

A Non-Perturbative Molecular Grafting Strategy for Stable and Potent Therapeutic Peptide Ligands

Kathleen M. Sicinski, Vittorio Montanari, Venkata S. Raman, Jamie R. Doyle, Benjamin N. Harwood, Yi Chi Song, Micaella P. Fagan, Maribel Rios, David R. Haines, Alan S. Kopin, Martin Beinborn,* and Krishna Kumar*



Cite This: *ACS Cent. Sci.* 2021, 7, 454–466



Read Online

ACCESS |



Metrics & More

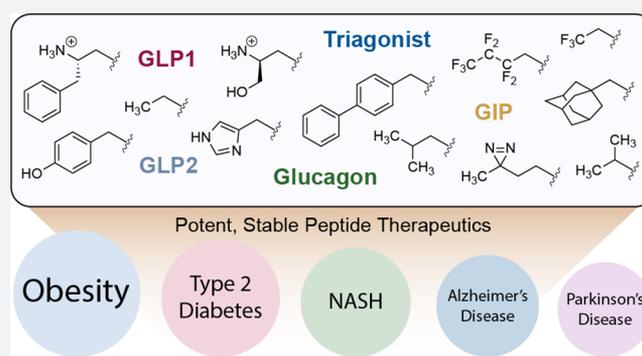


Article Recommendations



Supporting Information

ABSTRACT: The gut-derived incretin hormone, glucagon-like peptide-1 (GLP1), plays an important physiological role in attenuating post-prandial blood glucose excursions in part by amplifying pancreatic insulin secretion. Native GLP1 is rapidly degraded by the serine protease, dipeptidyl peptidase-4 (DPP4); however, enzyme-resistant analogues of this 30-amino-acid peptide provide an effective therapy for type 2 diabetes (T2D) and can curb obesity via complementary functions in the brain. In addition to its medical relevance, the incretin system provides a fertile arena for exploring how to better separate agonist function at cognate receptors versus susceptibility of peptides to DPP4-induced degradation. We have discovered that novel chemical decorations can make GLP1 and its analogues completely DPP4 resistant while fully preserving GLP1 receptor activity. This strategy is also applicable to other therapeutic ligands, namely, glucose-dependent insulinotropic polypeptide (GIP), glucagon, and glucagon-like peptide-2 (GLP2), targeting the secretin family of receptors. The versatility of the approach offers hundreds of active compounds based on any template that target these receptors. These observations should allow for rapid optimization of pharmacological properties and because the appendages are in a position crucial to receptor stimulation, they proffer the possibility of conferring “biased” signaling and in turn minimizing side effects.



INTRODUCTION

Peptides derived from the processing of the polypeptide preproglucagon define the axis of a synchronized, choreographed system of maintaining glucose homeostasis and regulating other aspects of physiological function.¹ In particular, three distinct peptides, glucagon-like peptide-1 (GLP-1; referred to as GLP1 in this paper), glucagon-like peptide-2 (GLP2), and glucagon are processed from the precursor preproglucagon into mature functioning entities and play distinct roles in glucose metabolism.² An additional peptide, produced by K-cells in the GI tract, glucose-dependent insulinotropic polypeptide (GIP), complements this axis. GLP1 is produced in enteroendocrine cells of the gut, and processed to the active form, GLP1(7–36)NH₂. Upon food ingestion, active GLP1 is secreted into the bloodstream and initiates multifaceted actions via activation of its cognate GLP1 receptor (GLP1R) expressed in several target tissues that synergistically work to restore normoglycemia.³ GLP1R stimulation triggers cAMP production and amplification of glucose-induced insulin secretion from pancreatic β -cells resulting in lowered blood glucose levels. This action also results in diminution of glucagon secretion and β -cell proliferation. In parallel, this peptide hormone delays gastric

emptying through vagal afferents and suppresses appetite via neurons in the paraventricular nucleus of the hypothalamus.⁴ Based in large part on the function of endogenous GLP1, physiologic glycemia is markedly lower after an oral glucose load compared to an equivalent parenteral load—this phenomenological observation has been termed the “incretin” effect.⁵ Because insulin secretion is glucose-dependent, this process is self-regulated, and the risk of hypoglycemia remains low.¹

The efficacy of GLP1 as a treatment for type 2 diabetes (T2D) has long been validated in human subjects, as infusion of exogenous GLP1 to diabetic patients results in near-normalization of blood sugar levels.⁶ However, native GLP1 cannot be used as a therapeutic, as it is rapidly inactivated by the ubiquitous serine protease dipeptidyl peptidase-4 (DPP4), which cleaves the N-terminal residues His7-Ala8 from the

Received: September 16, 2020

Published: February 15, 2021



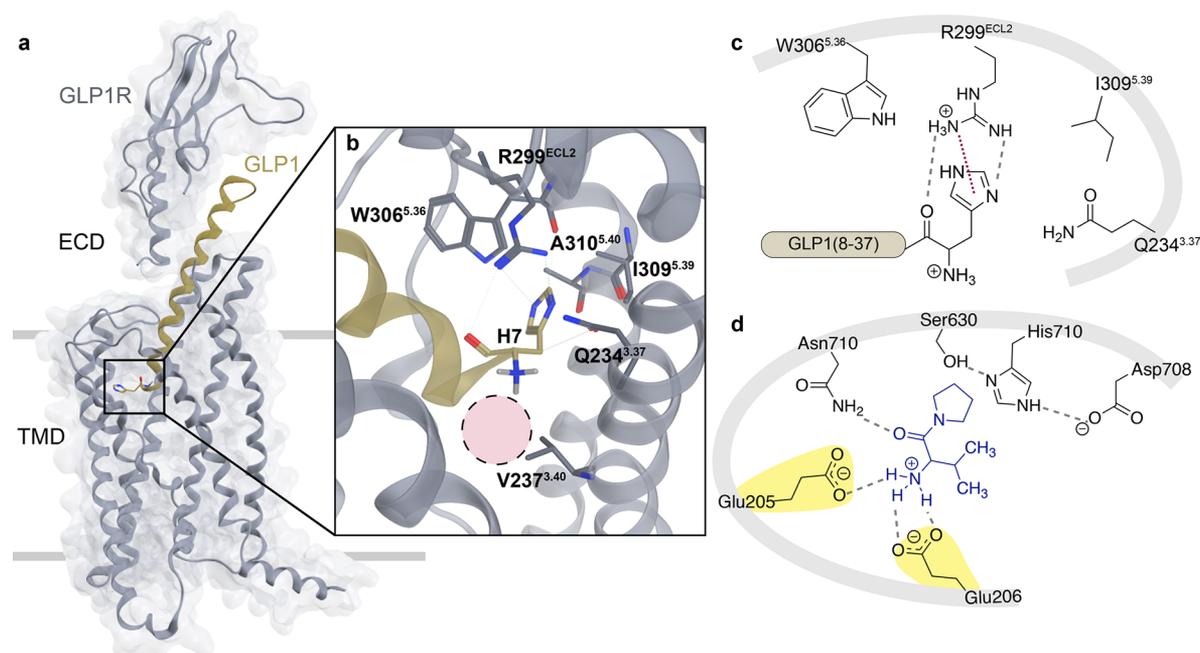


Figure 1. (a) Cryo-EM structure of the GLP1:GLP1R complex (PDB: 5VAI).¹⁰ GLP1 (ribbon, gold) bound to cognate G protein-coupled receptor, GLP1R (ribbon and surface, gray), with solid gray lines representing approximate locations of the cellular membrane that separate the extracellular domain (ECD) from the transmembrane domain (TMD) of GLP1R. (b) Illustration of receptor amino acids within 4 Å of the *N*-terminal histidine of GLP1. Residue numbers denote the Wootten nomenclature¹¹ for class B GPCRs. To note, human GLP1R contains R310^{5,40}, but this PDB structure contains an alanine mutation at this position. The dashed pink circle indicates approximate space where *N*-terminal decorations may reside. (c) Flattened 2D rendering of the interactions of GLP1R with His7 of GLP1. R299^{ECL2} forms two hydrogen bonds with the *N*-terminal histidine (dashed line, gray) and a putative cation- π interaction of the guanidine group of R299^{ECL2} with the imidazole of His7 of GLP1 (dotted line, maroon). Select neighboring side chains of the receptor are shown at approximate positions relative to GLP1. (d) 2D depiction of DPP4 active site, with a known inhibitor bound (valine-pyrrolidide, blue, PDB: 1N1M). Important electrostatic interactions (dashed lines, gray) occur between the primary amine of the substrate and Glu205 and Glu206 of DPP4 (highlighted yellow). The carbonyl of the first amide bond is anchored by a hydrogen bond to Asn710 (dashed line, gray). The catalytic triad (Ser630, His710, and Asp708) forms a hydrogen bonding network (dashed lines, gray) and is positioned proximal to the pyrrole ring. If the structure bound were a peptide substrate, the labile amide bond would be located close to the canonical nucleophile Ser630.

intact peptide ($t_{1/2} \sim 90$ s) resulting in a shorter inactive fragment, GLP1(9–36)NH₂.⁷ This observation has motivated development of GLP1 based T2D therapeutics with enhanced protease resistance and delayed renal clearance. The first GLP1 mimetic clinically approved was exenatide, an equipotent reptilian homologue that shows 50% amino acid similarity with human GLP1.⁸ Success of glycemic control by exenatide inspired the development of liraglutide, a lipidated form of GLP1.⁹ Liraglutide forms a heptamer in solution¹² and also binds the abundant plasma protein albumin resulting in delayed renal clearance.¹³ These compounds require twice- and once-daily injections, respectively. Further development has led to a derivative that is attached to the Fc fragment of IgG4 (dulaglutide)¹⁴ and another that uses addition of a lipid to enhance albumin binding along with substitution of aminoisobutyric acid (Aib) at position two (semaglutide)^{9,15} rendering these two compounds functional *in vivo* a week after injection. The future of this class of compounds has been further cemented by new guidance from the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD) which recommends GLP1R agonists as the first injectable before insulin for T2D.¹⁶ Small molecule inhibitors of DPP4 have also entered clinical use,¹⁷ although these compounds do not result in weight loss, and the long-term safety profile is still uncertain as there are more than 30 known endogenous substrates of DPP4.¹⁸ Recent efforts to overcome the liability of enzyme catalyzed hydrolysis have seen

widespread use of Aib in position 8, but a previous compound (taspoglutide)¹⁹ containing this noncanonical amino acid was withdrawn from Phase III clinical trials due to allergic reactions at the injection site.²⁰ Taspoglutide contains two Aib residues, at positions 8 and 35, to provide protection from proteases. Given that compounds in the clinic (e.g., semaglutide) contain Aib at position 8, and have good safety profiles, it is unclear whether the drug substance itself or the pharmaceutical formulation is responsible for the adverse reaction in the case of taspoglutide.²¹

Prior attempts at making GLP1 and related peptides more stable to proteolysis have relied on a number of strategies. The use of an α/β peptide scaffold that changes the backbone of the construct,²² incorporation of fluorinated amino acids at strategic positions,²³ modification of side chains with saccharides,²⁴ and the use of aminoisobutyric acid (Aib) in position 8 (second from the *N*-terminus)²⁵ have been the most fruitful. While these strategies provide proteolytic stability to various degrees, previous attempts at modifying the sacrosanct *N*-terminal histidine residue, crucial to receptor activation, have led to greatly compromised GLP1 function.²⁶

We describe in this report that *N*-alkyl modifications of several peptide ligands (GLP1, GIP, glucagon, GLP2, and a “triagonist” peptide) of the secretin family of receptors are refractory to DPP4 proteolysis while simultaneously maintaining full activity at respective cognate receptors. Because the strategy is applicable to all developed peptide ligand templates,

modulation of activity and retaining stability is possible with *N*-terminal alkylation. Given the large number of low-molecular-weight aldehydes (estd ~25,000) commercially available and the ease of installation, the number of derivatives for each peptide template to be useful clinically can easily number in the hundreds providing a blueprint for pharmacokinetic and pharmacodynamic optimization. In addition, because the *N*-terminus of the peptide ligand interacts deep in the membrane embedded region of the receptor, so as to say in the “belly of the beast”, it provides an avenue to select constructs that may trigger “biased” signaling resulting in fewer side effects.^{11,27} We note that Aib residues commonly employed to protect from DPP4 cleavage are used internally within the sequence, and do not confer ability to interact with the hitherto undescribed ligand binding pocket.

