of the STEP-HFpEF Trial. *Circulation* **2024**, *149*, 204–216. doi:10.1161/circulationaha.123.067505.
18. Nadkarni, P.; Chepurny, O.G.; Holz, G.G. Regulation of Glucose Homeostasis by GLP-1. In *Progress in Molecular Biology and Translational Science*; Elsevier, 2014; pp. 23–65.
19. Bucheit, J.D.; Pamulapati, L.G.; Carter, N.; Malloy, K.; Dixon, D.L.; Sisson, E.M. Oral Semaglutide: A Review of the First Oral Glucagon-Like Peptide 1 Receptor Agonist. *Diabetes Technology & Therapeutics* **2020**, *22*, 10–18.
20. Granhall, C.; Donsmark, M.; Blicher, T.M.; Golor, G.; Sondergaard, F.L.; Thomsen, M.; Bakdal, T.A. Safety and Pharmacokinetics of Single and Multiple Ascending Doses of the Novel Oral Human GLP-1 Analogue, Oral Semaglutide, in Healthy Subjects and Subjects with Type 2 Diabetes. *Clinical Pharmacokinetics* **2018**, *58*, 781–791. doi:10.1007/s40262-018-0728-4.
21. Garg, S.K.; Kaur, G.; Haider, Z.; Rodriquez, E.; Beatson, C.; Snell-Bergeon, J. Efficacy of Semaglutide in Overweight and Obese Patients with Type 1 Diabetes. *Diabetes Technology & Therapeutics* **2024**, *26*, 184–189. doi:10.1089/dia.2023.0490.
22. Europe, T.L.R.H. Semaglutide and beyond a turning point in obesity pharmacotherapy. *The Lancet Regional Health Europe* **2024**, *37*, 100860. doi:10.1016/j.lanepe.2024.100860.
23. Fuji, H.; Qi, F.; Qu, L.; Takaesu, Y.; Hoshino, T. Prediction of Ligand Binding Affinity to Target Proteins by Molecular Mechanics Theoretical Calculation. *Chemical and Pharmaceutical Bulletin* **2017**, *65*, 461–468.
24. Berman, H.; Henrick, K.; Nakamura, H. Announcing the worldwide Protein Data Bank. *Nature Structural & Molecular Biology* **2003**, *10*, 980–980.
25. Li, W. Half-a-century Burial of ρ, θ and φ in PDB **2021**. doi:10.20944/preprints202103.0590.v1.
26. Li, W. Visualising the Experimentally Uncharted Territories of Membrane Protein Structures inside Protein Data Bank **2020**.
27. Li, W. A Local Spherical Coordinate System Approach to Protein 3D Structure Description **2020**.
28. Li, W. Structurally Observed Electrostatic Features of the COVID-19 Coronavirus-Related Experimental Structures inside Protein Data Bank: A Brief Update **2020**.
29. Li, W. How do SMA-linked mutations of *SMN1* lead to structural/functional deficiency of the SMA protein? *PLOS ONE* **2017**, *12*, e0178519.
30. Li, W. Extracting the Interfacial Electrostatic Features from Experimentally Determined Antigen and/or Antibody-Related Structures inside Protein Data Bank for Machine Learning-Based Antibody Design **2020**.
31. Li, W. Structural Identification of the Electrostatic Hot Spots for Severe Acute Respiratory Syndrome Coronavirus Spike Protein to Be Complexed with Its Receptor ACE2 and Its Neutralizing Antibodies **2020**.
32. Li, W. Calcium Channel Trafficking Blocker Gabapentin Bound to the α-2-δ-1 Subunit of Voltage-Gated Calcium Channel: A Computational Structural Investigation **2020**.
33. Li, W. Inter-Molecular Electrostatic Interactions Stabilizing the Structure of the PD-1/PD-L1 Axis: A Structural Evolutionary Perspective **2020**.