RESULTS AND DISCUSSION

Peptides are inherently susceptible to proteolytic cleavage, and this deficit has usually been addressed by introduction of unnatural or noncanonical amino acids, by backbone or side chain modification, or by addition of exogenous groups that provide serum stability by binding to plasma components. In the case of GLP1, it has been challenging to decouple enzyme mediated hydrolysis from receptor activation. The ubiquitous serine protease DPP4 removes the *N*-terminal dipeptide His7Ala8 of GLP1 with a short half-life ($t_{1/2} < 2$ min), resulting in a truncated peptide, GLP1(9–37)NH₂, with greatly diminished efficacy and potency. The active site of DPP4 has an arrangement of functionality that positions two glutamates for substrate recognition (Glu205 and Glu206), and mutating either or both glutamate(s) to alanine(s) renders the enzyme incompetent at catalysis (Figure 1d). The carboxylate groups from Glu205/Glu206 form electrostatic interactions between the positively charged amino terminus of the substrate and position the labile amide proximal to the catalytic hydroxyl of Ser630 that is primed for nucleophilic attack. Previous studies have focused on amino acid alterations of GLP1 to increase stability toward DPP4, but this strategy has usually come at the expense of GLP1R potency and/or efficacy. The loss in receptor activity underscores the importance of His7 in GLP1. Therefore, instead of altering His7, we first focused our attention on abolishing the *N*-terminal charge in GLP1 to prevent enzyme recognition.

Acylation of the *N*-Terminus of GLP1 Provides DPP4 Protection but Diminishes Receptor Activity. An obvious strategy to overcome GLP1 degradation by DPP4 is removal of the *N*-terminal charge, crucial to substrate recognition, via acylation, as in 1-GLP1 (Figure 1). Prior reports have established that 1-GLP1 is resistant to DPP4 proteolysis; however, it results in a 60- to 97-fold decrease^{28,29} in potency at the receptor. In our hands, 1-GLP1 stimulated GLP1R with an EC₅₀ of 116 pM, a 58-fold loss in potency compared to native GLP1 (EC₅₀ = 2.5 pM; Table 1, Figure S2). This result is consistent with other examples of *N*-acylation of GLP1, for example, with pyroglutamyl²⁶ or hexanoyl²⁸ groups, that show vastly reduced affinity and activation of GLP1R. All acyl groups suffer from this fate, as the receptor has evolved to recognize substrates that are processed via proteolytic cleavage of the precursor to reveal the canonical peptide of correct length, which thus lacks an acyl (amide) bond at the *N*-terminus. Processing in intestinal L-cells of a larger polypeptide, proglucagon, first results in production of GLP1(1–37) which is further cleaved to active fragments GLP1(7–37) or

Table 1. Receptor Activation of *N*^α-Modified Peptides^a

Peptide ^b	pEC ₅₀ ± SEM ^c	EC ₅₀ (pM) ^c	<i>n</i> ^d
GLP1R			
GLP1	11.6 ± 0.1	2.5	3
1-GLP1 ^e	10.0 ± 0.2	115.9	3
2-GLP1	11.5 ± 0.1	2.8	3
3-GLP1	11.2 ± 0.1	6.7	3
4-GLP1	11.4 ± 0.3	6.9	3
5-GLP1	11.4 ± 0.07	3.8	3
6-GLP1	12.0 ± 0.2	1.1	2
7-GLP1	11.3 ± 0.2	6.5	3
8-GLP1	10.8 ± 0.2	17.6	2
9-GLP1	11.7 ± 0.4	2.8	3
10-GLP1	11.4 ± 0.02	4.4	3
11-GLP1	11.6 ± 0.7	4.7	3
12-GLP1	11.6 ± 0.1	2.7	2
13-GLP1	11.3 ± 0.2	5.3	2
14-GLP1	10.6 ± 0.1	25.3	2
15-GLP1	11.4 ± 0.1	3.8	2
16-GLP1	11.0 ± 0.03	9.6	2
17-GLP1	10.4 ± 0.05	41.6	3
18-GLP1	11.1 ± 0.06	8.4	4
19-GLP1	7.5 ± 0.1	3.3 × 10 ⁴	2
Liraglutide	11.6 ± 0.2	3.2	3
2-Liraglutide	11.4 ± 0.03	4.0	3
Exenatide	11.6 ± 0.02	2.6	3
2-Exenatide	11.6 ± 0.06	2.9	3
2-Triagonist(Ala2)	11.2	6.9	1
GIP	6.92	1.2 × 10 ⁵	1
Glucagon	9.31	490	1
GIPR			
GIP	12.2 ± 0.04	0.6	3
1-GIP	10.3 ± 0.2	54.6	2
2-GIP	12.6 ± 0.02	0.3	3
Glucagon	8.07	8.6 × 10 ³	1
GLP1	8.30	5.0 × 10 ³	1
2-Triagonist(Ala2)	11.5	3.3	1
GCGR			
Glucagon	11.5 ± 0.04	3.4	5
2-Glucagon	11.6 ± 0.1	2.3	5
7-Glucagon	9.49 ± 0.1	325	2
GLP1	6.18	6.6 × 10 ⁵	1
GIP	6.30	5.1 × 10 ⁵	1
2-Triagonist(Ala2)	11.4	4.5	1
GLP2R			
GLP2	10.7 ± 0.1	21	3
2-GLP2	11.2 ± 0.1	7	3

^aPotency of synthesized peptides using HEK293 cells expressing GLP1R, GIPR, GCGR, or GLP2R and luciferase reporter system. Results are separated by target receptors. ^bSee Figure 2 for peptide structures. ^cEC₅₀ is the concentration of peptide required for half-maximal activity of the targeted receptor. pEC₅₀ = $-\log(\text{EC}_{50}) \pm$ standard error of the mean (SEM) of independent experiments where applicable. ^dNumber of independent experiments that were run in triplicate or quadruplicate. ^ePeptides were incubated at 37 °C overnight before stimulating transfected cells.

GLP1(7–36)NH₂. Improperly processed fragments such as GLP1(1–37), GLP1(8–37), and GLP1(6–37) are unable to

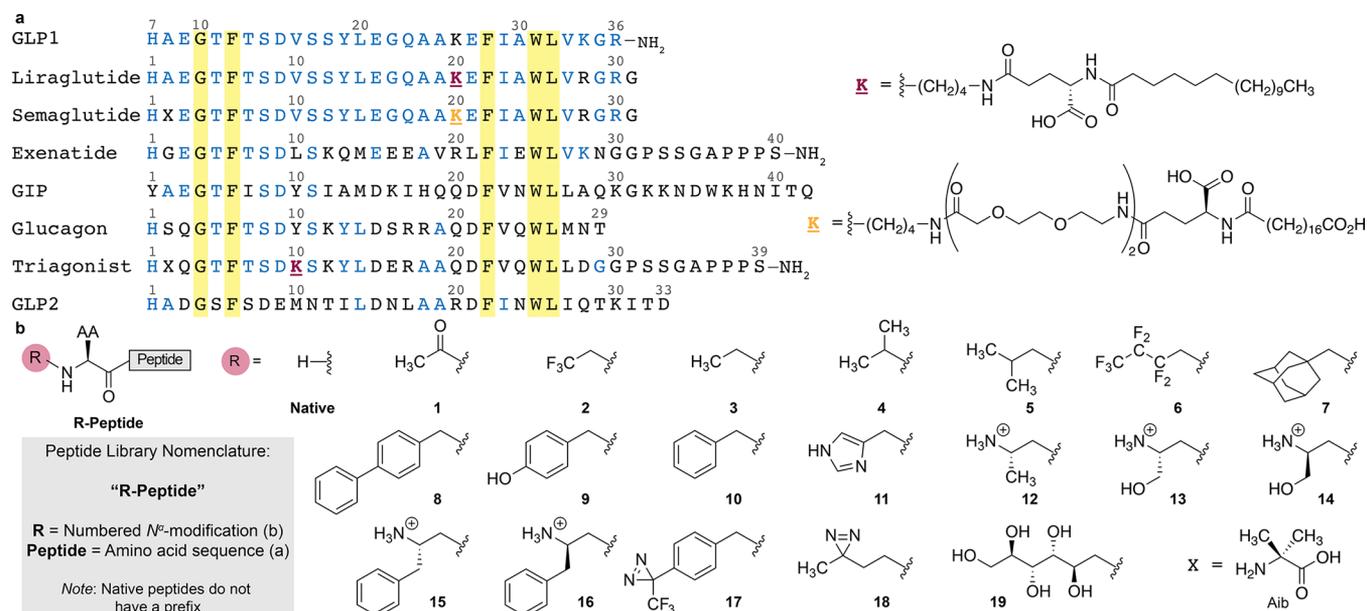


Figure 2. Library of *N*-terminally modified peptides. (a) Alignment of GLP1 and related peptides with positions and numbering above each residue (gray). GLP1 starts with amino acid 7 based on established convention.³ Blue residues are homologous to GLP1, and residues highlighted yellow are conserved between all peptides. Liraglutide and triagonist contain a lysine (**K**, maroon) modified with a γ -glutamic acid spacer and by palmitoylation (right). Semaglutide contains a modified lysine (**K**, orange) with two oligoethylene glycol (OEG) spacers, γ -glutamic acid, and octadecanedioic acid (right). "X" denotes the noncanonical amino acid, aminoisobutyric acid (Aib, bottom right). (b) Native amino acid sequences are modified with *N*-terminal chemical modifications 1–19 resulting in a library of peptides, nominally "R-Peptide" where "R" is the number referencing the *N*-terminus modification and "Peptide" indicates the template sequence as in (a). Semaglutide and triagonist peptides were also assembled with Aib2Ala mutation denoted as R-semaglutide(Ala2) and R-triagonist(Ala2), respectively.

induce insulin secretion at normal levels.²⁹ It is only once GLP1 is proteolyzed to display the free *N*-terminal histidine in GLP1(7–37) or GLP1(7–36)NH₂ that native insulin release action is observed.³⁰

We show here that removal of charge by alkylation (for example, with a trifluoroethyl group) is a viable strategy. Since this modification does not introduce an acyl (amide) bond at the *N*-terminus, it is tolerated by the receptor. A second approach is to increase steric bulk at the *N*-terminus so that the binding of the free amine in the enzyme active site is compromised. Surprisingly, the receptor is again accommodating of a large variety of such bulky groups introduced as *N*-alkyl tethers.

We then investigated alternative strategies that are capable of altering the p*K*_a of the *N*-terminal amine. The cryo-EM structure of a GLP1–GLP1R complex (PDB: SVAI)¹⁰ provides a basis to inform which His7 modifications on the ligand could be accommodated within the membrane embedded region of the receptor (Figure 1a). Within the binding pocket of GLP1R, there are receptor side chains that make key noncovalent interactions with His7 of GLP1 (Figures 1b, S1A). Notably, R299^{ECL2} is involved in a π -cation interaction with the aromatic imidazole ring of His7, as well as W306^{S.36} and I309^{S.39} that form hydrogen bonds, but the primary *N*-terminal amine of GLP1 in this structure is devoid of contacts with the receptor (Figure 1c). The only residue within 4 Å of the free amine is V237^{S.40}, and this hydrophobic residue is not proximal enough to GLP1 to contribute toward intermolecular interactions between receptor and ligand. Although this amine is not perceived to be interacting with GLP1R, removal of the primary amine results in a 15-fold loss in potency illustrating the importance of the amino terminus.³¹ Not only has the free positive amino terminus been determined to be

crucial, but the methyl imidazole side chain of His7 has also been recalcitrant to modification efforts. For example, substitution of His7 with Ala, Arg, Lys, or Tyr results in losses in both potency (174- to 1280-fold) and efficacy (12–66% of native peptide).³² Further interrogation with six analogues that differ in the *number* and *position* of nitrogens in the methyl imidazole side chain revealed potency losses from 2.1-fold (H_{2,2}-GLP1(8–36), Figure S1B) to 475-fold (H_{3,1}-GLP1(8–36), Figure S1B) without a significant improvement in DPP4 resistance. Further, a His7Tyr modification leads both to decrease in receptor affinity (10-fold) and EC₅₀ values of stimulatory activity (13-fold).²⁸ These observations serve as a foundation to make chemical modifications at this location to prevent DPP4 proteolysis while simultaneously retaining receptor efficacy and potency.

We hypothesized that modification of the amino terminus via alkylation through the agency of a trifluoroethyl group would afford an equipotent and equiefficacious GLP1 analogue that is impervious to DPP4 catalyzed inactivation. We envisioned that proteolytic stability would result from the electron withdrawing character and steric bulk of the 2,2,2-trifluoroethyl group. Alkylation of the α -amine of glycine with the trifluoroethyl moiety reduces the p*K*_a of the ammonium species³³ from 9.6 to 5.3, and we envisioned that such a modification would result in removal of charge and simultaneously eliminate the acyl functionality at the *N*-terminus that GLP1R recognizes as an improperly processed ligand.

***N*-Trifluoroethyl Modification of GLP1 and of Related Receptor Agonists Result in Maintenance of Receptor Activity.** We assessed the potency and efficacy of 2-GLP1 (Figure 2) at GLP1R by transfecting HEK293 cells with three cDNA constructs that encode for the target G protein-coupled

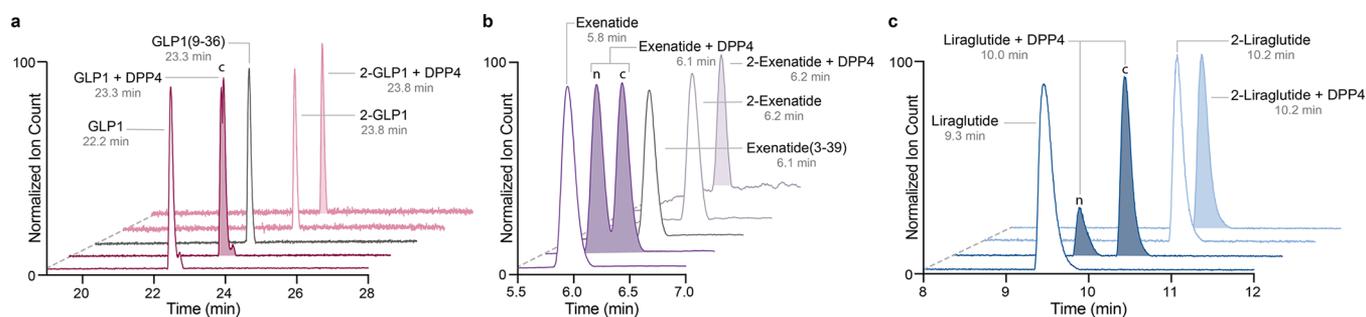


Figure 3. (a) LC-MS/MS total ion chromatogram depicting the stability of GLP1 (maroon) and 2-GLP1 (pink) with (shaded) and without (nonshaded) DPP4. GLP1 incubated with DPP4 shows the same retention time as control GLP1(9–36) (gray) indicating excision of dipeptide His7Ala8 to give cleaved, c, peptide. 2-GLP1 incubated with DPP4 exhibits no change in retention time or mass. (b) LC-MS/MS total ion chromatogram depicting the stability of exenatide (purple) and 2-exenatide (light purple) with (shaded) and without DPP4 (nonshaded). Exenatide incubated with DPP4 results in a mixture of cleaved, c, and native (unreacted), n, exenatide. The retention time of cleaved exenatide, c, is the same as control exenatide(3–39) (gray). 2-Exenatide incubated with DPP4 is unreactive with no change in retention time or mass. (c) LC-MS/MS total ion chromatogram depicting the stability of liraglutide (navy) and 2-liraglutide (light blue) with (shaded) and without DPP4 (nonshaded). Liraglutide incubated with DPP4 results in a mixture of cleaved, c, and native (unreacted), n, liraglutide. 2-Liraglutide incubated with DPP4 does not undergo reaction with unchanged retention time and mass.

receptor, GLP1R, CRE_{6x}-luciferase reporter, and β -galactosidase to determine transfection efficiency. After 24 h, 2-GLP1 was titrated and the amount of luciferase produced after 4–6 h was used to calculate EC₅₀ values, that were found to be comparable to native GLP1 (Table 1). This was a surprising and remarkable discovery given that the free amino terminus of GLP1 has long been considered crucial for receptor binding and activation.²⁸ Additionally, 2-GLP1 was entirely refractory to DPP4 catalyzed hydrolysis (*vide infra*). To the best of our knowledge, this is the first instance that an *N*-terminal modification of GLP1 has resulted in an equipotent and equally efficacious GLP1R agonist.

GLP1R agonists currently used clinically (liraglutide, exenatide, and semaglutide) operate by similar modes of binding and activation; we therefore evaluated the effect of *N*-terminal alkylation on these peptide ligands. For these GLP1R agonists, the potency and efficacy toward the receptor were not altered. The potency of 2-exenatide (2.9 pM) was essentially unchanged from that of exenatide (Figure 2, Table 1). A similar trend was seen for 2-liraglutide and 2-semaglutide-(Ala2) (Figure 2, Table 1, Figure S3). The fact that this same *N*-terminal modification is useful in multiple GLP1R agonists confirms the robustness of this chemical grafting method and fueled our investigation to probe for stability against DPP4 action.

LC-MS Analysis of DPP4 Stability for GLP1R Agonists.

After having determined that *N*-trifluoroethyl histidine modified GLP1 is accommodated by GLP1R, we investigated if the change in the *N*-terminal p*K*_a influenced DPP4 proteolysis.³³ Native GLP1 or 2-GLP1 was incubated with and without DPP4 at 37 °C and the reaction mixture analyzed by ESI LC-MS after 16 h. In the presence of DPP4, native GLP1 was cleaved resulting in a distinct shift in retention time, from 22.2 to 23.3 min, with an accompanying reduction in mass by 208.1 Da indicating removal of the His7Ala8 dipeptide from the *N*-terminus. The product of GLP1 incubated with DPP4 eluted at the same time as the control GLP1(9–37)NH₂, further confirming that the product resulted from dipeptide cleavage (Figure 3a). When 2-GLP1 was incubated with DPP4 overnight, no change in retention time or mass was observed, illustrating that the *N*-trifluoroethyl alkylation of GLP1 prevented DPP4 hydrolysis likely due to both decrease

in p*K*_a and addition of steric bulk on the *N*-terminal amine. This modification abolishes the interaction between the amino terminus of 2-GLP1 and Glu205 and Glu206 in the DPP4 active site, rendering it refractory to enzyme recognition and hydrolysis. We also explored whether the modified peptides are inhibitors of the enzyme or are just plainly not recognized as substrates. Steady-state kinetics with a known inhibitor (linagliptin) and substrate (GlyPro-pNA, Figure S4) reveal that these compounds are neither inhibitors (no reduction in *K*_m) nor bind to the enzyme (no change in *V*_{max}). These data suggest that either removal of charge or addition of bulk at the *N*-terminus of GLP1 abrogates recognition by DPP4, the frontline protease.

Exenatide is more resistant to DPP4 than GLP1; however, it is still susceptible to enzymatic degradation as judged by LC-MS analysis. After overnight incubation with DPP4, exenatide was partially cleaved resulting in a mixture containing an additional product (*t*_R = 6.1 min). The elution profile and identified mass (−194.1 g/mol) corresponded to the peptide fragment exenatide(3–39) (Figure 3b). When 2-exenatide was subjected to DPP4 incubation, no cleavage products were observed. Our results can be explained with the same rationale as 2-GLP1, where recognition of the substrate and cleavage are prevented by removal of charge and addition of steric bulk at the *N*-terminus.

Liraglutide contains the same *N*-terminal motif as GLP1, and in the absence of serum albumin, it is susceptible to proteolysis. When we incubated liraglutide with DPP4 overnight, we saw MS signatures of two unique compounds by LC-MS (Figure 3c). The first compound eluted at the same time (*t*_R = 9.3 min) and had the same MW as liraglutide. The second compound eluted later and had a mass that was 208.1 Da lower than liraglutide, suggestive of DPP4 removal of the HisAla dipeptide. When the *N*-terminally modified peptide 2-liraglutide was incubated with DPP4 under similar conditions, no change in mass or retention time was observed indicating the failure of DPP4 to process and inactivate substrate.