34. Li, W. Structural and Functional Consequences of the SMA-Linked Missense Mutations of the Survival Motor Neuron Protein: A Brief Update. In *Novel Aspects on Motor Neuron Disease*; IntechOpen, 2019.
35. Li, W. Designing Nerve Growth Factor Analogues to Suppress Pain Signal Transduction Mediated by the p75NTR-NGF-TrkA Complex: A Structural and Biophysical Perspective **2024**. doi:10.20944/preprints202403.1756.v1.

36. Aroda, V.R.; Rosenstock, J.; Terauchi, Y.; Altuntas, Y.; Lalic, N.M.; Villegas, E.C.M.; Jeppesen, O.K.; Christiansen, E.; Hertz, C.L.; Haluzík, M. PIONEER 1: Randomized Clinical Trial of the Efficacy and Safety of Oral Semaglutide Monotherapy in Comparison With Placebo in Patients With Type 2 Diabetes. *Diabetes Care* **2019**, *42*, 1724–1732.
37. DeLano, W.L. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography* **2002**, *40*, 82–92.
38. Blüher, M.; Rosenstock, J.; Hoefler, J.; Manuel, R.; Hennige, A.M. Dose–response effects on HbA1c and bodyweight reduction of survodutide, a dual glucagon/GLP-1 receptor agonist, compared with placebo and open-label semaglutide in people with type 2 diabetes: a randomised clinical trial. *Diabetologia* **2023**, *67*, 470–482. doi:10.1007/s00125-023-06053-9.
39. Lau, J.; Bloch, P.; Schäffer, L.; Pettersson, I.; Spetzler, J.; Kofoed, J.; Madsen, K.; Knudsen, L.B.; McGuire, J.; Steensgaard, D.B.; Strauss, H.M.; Gram, D.X.; Knudsen, S.M.; Nielsen, F.S.; Thygesen, P.; Reedtz-Runge, S.; Kruse, T. Discovery of the Once-Weekly Glucagon-Like Peptide-1 (GLP-1) Analogue Semaglutide. *Journal of Medicinal Chemistry* **2015**, *58*, 7370–7380.
40. Anderson, S.L.; Beutel, T.R.; Trujillo, J.M. Oral semaglutide in type 2 diabetes. *Journal of Diabetes and its Complications* **2020**, *34*, 107520.
41. Pieber, T.R.; Bode, B.; Mertens, A.; Cho, Y.M.; Christiansen, E.; Hertz, C.L.; Wallenstein, S.O.R.; Buse, J.B.; Akın, S.; Aladağ, N.; Arif, A.A.; Aronne, L.J.; Aronoff, S.; Ataoglu, E.; Baik, S.H.; Bays, H.; Beckett, P.L.; Berker, D.; Bilz, S.; Bode, B.; Braun, E.W.; Buse, J.B.; Canani, L.H.S.; Cho, Y.M.; Chung, C.H.; Colin, I.; Condit, J.; Cooper, J.; Delgado, B.; Eagerton, D.C.; Ebrashy, I.N.E.; Hefnawy, M.H.M.F.E.; Eliaschewitz, F.G.; Finneran, M.P.; Fischli, S.; Fließer-Görzer, E.; Geohas, J.; Godbole, N.A.; Golay, A.; de Lapertosa, S.G.; Gross, J.L.; Gulseth, H.L.; Helland, F.; Høivik, H.O.; Issa, C.; Kang, E.S.; Keller, C.; Khalil, S.H.A.; Kim, N.H.; Kim, I.J.; Klaff, L.J.; Laimer, M.; LaRocque, J.C.; Lederman, S.N.; Lee, K.W.; Litchfield, W.R.; Manning, M.B.; Mertens, A.; Morawski, E.J.; Murray, A.V.; Nicol, P.R.; O'Connor, T.M.; Oğuz, A.; Ong, S.; özdemir, A.; Palace, E.M.; Palchick, B.A.; Pereles-Ortiz, J.; Pieber, T.; Prager, R.; Preumont, V.; Riffer, E.; Rista, L.; Rudofsky, G.; Sarı, R.; Scheen, A.; Schultes, B.; Seo, J.A.; Shelbaya, S.A.; Sivalingam, K.; Sorli, C.H.; Stäuble, S.; Streja, D.A.; T'Sjoen, G.; Tetiker, T.; Gaal, L.V.; Vercammen, C.; Warren, M.L.; Weinstein, D.L.; Weiss, D.; White, A.; Winnie, M.; Wium, C.; Yavuz, D. Efficacy and safety of oral semaglutide with flexible dose adjustment versus sitagliptin in type 2 diabetes (PIONEER 7): a multicentre, open-label, randomised, phase 3a trial. *The Lancet Diabetes & Endocrinology* **2019**, *7*, 528–539.