***N*-Trifluoroethyl Modification Is Accommodated by Most Secretin Family Receptors.** Three peptide ligands—glucagon, GIP, and GLP2—fall within the same secretin family as GLP1 and contain five conserved amino acids: Gly4, Thr6, Phe22, Trp31, and Leu32 (Figure 2). Each peptide elicits

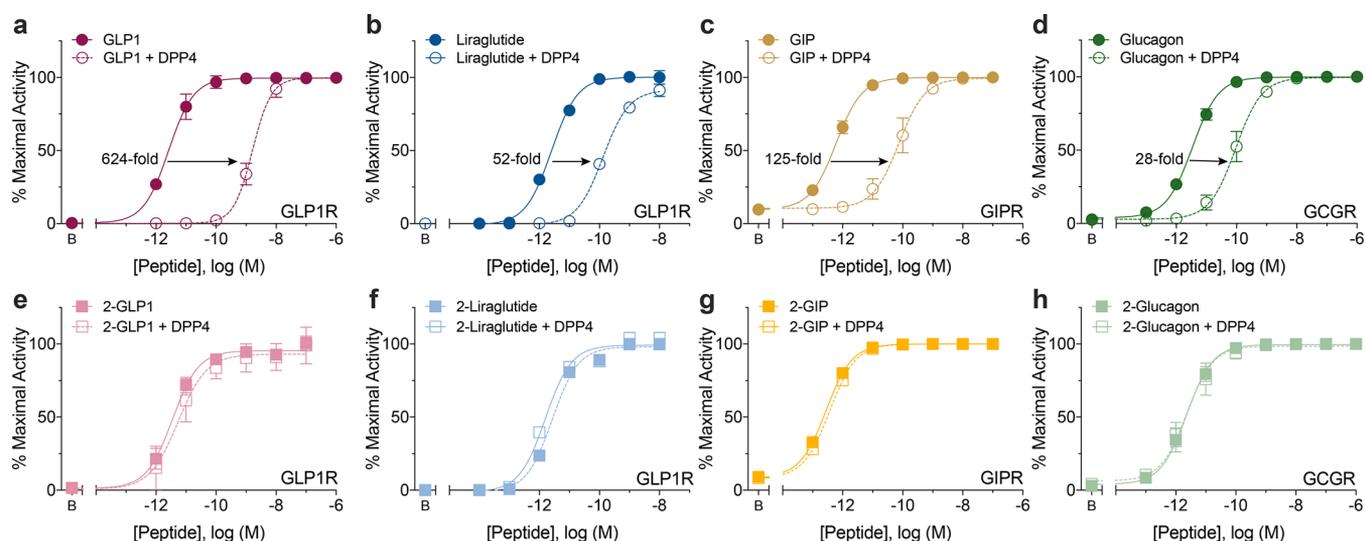


Figure 4. Representative concentration–response curves of unmodified peptides (GLP1, liraglutide, GIP, and glucagon; a–d) or *N*-trifluoroethyl analogues (2-GLP1, 2-liraglutide, 2-GIP, and 2-glucagon; e–h) incubated overnight with DPP4 or vehicle prior to diluting into microtiter plates containing HEK293 cells overexpressing receptors (GLP1R, GIPR, or GCGR) and reporter CRE_{6x}-luciferase. Luciferase production corresponds directly to activation of cognate GPCR via a cAMP dependent pathway, normalized to 100% maximal activity, and resultant fold-loss in EC₅₀ upon DPP4 incubation is listed when applicable. Error bars represent SEM for three independent experiments (*n* = 3).

various physiological responses from increasing blood glucose to stimulating intestinal growth. Although the first three *N*-terminal amino acids are dissimilar, all peptides are cleaved by DPP4. We sought to ascertain if fluoroalkylation would be compatible with the cognate receptors for glucagon, GIP, and GLP2 secretin peptides.

Glucose-dependent insulinotropic polypeptide (GIP), like GLP1, is a gut hormone, and functionally complements the latter in accounting for the incretin effect after food ingestion.³⁴ GIP works through a different class B receptor (GIPR)^{35,36} and has overlapping (amplification of insulin release from pancreatic β -cells) as well as distinct (modulation of fat cell metabolism) functions compared to GLP1.^{7,5} In the diabetic state, the function of GIP is greatly diminished,³⁷ possibly as a result of GIPR down-regulation.^{38,39} However, there is now increasing evidence that this defect is reversible once chronic hyperglycemia is alleviated,⁴⁰ e.g., after treatment with GLP1 analogues. There are some sequence similarities between GLP1 and GIP, but the *N*-terminus is occupied by a Tyr instead of His and this change provides specificity at GIPR. We synthesized an α -amino-modified tyrosine to incorporate into the full-length peptides yielding 2-GIP. Again, utilizing a luciferase-based assay, we transfected HEK293 cells with cDNAs encoding for the receptor of interest, GIPR, CRE_{6x}-luciferase reporter, and a β -galactosidase control.²³ The day following transfection, cells were agonized with native GIP or 2-GIP and production of luciferase was quantified. 2-GIP showed similar potency as native GIP at the cognate receptor, GIPR (EC₅₀ = 0.6 pM and 0.3 pM, respectively, Table 1). Similar to GLP1, this small molecular modification resulted in a GIP analogue with complete proteolytic stability without compromising receptor stimulatory activity (*vide infra*). The minimal change in potency and efficacy of 2-GIP cements the view that this molecular grafting method by alkylation is suitable for use at multiple receptors of the secretin family.

Glucagon originates from pancreatic islets and, in contrast to the incretins, elevates glycemia by inducing hepatic glucone-

genesis and mobilization of glucose into blood. However, during balanced pharmacological co-stimulation of GLP1R, GIPR, and glucagon receptors with low concentrations of corresponding agonists, glucagon further amplifies the anorectic effect of the incretins by enhancing energy expenditure via receptors on adipocytes and in the brain, thereby further promoting a loss in body weight.^{1,41} The glycemic liability of glucagon disappears in combination drug regimens with GLP1 and GIP mimetics where the latter are dominant in reducing blood glucose. A recently engineered GLP1/GIP/glucagon triple agonist was highly effective in treating diabetes and obesity in mice.⁴² Akin to GIP and GLP1, glucagon is also a substrate for DPP4 albeit suffers hydrolysis at a slower rate ($t_{1/2}$ = 5–6 min).^{43,44} Glucagon possesses the same *N*-terminal HisXaa motif as GLP1 and the *N*-terminal histidine was previously found to be essential for receptor activation.⁴⁵ We synthesized glucagon with the *N*-trifluoroethyl modification, 2-glucagon, and tested its ability to bind and activate its cognate glucagon receptor (GCGR). Consistent with our findings above, the addition of the *N*-terminal modification did not influence potency—2-glucagon was slightly more potent than native glucagon, with an EC₅₀ of 2.3 pM compared to 3.4 pM of native glucagon in cellular assays (Table 1).

Glucagon-like peptide-2 (GLP2) is an agonist of another secretin family receptor, GLP2R. The 33-residue peptide is cosecreted with GLP1 and is a mediator of mucosal proliferation and enhances the activity of several absorptive enzymes resulting in optimal intestinal uptake of nutrients.^{46,47} We synthesized an *N*-terminally modified GLP2 analogue to determine if this trifluoroethyl decoration would also be useful at its cognate receptor GLP2R. After synthesizing 2-GLP2, cells overexpressing GLP2R were treated with 2-GLP2 or native GLP2. We discovered that the fluorinated derivative 2-GLP2 was more potent than native GLP2 with EC₅₀ of 7 pM compared to 21 pM of native GLP2 and was completely refractory to DPP4 catalyzed proteolysis. This was the most impressive improvement in potency of a secretin peptide that

Table 2. Stability towards DPP4 Proteolysis of N^α-Modified Peptides^a

Peptide ^b	-DPP4			+DPP4			Fold-shift (↓) ^e
	pEC ₅₀ ± SEM ^c	EC ₅₀ (pM) ^c	n ^d	pEC ₅₀ ± SEM ^c	EC ₅₀ (pM) ^c	n ^d	
GLP1R							
GLP1	11.6 ± 0.2	3.5	5	8.83 ± 0.3	2.8 × 10 ³	5	800.0
1-GLP1	9.98 ± 0.2	99.1	3	9.93 ± 0.2	106.4	3	1.1
2-GLP1	11.5 ± 0.1	3.5	3	11.3 ± 0.1	5.4	3	1.5
3-GLP1	11.4 ± 0.2	5.5	3	11.2 ± 0.1	6.4	3	1.2
4-GLP1	11.4 ± 0.5	7.4	2	11.0 ± 0.4	17.1	2	2.3
5-GLP1	11.3 ± 0.1	4.8	3	11.3 ± 0.1	5.0	3	1.0
6-GLP1	11.8	1.55	1	11.5	3.23	1	2.1
7-GLP1	11.5 ± 0.3	3.6	2	11.5 ± 0.3	3.1	2	0.9
8-GLP1	11.1	7.2	1	10.9	12.4	1	1.7
10-GLP1	11.4 ± 0.07	4.0	3	11.4 ± 0.08	4.6	3	1.2
12-GLP1	11.2 ± 0.5	10.8	2	10.9 ± 0.3	14.0	2	1.3
19-GLP1	7.53 ± 0.09	3.0 × 10 ⁴	2	6.91 ± 0.1	1.3 × 10 ⁵	2	4.3
Liraglutide	11.2 ± 0.2	8.6	4	9.59 ± 0.1	283	4	32.9
2-Liraglutide	11.0 ± 0.3	12.8	3	11.1 ± 0.05	8.7	3	0.7
Exenatide	11.5 ± 0.1	3.3	3	11.3 ± 0.05	5.6	3	1.7
2-Exenatide	11.4 ± 0.07	4.3	3	11.4 ± 0.02	4.4	3	1.0
GIPR							
GIP	11.9 ± 0.08	1.4	3	9.77 ± 0.3	226.5	3	161.8
2-GIP	12.2 ± 0.04	0.6	3	12.1 ± 0.03	0.90	3	1.5
GCCR							
Glucagon	11.3 ± 0.1	5.9	4	9.42 ± 0.4	870	4	147.5
2-Glucagon	11.4 ± 0.2	4.9	4	11.5 ± 0.2	3.7	4	0.8
GLP2R							
GLP2	10.4 ± 0.06	39.3	4	8.40 ± 0.02	4.0 × 10 ³	3	101.8
2-GLP2	11.1 ± 0.2	9.6	2	11.1 ± 0.05	9.0	2	0.9

^aPotency of synthesized peptides using HEK293 cells expressing GLP1R, GIPR, GCCR, or GLP2R and luciferase reporter system. Results are separated by target receptors. Peptides were incubated at 37 °C overnight with and without DPP4 before incubation of transfected cells. ^bSee Figure 2 for structures of synthesized peptides. ^cEC₅₀ is the concentration of peptide required for half maximal activity of the targeted receptor. pEC₅₀ = -log(EC₅₀) ± standard error of the mean (SEM) of independent experiments where applicable. ^dNumber of independent experiments conducted in triplicate or quadruplicate. ^eCalculated by the ratio (EC₅₀ with DPP4)/(EC₅₀ without DPP4)

we observed. Taken together, we have shown that multiple peptides within the secretin family tolerate or are actually functionally enhanced by the small N-trifluoroethyl modification.

Because variations in potency was observed for alkylated secretin peptides at their cognate receptors, we then modified a peptide known to activate three of the secretin receptors (GLP1R, GIPR, and GCCR). Previous work by Finan *et al.* noticed a similarity in amino acid sequences between GLP1, GIP, and glucagon and generated a single peptide constructed with a synergistic effect on body weight reduction.⁴² This peptide, named “triagonist”, contains Aib at position two to prevent DPP4 cleavage. Our goal was to eliminate Aib and utilize our novel N-terminal modification to retain receptor stimulatory activity while simultaneously preserving DPP4 resistance. We synthesized a N-trifluoroethyl triagonist with alanine in position 2, **2-triagonist(Ala2)**. Fluoroalkylation of the triagonist resulted in potencies in the single picomolar range, remarkably close to the respective native peptide ligands at their cognate receptors (Table 1, Figure S5).

In order to examine the scope and limitation of receptors that are able to recognize ligands modified through this strategy, we focused on the δ-opioid receptor and its native

ligand, MetEnk, which can also be inactivated by DPP4.^{48,49} In order to test if the trifluoroethyl modification protects the native peptide from hydrolysis and retain activity at the receptor, the activity of **2-MetEnk** was tested at the native δ-receptor. This small modification abolished activity of the native peptide completely, most likely due to the importance of the primary amine for receptor activation (Figure S6) and thus circumscribing the scope of such alterations to the secretin family of receptors and their ligands.

N-Trifluoroethyl Alkylation Confers Superior DPP4 Stability. We have established that altering the pK_a of the N-terminal amine of GLP1 contributes to increased DPP4 stability, as judged by LC-MS. We then investigated how EC₅₀ is modulated upon incubation with DPP4, which provides a more sensitive analysis of the remaining full-length peptide. We incubated **2-GLP1** and GLP1 with DPP4 at 37 °C overnight. The peptides were then added directly to HEK293 cells overexpressing GLP1R. The receptor was stimulated for 4–6 h and production of luciferase was quantified and concentration response curves with and without DPP4 incubation were plotted. When GLP1 was incubated with DPP4 overnight, there was a shift in potency (EC₅₀) from 3 pM to 2.8 nM corresponding to a massive 800-fold shift due to proteolysis

(Figure 4a and Table 2). A significant difference in potency was also observed after incubation of liraglutide with DPP4 shifting an average of 33-fold from an EC_{50} of 8.6 pM to 283 pM (Figure 4b, Table 2). Only a minimal loss (~ 1.7 -fold) in potency of exenatide was observed after DPP4 incubation from 3.3 to 5.6 pM (Table 2).

Because DPP4 is known to cleave multiple substrates, we interrogated the protease stability of secretin peptides: GIP, glucagon, and GLP2. Peptides were preincubated with DPP4, then titrated into a microtiter plate containing HEK293 cells overexpressing the receptor of interest (GIPR, GCGR, or GLP2R). Each peptide showed diminution in receptor stimulatory activity after DPP4 incubation. The tabulated loss for GIP was 162-fold, for glucagon 147-fold, and for GLP2 102-fold (Table 2, Figure 4c,d).

After determining the loss in potency upon DPP4 incubation, we assessed if this loss could be prevented by *N*-trifluoroethyl alkylation. All fluorinated peptides 2-GLP1, 2-liraglutide, 2-exenatide, 2-GIP, 2-glucagon, and 2-GLP2 showed retention of potency and efficacy upon preincubation with DPP4. An alternative to Aib protease protection in semaglutide is made available by these results, by retaining alanine at position 2 and alkylation of the *N*-terminus to give 2-semaglutide(Ala2) (Figure S3b,c). This pointed to a new, vigorous approach to prevent DPP4 proteolysis among various peptides of the secretin family without compromising receptor stimulatory function.

***N*-Terminal Modifications on GLP1 Extend beyond $-CH_2CF_3$.** In order to establish the molecular parameters that best complement the receptor binding pocket in the membrane embedded region that is assumed to be *elastic*, we varied the alkyl tethers for various physicochemical attributes such as size (7-GLP1, 8-GLP1, 10-GLP1, and 19-GLP1), hydrophobicity (7-GLP1, 8-GLP1, and 10-GLP1), geometry (7-GLP1 and 8-GLP1), charge (12-GLP1, 13-GLP1, 14-GLP1, 15-GLP1, and 16-GLP1), polarizability (2-GLP1, 6-GLP1, 8-GLP1, 9-GLP1, 10-GLP1, and 11-GLP1), stereochemistry (12-GLP1, 13-GLP1, 14-GLP1, 15-GLP1, and 16-GLP1), and electronegativity (2-GLP1, 3-GLP1, and 6-GLP1). A polar appendage with multiple hydroxyl groups (mannitol, 19-GLP1) was also tested.

A structure activity relationship (SAR) approach was employed to explore an assortment of *N*-terminal decorations. The majority of *N*-alkyl groups were installed by reductive amination directly on resin by reaction of the α -amino group and the corresponding aldehyde.⁵⁰ After incubation of the peptide and aldehyde, the resulting imine was reduced with $NaBH_4$ to afford the alkylated product creating a structurally diverse library of amine decorations (Figure 2).

The ethyl, propyl, and isobutyl modifications of GLP1 introduce small aliphatic groups on the *N*-terminus to afford 3-GLP1, 4-GLP1, and 5-GLP1. These modifications resulted in constructs that are close in potencies to native GLP1 and were able to stimulate the receptor with full efficacy. In addition, the largest group in 5-GLP1 resulted in the most potent compound of the group suggesting that larger nonpolar groups are able to penetrate the binding pocket and activate the receptor.

After observing the number of small aliphatic groups that GLP1R is able to accommodate, we investigated the size limit of these modifications. Introduction of a larger perfluoroalkyl group as in 6-GLP1 resulted in equipotent GLP1R activity ($EC_{50} = 1.1$ pM). We placed a large hydrophobic methyl

adamantyl group on the *N*-terminus of GLP1 (7-GLP1), and surprisingly, this did not significantly diminish activation of GLP1R (~ 3 -fold loss in potency). Extension with a flat biphenyl aromatic ring system 8-GLP1 resulted in a 9-fold change in EC_{50} to 17.6 pM. Based on these observations, one can conclude that large nonpolar groups are compatible with GLP1R, but geometric attributes do play a role in receptor activation. We estimate that the receptor is able to accommodate aliphatic groups up to a volume of 130 \AA^3 .

The observation that a greater reduction in potency with 8-GLP1 containing a biphenyl group led us to remove the second geometrically flat aromatic ring to yield a benzyl functionality (10-GLP1). The potency recovered from a 9-fold loss to only a 2-fold loss in potency compared to GLP1. The methyl imidazole containing 11-GLP1 was investigated to determine its effect and influence on receptor binding and activation. *A priori*, the expectation was an increase in potency because of the crucial role the methyl imidazole side chain of His7 had for receptor binding and activation; however, no increased potency was observed. The EC_{50} remained essentially unchanged compared to native GLP1 at 3 pM.

As the positively charged amino terminus of GLP1 was previously shown to be important for receptor binding and activation, we used alanine aldehyde to yield construct 12-GLP1. This resulted in essentially no change in potency at GLP1R ($EC_{50} = 2.7$ pM). More intriguingly, stereochemical differences play an important role in receptor activation. We used reductive amination with the *R* or *S* stereoisomers of Garner's aldehyde to retain stereochemistry at the α -carbon and produce stereoisomeric constructs equivalent to those that would be obtained from the *D*-serine (13-GLP1) and *L*-serine (14-GLP1) aldehydes after cleavage from peptide resin.⁵¹ Remarkably, there was a 6.3-fold difference in potencies, 4.0 pM and 25 pM, in favor of the *S* isomer. We also conducted a similar comparison with the aldehyde of *L*-phenylalanine (15-GLP1) and *D*-phenylalanine (16-GLP1). A difference of 3-fold was observed between the two stereoisomers, with the *S* chirality favored once again. This highlights the fact that stereochemical differences in the ligand can be exploited to fine-tune receptor activation and that the binding pocket has chirality that is responsive to changes in the three-dimensional disposition of functionality.

In order to gain a full understanding of the dynamics of this pocket, we synthesized two diazirine moieties to generate a probe capable of cross-linking with the receptor. We synthesized the commonly used trifluoromethylphenyl diazirine 17-GLP1 and assessed the potency at the receptor. We observed a modestly diminished potency ($EC_{50} = 42$ pM) which fell in line with the biphenyl compound 8-GLP1, suggesting that the large geometrically flat phenyl ring cannot bear too much steric bulk and still fit within the receptor pocket. This prompted the synthesis of a smaller alkyl diazirine, 18-GLP1, resulting in a construct with significantly improved potency of 8.4 pM.

Finally, we interrogated whether the receptor is tolerant of polar uncharged functionality. We tested the mannitol derivative, 19-GLP1, for potency and efficacy. With five hydroxyl groups positioned in the pocket, the potency of this construct was massively diminished with an EC_{50} of 33 nM (13,200-fold loss) while efficacy was fully retained. This suggests that 19-GLP1 is able to bind the receptor, but the appendage is unable to penetrate the receptor binding pocket in any meaningful way.