42. Zhang, X.; Belousoff, M.; Danev, R.; Sexton, P.; Wootten, D. Semaglutide-bound Glucagon-Like Peptide-1 (GLP-1) Receptor in Complex with Gs protein, 2021. doi:10.2210/pdb7ki0/pdb.
43. Webb, B.; Sali, A. Protein Structure Modeling with MODELLER. In *Methods in Molecular Biology*; Springer US, 2020; pp. 239–255.
44. Vangone, A.; Bonvin, A.M. Contacts-based prediction of binding affinity in protein-protein complexes. *eLife* **2015**, *4*.
45. Xue, L.C.; Rodrigues, J.P.; Kastritis, P.L.; Bonvin, A.M.; Vangone, A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics* **2016**, p. btw514.
46. Li, W.; Vottevor, G. Towards a Truly General Intermolecular Binding Affinity Calculator for Drug Discovery & Design **2023**. doi:10.20944/preprints202208.0213.v2.
47. Carris, N.W.; Wallace, S.; DuCoin, C.G.; Mhaskar, R.; Stern, M.; Bunnell, B. Discontinuing semaglutide after weight loss: strategy for weight maintenance and a possible new side effect. *Canadian Journal of Physiology and Pharmacology* **2024**. doi:10.1139/cjpp-2023-0464.
48. Trujillo, J.M.; Nuffer, W.; Smith, B.A. GLP-1 receptor agonists: an updated review of head-to-head clinical studies. *Therapeutic Advances in Endocrinology and Metabolism* **2021**, *12*, 204201882199732. doi:10.1177/2042018821997320.
49. He, Y.; Su, J.; Lan, B.; Gao, Y.; Zhao, J. Targeting off-target effects: endoplasmic reticulum stress and autophagy as effective strategies to enhance temozolomide treatment. *OncoTargets and Therapy* **2019**, *Volume 12*, 1857–1865. doi:10.2147/ott.s194770.
50. Smits, M.M.; Van Raalte, D.H. Safety of Semaglutide. *Frontiers in Endocrinology* **2021**, *12*. doi:10.3389/fendo.2021.645563.

 doi:10.20944/preprints202405.0258.v1

51. Shetty, R.; Basheer, F.T.; Poojari, P.G.; Thunga, G.; Chandran, V.P.; Acharya, L.D. Adverse drug reactions of GLP-1 agonists: A systematic review of case reports. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* **2022**, *16*, 102427. doi:10.1016/j.dsx.2022.102427.
52. Li, W. Towards a General Intermolecular Binding Affinity Calculator **2022**.
53. Li, W. High-Throughput Extraction of Interfacial Electrostatic Features from GLP-1-GLP-1R Complex Structures: A GLP-1-GLP-1R-Based Mini GIBAC Perspective **2024**. doi:10.20944/preprints202402.1519.v1.
54. Li, W.; Shi, G. How $Ca_V1.2$-bound verapamil blocks Ca^{2+} influx into cardiomyocyte: Atomic level views. *Pharmacological Research* **2019**, *139*, 153–157.