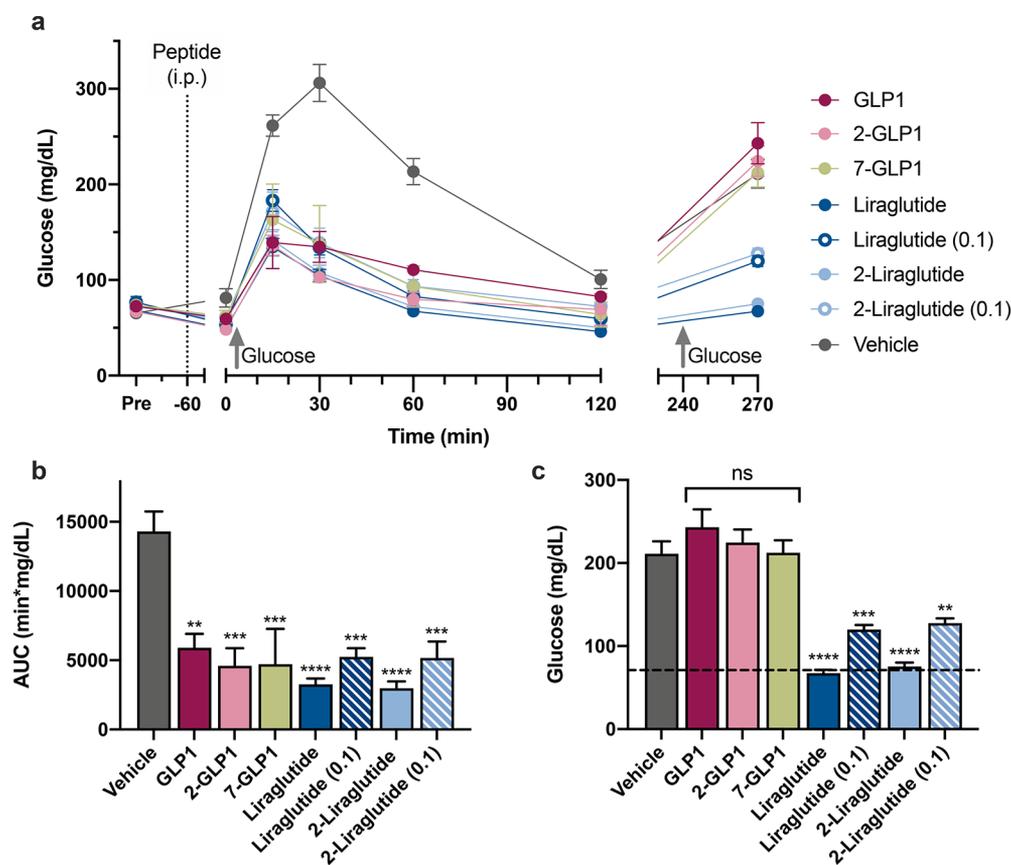


Figure 5. *N*-Trifluoroethyl alkylation and lipidation of 2-liraglutide performs as well as liraglutide at lowering blood sugar levels *in vivo*. (a) Measured glucose levels by tail vein prick for an oral glucose tolerance test (OGTT) of fasted mice treated intraperitoneally (i.p., dotted line) with vehicle, GLP1, 2-GLP1, 7-GLP1, liraglutide, or 2-liraglutide at (1 mg/kg or 0.1 mg/kg as noted). Glucose bolus was administered orally at time 0 and 240 min (gray, upward arrow). (b) Average area under the curve (AUC) calculated from 0 to 120 min in part a. (c) Glucose levels 30 min past second glucose challenge that occurred 5 h after first OGTT. Error represents the average \pm SEM ($n = 5$). *P*-values compared to vehicle: ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

Overall, these peptide constructs provide a framework for understanding the types of modification that are topologically compatible with the GLP1R binding pocket, the volume that is available, the functional groups that are complementary, and the three-dimensional chiral space that the binding region presents to ligands. These observations should help guide design of a diverse library of high potency, full efficacy peptides that are also resistant to DPP4 catalyzed hydrolysis (Table 2). It is likely that subtle changes in the interaction of these ligands with the GLP1R binding pocket will result in peptides that trigger biased signaling.

Stability Conferred by Fluoroalkylation Extends to Protease Family Related to DPP4. DPP4 belongs to a family of homologous serine proteases that are active against the same *N*-terminal motif and cleave GLP1 resulting in the inactive fragment GLP1(9–37)NH₂.⁵² We interrogated if trifluoroethylation of GLP1 increases stability toward enzymes related to DPP4 (members of the S9B prolyl oligopeptidase subfamily that have similar arrangement of active sites and recognition of substrates by two glutamate residues), namely, dipeptidyl peptidase-9 (DPP9) and fibroblast activation protein α (FAP). Peptidases DPP4, DPP9, and FAP all cleave similar substrates: GLP1, GLP2, neuropeptide Y (NPY), and peptide YY (PYY).⁵³ Prior reports have documented the ability of DPP9 in cleaving GLP1 efficiently with a half-life of 6 min⁵⁴ *in vitro* as compared FAP, which gives GLP1 a $t_{1/2}$ of 22 h.⁵³

After overnight incubation of GLP1 or 2-GLP1 peptides with vehicle, FAP, or DPP9, samples were titrated into a microtiter plate containing the luciferase reporter HEK293 cells overexpressing GLP1R. After the receptors were stimulated for 4–6 h, the production of luciferase was quantified as an index of GLP1R activation. GLP1 was cleaved by FAP and DPP9 with DPP9 cleaving more peptide due to the shorter half-life followed by FAP (Figure S7). This was not the case for 2-GLP1, where we did not observe any change in potency. This result indicates that 2-GLP1 is refractory to DPP9 and FAP action and expands the protective umbrella through this modification to other enzymes within the protease family.

***In Vivo* Glucose Tolerance Test.** Blood glucose levels are regulated by the ability of GLP1 to promote insulin secretion, and we explored if our compounds achieve the same. We compared the ability of select peptides (2-GLP1, 7-GLP1, liraglutide, and 2-liraglutide) and GLP1 to stimulate insulin secretion and normalize glucose levels by performing *in vivo* glucose tolerance tests (Figure 5). Before peptide administration, the blood glucose of C57BL/6J mice was measured establishing an average fasting glucose level of 71 mg/dL. Mice were injected peritoneally with vehicle, GLP1, 2-GLP1, 7-GLP1, liraglutide, or 2-liraglutide at doses of 1 or 0.1 mg/kg. After 1 h, mice were administered an oral glucose bolus and mice administered vehicle showed a dramatic spike in blood glucose levels after 30 min (Figure 5a). All mice administered

GLP1R agonists showed a less pronounced increase in blood glucose, which returned to basal glucose levels at about 120 min post glucose bolus. The area under the curve (AUC) between 0 and 120 min was significantly different for all GLP1R agonists compared to the vehicle (Figure 5b). To determine circulation longevity of each peptide, we administered a second bolus 5 h after the first glucose challenge. Lipidated analogues, liraglutide and **2-liraglutide**, initially dosed at 1 mg/kg remained in circulation and prevented blood glucose level excursion post 30 min after the second glucose challenge (Figure 5c). No significant difference between **GLP1**, **2-GLP1**, **7-GLP1**, and vehicle was observed and **2-liraglutide** performed just as well as **liraglutide** at both doses. This demonstrates the utility of *N*-alkylated peptides that are also lipidated to avoid renal clearance, in maintaining blood glucose levels to the same extent as compounds currently in clinical use. These findings also underscore the initial expectation that these minimal modifications do not affect *in vivo* biological function.

Serum Stability. Although Liraglutide has been shown to have increased serum stability, it is still metabolically processed by DPP4 (Figures 3c and 4b). Compound **2-liraglutide**, which is only minimally modified from liraglutide with the sole addition of the *N*-terminal alkyl tether, but otherwise has the same sequence, linker, and lipid side chain, has a longer systemic half-life that is similar to semaglutide, but without the use of Aib at position 2.²⁵ Compounds were administered to rats via oral gavage (5 nmol/kg), and blood was collected sublingually to quantify the amount of peptide that remained in circulation via bioassay. The observed half-lives of liraglutide and **2-liraglutide** were 3 and 5.5 h (Figure S8). These results indicate that the designed compounds survive in the bloodstream intact without damage by other metabolic processes and are active at the cognate receptor. The protracted lifetime is similar to semaglutide, the leading peptide-based compound in clinical use.²⁵

CONCLUSION

Decades of research on the design and modification of incretins has yielded functional templates that have resulted in several therapeutics for type 2 diabetes and related comorbidities.²¹ The *N*-terminus of GLP1 has previously been assumed to be intolerant of modifications as it resides in a peptide domain crucial for intermolecular interactions with the receptor. This is mainly because the range of ligand modifications hitherto have relied primarily on acyl functionalities that the receptor has evolved to recognize as improperly processed fragments. We demonstrate in this study that not only is the *N*-terminal amenable to chemical modification but it is also tolerant to a wide range of functionality if the group is attached by alkylation, whereas acylation is deleterious to activity. The compounds described in this study are as efficacious and long-lived as the leading compounds in the clinic today. In addition, the strategy is generalizable to all templates for a range of receptors. Because the number of active compounds that can be generated for each template is in the hundreds, optimization of pharmacological properties (e.g., ligand bias and oral bioavailability) should be possible.

The use of native GLP1 as a therapeutic is hindered by its extreme lability, in that it is very short-lived with a half-life *in vivo* of less than 2 min, as it is degraded by the ubiquitous serine protease DPP4. We demonstrate in this study that the simultaneous goal of rendering substrates refractory to enzyme

catalyzed hydrolysis and maintaining activity at the receptor is possible. We conducted activity assays (cAMP production) as well as LC ESI-MS to establish the resistance of the peptide constructs to DPP4. Through post-incubation assays, we show that the designed peptides have protracted lifetimes *in vivo* and are also able to restore normoglycemia over a prolonged period of time. We further show that the strategy is not only applicable to GLP1 but also to leading compounds currently in clinical use including liraglutide, exenatide, and semaglutide.

We extended the scope of such *N*-terminal modifications to other receptors within the class B family, namely, GCGR, GIPR, and GLP2R. These related receptors are also therapeutically relevant and provide avenues for design and development to a range of clinically useful peptide constructs. Collectively, this strategy is broadly applicable for modification of ligands which activate receptors that define the axis of glucose metabolism.

The results presented here open up an array of opportunities that have traditionally not been accessible. Class B G protein-coupled receptors (GPCRs) which encompass GLP1R, GIPR, GCGR, and GLP2R all interact with intracellular partner proteins upon activation by ligand. Several G proteins such as $G\alpha_s$, $G\alpha_i$, $G\alpha_q$, $G\alpha_o$, and arrestins (β -arrestin-1, β -arrestin-2) can serve in the effector role.^{11,55,56} These partner/effector G proteins activate one or more signaling pathways such as mobilization of intracellular calcium and ERK1/2 phosphorylation ($G\alpha_s$, $G\alpha_q$, and $G\alpha_{12/10}$ or production of cAMP and insulin secretion ($G\alpha_s$).^{57,58} β -Arrestins control receptor internalization and/or mediate G protein independent signaling that lead to cell proliferation or apoptosis in part through activation of MAP kinases such as ERK1/2.⁵⁹ It is possible that binding of certain ligands to the receptor can selectively engage particular effectors over others resulting in “biased” signaling favoring one channel of activity over another. This has led to the idea that selective modulation of certain channels may result in ligands (therapeutics) that have fewer side effects sometimes accompanying receptor stimulation.⁵⁵ The molecular grafting method described in this study modifies the ligands at the crucial *N*-terminal residues and introduces new moieties in the receptor pocket interacting with a region of the receptor that is buried deep in the membrane. We believe this provides an avenue to achieve such signaling bias and studies along these lines are ongoing in our laboratories.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscentsci.0c01237>.

Two-dimensional diagram of His7 interacting with GLP1R (Figure S1A), His7 mutations of GLP1 and calculated EC₅₀ (Figure S1B), concentration response curve of 1-GLP1 with DPP4 incubation (Figure S2), concentration response curve of semaglutide and 2-Semaglutide(Ala2) with DPP4 incubation (Figure S3), Michaelis–Menten kinetics of DPP4 (Figure S4), concentration response curves of 2-triagonist(Ala2) (Figure S5), MetEnk and 2-MetEnk (Figure S6), GLP1 and 2-GLP1 incubated with FAP and DPP9 (Figure S7), and *in vivo* serum stability of 2-liraglutide (Figure S8); safety information; materials and methods; and analytical characterization of synthesized compounds and peptides (Table S1) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Martin Beinborn – Department of Chemistry, Tufts University, Medford, Massachusetts 02155, United States; Molecular Cardiology Research Institute, Tufts Medical Center, Boston, Massachusetts 02111, United States; Email: martin.beinborn@tufts.edu

Krishna Kumar – Department of Chemistry, Tufts University, Medford, Massachusetts 02155, United States; orcid.org/0000-0002-0548-5014; Email: krishna.kumar@tufts.edu

Authors

Kathleen M. Sicinski – Department of Chemistry, Tufts University, Medford, Massachusetts 02155, United States

Vittorio Montanari – Department of Chemistry, Tufts University, Medford, Massachusetts 02155, United States

Venkata S. Raman – Department of Chemistry, Tufts University, Medford, Massachusetts 02155, United States

Jamie R. Doyle – Molecular Cardiology Research Institute, Tufts Medical Center, Boston, Massachusetts 02111, United States

Benjamin N. Harwood – Molecular Cardiology Research Institute, Tufts Medical Center, Boston, Massachusetts 02111, United States

Yi Chi Song – Department of Chemistry, Tufts University, Medford, Massachusetts 02155, United States

Micaella P. Fagan – Department of Neuroscience, Tufts University School of Medicine, Boston, Massachusetts 02111, United States

Maribel Rios – Department of Neuroscience, Tufts University School of Medicine, Boston, Massachusetts 02111, United States

David R. Haines – Department of Chemistry, Wellesley College, Wellesley, Massachusetts 02481, United States

Alan S. Kopin – Molecular Cardiology Research Institute, Tufts Medical Center, Boston, Massachusetts 02111, United States

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acscentsci.0c01237>

Notes

The authors declare the following competing financial interest(s): K.K. is Scientific Founder of Velum, Inc. K.K., V.M., M.B., V.R. are named inventors in patent applications (WO2017075505 and WO2017075522 (additionally A.S.K.) to Tufts University) related to the work described in this paper. All other authors declare no competing interests.

■ ACKNOWLEDGMENTS

We thank Prof. Rebecca Scheck (Tufts University) for help with LC ESI-MS/MS. Thanks to Kathryn A. Lipford, Rebecca Roush and Dr. Michael J. Hanley for compound synthesis and assays described in Supplemental Figure S1B. The ESI-MS and NMR facilities at Tufts University were established by grants from the NSF (0320783 and 0821508). This work was supported in part by National Institutes of Health grants GM133272, GM130257 (to K.K.), and AG061909 (M.B.). B.N.H. was supported in part by NIH 5T32HL069770 (Karas). We thank Tufts University for ongoing and past support.

■ REFERENCES

- (1) Campbell, J. E.; Drucker, D. J. Pharmacology, physiology, and mechanisms of incretin hormone action. *Cell Metab.* **2013**, *17*, 819–837.
- (2) Heinrich, G.; Gros, P.; Habener, J. F. Glucagon gene sequence. Four of six exons encode separate functional domains of rat proglucagon. *J. Biol. Chem.* **1984**, *259*, 14082–14087.
- (3) Muller, T. D.; Finan, B.; Bloom, S. R.; D'Alessio, D.; Drucker, D. J.; Flatt, P. R.; Fritsche, A.; Gribble, F.; Grill, H. J.; Habener, J. F.; Holst, J. J.; Langhans, W.; Meier, J. J.; Nauck, M. A.; Perez-Tilve, D.; Pocai, A.; Reimann, F.; Sandoval, D. A.; Schwartz, T. W.; Seeley, R. J.; Stemmer, K.; Tang-Christensen, M.; Woods, S. C.; DiMarchi, R. D.; Tschop, M. H. Glucagon-like peptide 1 (GLP-1). *Mol. Metab.* **2019**, *30*, 72–130.
- (4) Vrang, N.; Larsen, P. J. Preproglucagon derived peptides GLP-1, GLP-2 and oxyntomodulin in the CNS: role of peripherally secreted and centrally produced peptides. *Prog. Neurobiol.* **2010**, *92*, 442–462.
- (5) Wu, T.; Rayner, C. K.; Horowitz, M. *Incretins*; Springer International Publishing: 2015; pp 137–171.
- (6) Manandhar, B.; Ahn, J. M. Glucagon-like peptide-1 (GLP-1) analogs: recent advances, new possibilities, and therapeutic implications. *J. Med. Chem.* **2015**, *58*, 1020–1037.
- (7) Kim, W.; Egan, J. M. The role of incretins in glucose homeostasis and diabetes treatment. *Pharmacol. Rev.* **2008**, *60*, 470–512.
- (8) Knudsen, L. B. Glucagon-like peptide-1: the basis of a new class of treatment for type 2 diabetes. *J. Med. Chem.* **2004**, *47*, 4128–4134.
- (9) Madsen, K.; Knudsen, L. B.; Agerso, H.; Nielsen, P. F.; Thogersen, H.; Wilken, M.; Johansen, N. L. Structure-activity and protraction relationship of long-acting glucagon-like peptide-1 derivatives: importance of fatty acid length, polarity, and bulkiness. *J. Med. Chem.* **2007**, *50*, 6126–6132.
- (10) Zhang, Y.; Sun, B.; Feng, D.; Hu, H.; Chu, M.; Qu, Q.; Tarrasch, J. T.; Li, S.; Sun Kobilka, T.; Kobilka, B. K.; Skiniotis, G. Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein. *Nature* **2017**, *546*, 248–253.
- (11) Wootten, D.; Simms, J.; Miller, L. J.; Christopoulos, A.; Sexton, P. M. Polar transmembrane interactions drive formation of ligand-specific and signal pathway-biased family B G protein-coupled receptor conformations. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 5211–5216.
- (12) Steensgaard, D. B.; Thomsen, J. K.; Olsen, H. B.; Knudsen, L. B. The molecular basis for the delayed absorption of the once-daily human GLP-1 analogue, liraglutide. *Diabetes* **2008**, *57*, A164–A164.
- (13) Knudsen, L. B.; Nielsen, P. F.; Huusfeldt, P. O.; Johansen, N. L.; Madsen, K.; Pedersen, F. Z.; Thogersen, H.; Wilken, M.; Agerso, H. Potent derivatives of glucagon-like peptide-1 with pharmacokinetic properties suitable for once daily administration. *J. Med. Chem.* **2000**, *43*, 1664–1669.
- (14) Jimenez-Solem, E.; Rasmussen, M. H.; Christensen, M.; Knop, F. K. Dulaglutide, a long-acting GLP-1 analog fused with an Fc antibody fragment for the potential treatment of type 2 diabetes. *Curr. Opin. Mol. Ther.* **2010**, *12*, 790–797.
- (15) Knudsen, L. B.; Lau, J. The Discovery and Development of Liraglutide and Semaglutide. *Front. Endocrinol. (Lausanne, Switz.)* **2019**, *10*, 155.
- (16) Davies, M. J.; D'Alessio, D. A.; Fradkin, J.; Kernan, W. N.; Mathieu, C.; Mingrone, G.; Rossing, P.; Tsapas, A.; Wexler, D. J.; Buse, J. B. Management of Hyperglycemia in Type 2 Diabetes, 2018. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes Care* **2018**, *41*, 2669–2701.
- (17) Trzaskalski, N. A.; Fadzeyeva, E.; Mulvihill, E. E. Dipeptidyl Peptidase-4 at the Interface Between Inflammation and Metabolism. *Clin. Med. Insights: Endocrinol. Diabetes* **2020**, *13*, 1–10.
- (18) Mulvihill, E. E.; Drucker, D. J. Pharmacology, physiology, and mechanisms of action of dipeptidyl peptidase-4 inhibitors. *Endocr. Rev.* **2014**, *35*, 992–1019.
- (19) Sebokova, E.; Christ, A. D.; Wang, H.; Sewing, S.; Dong, J. Z.; Taylor, J.; Cawthorne, M. A.; Culler, M. D. Taspoglutide, an analog of

human glucagon-like Peptide-1 with enhanced stability and in vivo potency. *Endocrinology* **2010**, *151*, 2474–2482.

(20) Rosenstock, J.; Balas, B.; Charbonnel, B.; Bolli, G. B.; Boldrin, M.; Ratner, R.; Balena, R. The fate of taspoglutide, a weekly GLP-1 receptor agonist, versus twice-daily exenatide for type 2 diabetes: the T-emerge 2 trial. *Diabetes Care* **2013**, *36*, 498–504.

(21) Jones, L. H.; Price, D. A. Medicinal Chemistry of Glucagon-Like Peptide Receptor Agonists. *Prog. Med. Chem.* **2013**, *52*, 45–96.

(22) Johnson, L. M.; Barrick, S.; Hager, M. V.; McFedries, A.; Homan, E. A.; Rabaglia, M. E.; Keller, M. P.; Attie, A. D.; Saghatelian, A.; Bisello, A.; Gellman, S. H. A potent alpha/beta-peptide analogue of GLP-1 with prolonged action in vivo. *J. Am. Chem. Soc.* **2014**, *136*, 12848–12851.

(23) Meng, H.; Krishnaji, S. T.; Beinborn, M.; Kumar, K. Influence of selective fluorination on the biological activity and proteolytic stability of glucagon-like peptide-1. *J. Med. Chem.* **2008**, *51*, 7303–7307.

(24) Ueda, T.; Tomita, K.; Notsu, Y.; Ito, T.; Fumoto, M.; Takakura, T.; Nagatome, H.; Takimoto, A.; Mihara, S.; Togame, H.; Kawamoto, K.; Iwasaki, T.; Asakura, K.; Oshima, T.; Hanasaki, K.; Nishimura, S.; Kondo, H. Chemoenzymatic synthesis of glycosylated glucagon-like peptide 1: effect of glycosylation on proteolytic resistance and in vivo blood glucose-lowering activity. *J. Am. Chem. Soc.* **2009**, *131*, 6237–6245.

(25) Lau, J.; Bloch, P.; Schaffer, 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. *J. Med. Chem.* **2015**, *58*, 7370–7380.

(26) Green, B. D.; Mooney, M. H.; Gault, V. A.; Irwin, N.; Bailey, C. J.; Harriott, P.; Greer, B.; O'Harte, F. P.; Flatt, P. R. N-terminal His(7)-modification of glucagon-like peptide-1(7–36) amide generates dipeptidyl peptidase IV-stable analogues with potent anti-hyperglycaemic activity. *J. Endocrinol.* **2004**, *180*, 379–388.

(27) Wootten, D.; Reynolds, C. A.; Koole, C.; Smith, K. J.; Mobarec, J. C.; Simms, J.; Quon, T.; Coudrat, T.; Furness, S. G.; Miller, L. J.; Christopoulos, A.; Sexton, P. M. A Hydrogen-Bonded Polar Network in the Core of the Glucagon-Like Peptide-1 Receptor Is a Fulcrum for Biased Agonism: Lessons from Class B Crystal Structures. *Mol. Pharmacol.* **2016**, *89*, 335–347.

(28) Xiao, Q.; Giguere, J.; Parisien, M.; Jeng, W.; St-Pierre, S. A.; Brubaker, P. L.; Wheeler, M. B. Biological activities of glucagon-like peptide-1 analogues in vitro and in vivo. *Biochemistry* **2001**, *40*, 2860–2869.

(29) Buckley, D. I.; Habener, J. F.; Mallory, J. B.; Mojsov, S. GLP-1 Analogs Useful for Diabetes Treatment; European Patent EP 0512 042 B1; 1991.

(30) Ohneda, A.; Ohneda, K.; Ohneda, M.; Koizumi, F.; Ohashi, S.; Kawai, K.; Suzuki, S. The structure-function relationship of GLP-1 related peptides in the endocrine function of the canine pancreas. *Tohoku J. Exp. Med.* **1991**, *165*, 209–221.

(31) Gallwitz, B.; Ropeter, T.; Morys-Wortmann, C.; Mentlein, R.; Siegel, E. G.; Schmidt, W. E. GLP-1-analogues resistant to degradation by dipeptidyl-peptidase IV in vitro. *Regul. Pept.* **2000**, *86*, 103–111.

(32) Hareter, A.; Hoffmann, E.; Bode, H. P.; Goke, B.; Goke, R. The positive charge of the imidazole side chain of histidine(7) is crucial for GLP-1 action. *Endocr. J.* **1997**, *44*, 701–705.

(33) DesMarteau, D. D.; Montanari, V. Easy preparation of bioactive peptides from the novel N-alpha-trifluoroethyl amino acids. *Chem. Lett.* **2000**, *29*, 1052–1053.

(34) Baggio, L. L.; Drucker, D. J. Biology of incretins: GLP-1 and GIP. *Gastroenterology* **2007**, *132*, 2131–2157.

(35) Parthier, C.; Kleinschmidt, M.; Neumann, P.; Rudolph, R.; Manhart, S.; Schlenzig, D.; Fanghanel, J.; Rahfeld, J. U.; Demuth, H. U.; Stubbs, M. T. Crystal structure of the incretin-bound extracellular domain of a G protein-coupled receptor. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 13942–13947.

(36) Underwood, C. R.; Parthier, C.; Reedtz-Runge, S. Structural Basis for Ligand Recognition of Incretin Receptors. *Vitam. Horm.* **2010**, *84*, 251–278.

(37) Nauck, M. A.; Heimesaat, M. M.; Orskov, C.; Holst, J. J.; Ebert, R.; Creutzfeldt, W. Preserved incretin activity of glucagon-like peptide 1 [7–36 amide] but not of synthetic human gastric inhibitory polypeptide in patients with type-2 diabetes mellitus. *J. Clin. Invest.* **1993**, *91*, 301–307.

(38) Holst, J. J.; Gromada, J.; Nauck, M. A. The pathogenesis of NIDDM involves a defective expression of the GIP receptor. *Diabetologia* **1997**, *40*, 984–986.

(39) Vilsboll, T.; Krarup, T.; Madsbad, S.; Holst, J. J. Defective amplification of the late phase insulin response to glucose by GIP in obese Type II diabetic patients. *Diabetologia* **2002**, *45*, 1111–1119.

(40) Hojberg, P. V.; Vilsboll, T.; Rabol, R.; Knop, F. K.; Bache, M.; Krarup, T.; Holst, J. J.; Madsbad, S. Four weeks of near-normalisation of blood glucose improves the insulin response to glucagon-like peptide-1 and glucose-dependent insulinotropic polypeptide in patients with type 2 diabetes. *Diabetologia* **2009**, *52*, 199–207.

(41) Habegger, K. M.; Heppner, K. M.; Geary, N.; Bartness, T. J.; DiMarchi, R.; Tschop, M. H. The metabolic actions of glucagon revisited. *Nat. Rev. Endocrinol.* **2010**, *6*, 689–697.

(42) Finan, B.; Yang, B.; Ottaway, N.; Smiley, D. L.; Ma, T.; Clemmensen, C.; Chabenne, J.; Zhang, L.; Habegger, K. M.; Fischer, K.; Campbell, J. E.; Sandoval, D.; Seeley, R. J.; Bleicher, K.; Uhles, S.; Riboulet, W.; Funk, J.; Hertel, C.; Belli, S.; Sebokova, E.; Conde-Knape, K.; Konkar, A.; Drucker, D. J.; Gelfanov, V.; Pfluger, P. T.; Muller, T. D.; Perez-Tilve, D.; DiMarchi, R. D.; Tschop, M. H. A rationally designed monomeric peptide triagonist corrects obesity and diabetes in rodents. *Nat. Med.* **2015**, *21*, 27–36.

(43) Pospisilik, J. A.; Hinke, S. A.; Pederson, R. A.; Hoffmann, T.; Rosche, F.; Schlenzig, D.; Glund, K.; Heiser, U.; McIntosh, C. H.; Demuth, H. Metabolism of glucagon by dipeptidyl peptidase IV (CD26). *Regul. Pept.* **2001**, *96*, 133–141.

(44) Jaspan, J. B.; Polonsky, K. S.; Lewis, M.; Pensler, J.; Pugh, W.; Moossa, A. R.; Rubenstein, A. H. Hepatic-Metabolism of Glucagon in the Dog - Contribution of the Liver to Overall Metabolic Disposal of Glucagon. *Am. J. Physiol.* **1981**, *240*, E233–E244.

(45) Unson, C. G.; Macdonald, D.; Merrifield, R. B. The role of histidine-1 in glucagon action. *Arch. Biochem. Biophys.* **1993**, *300*, 747–750.

(46) Drucker, D. J. The Discovery of GLP-2 and Development of Teduglutide for Short Bowel Syndrome. *ACS Pharmacol. Transl. Sci.* **2019**, *2*, 134–142.

(47) Fosgerau, K.; Hoffmann, T. Peptide therapeutics: current status and future directions. *Drug Discovery Today* **2015**, *20*, 122–128.

(48) Marini, M.; Urbani, A.; Trani, E.; Bongiorno, L.; Roda, L. G. Interindividual variability of enkephalin-degrading enzymes in human plasma. *Peptides* **1997**, *18*, 741–748.

(49) Yakovleva, A. A.; Zolotov, N. N.; Sokolov, O. Y.; Kost, N. V.; Kolyasnikova, K. N.; Mischeeva, I. G. Dipeptidylpeptidase 4 (DPP4, CD26) activity in the blood serum of term and preterm neonates with cerebral ischemia. *Neuropeptides* **2015**, *52*, 113–117.

(50) Pels, K.; Kodadek, T. Solid-phase synthesis of diverse peptide tertiary amides by reductive amination. *ACS Comb. Sci.* **2015**, *17*, 152–155.

(51) Xing, X.; Fichera, A.; Kumar, K. A novel synthesis of enantiomerically pure 5,5,5,5',5',5'-hexafluoroleucine. *Org. Lett.* **2001**, *3*, 1285–1286.

(52) Bušek, P.; Malík, R.; Šedo, A. Dipeptidyl peptidase IV activity and/or structure homologues (DASH) and their substrates in cancer. *Int. J. Biochem. Cell Biol.* **2004**, *36*, 408–421.

(53) Keane, F. M.; Nadvi, N. A.; Yao, T. W.; Gorrell, M. D. Neuropeptide Y, B-type natriuretic peptide, substance P and peptide YY are novel substrates of fibroblast activation protein-alpha. *FEBS J.* **2011**, *278*, 1316–1332.

(54) Bjelke, J. R.; Christensen, J.; Nielsen, P. F.; Branner, S.; Kanstrup, A. B.; Wagtmann, N.; Rasmussen, H. B. Dipeptidyl peptidases 8 and 9: specificity and molecular characterization

compared with dipeptidyl peptidase IV. *Biochem. J.* **2006**, *396*, 391–399.

(55) Hager, M. V.; Johnson, L. M.; Wootten, D.; Sexton, P. M.; Gellman, S. H. beta-Arrestin-Biased Agonists of the GLP-1 Receptor from beta-Amino Acid Residue Incorporation into GLP-1 Analogues. *J. Am. Chem. Soc.* **2016**, *138*, 14970–14979.

(56) Made, V.; Babilon, S.; Jolly, N.; Wanka, L.; Bellmann-Sickert, K.; Diaz Gimenez, L. E.; Morl, K.; Cox, H. M.; Gurevich, V. V.; Beck-Sickinger, A. G. Peptide modifications differentially alter G protein-coupled receptor internalization and signaling bias. *Angew. Chem., Int. Ed.* **2014**, *53*, 10067–10071.

(57) Montrose-Rafizadeh, C.; Avdonin, P.; Garant, M. J.; Rodgers, B. D.; Kole, S.; Yang, H.; Levine, M. A.; Schwindinger, W.; Bernier, M. Pancreatic glucagon-like peptide-1 receptor couples to multiple G proteins and activates mitogen-activated protein kinase pathways in Chinese hamster ovary cells. *Endocrinology* **1999**, *140*, 1132–1140.

(58) Hallbrink, M.; Holmqvist, T.; Olsson, M.; Ostenson, C. G.; Efendic, S.; Langel, U. Different domains in the third intracellular loop of the GLP-1 receptor are responsible for G alpha(s) and G alpha(i)/G alpha(o) activation. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **2001**, *1546*, 79–86.

(59) Quoyer, J.; Longuet, C.; Broca, C.; Linck, N.; Costes, S.; Varin, E.; Bockaert, J.; Bertrand, G.; Dalle, S. GLP-1 mediates antiapoptotic effect by phosphorylating Bad through a beta-arrestin 1-mediated ERK1/2 activation in pancreatic beta-cells. *J. Biol. Chem.* **2010**, *285*, 1989–2002.