An integrated platform approach enables discovery of potent, selective and ligand-competitive cyclic peptides targeting the GIP receptor

Supplementary Information

Table of Contents

Protein expression and purification of hGIP-R	.2
Preparation of puromycin-fused mRNA library:	.3
In vitro selection of cyclic peptides binding to hGIP-R ECD:	.3
Confirmation of cyclisation after in vitro translation	.4
Next Generation sequencing	.5
Analysis of NGS data	.5
Sequence analysis:	.5
Peptide Synthesis	.6
Binding measurements	.8
GIP displacement assays1	1
Modelling1	2
Analysis of solubility1	4
In vitro plasma stability and quantification of in vivo samples by LC-MS:1	4
Pharmacokinetic studies1	5
Pharmacokinetic analysis1	6
Supplementary Info References1	6
Analytical data on purified peptides1	17

Protein expression and purification of hGIP-R

The DNA sequence encoding the extra cellular domain of human GIPR (22-138, UniProtKB - P48546) was synthesized (Taihe Biotechnology, Beijing, China). The signal peptide of CD33 was cloned to its N-terminus, and an avi-tag (GLNDIFEAQKIEWHE for in vitro biotinylation) followed by a HPC4 tag (EDQVDPRLIDGK to favour purification) were cloned to its C-terminus. This fragment was further cloned into pJSV002 plasmid for transient over-expression in HEK293 6E cells (Thermo Fisher Scientific). Plasmid transfection into HEK 293-6E cells was performed using transfection reagents 293fectin[™] (Gibco #12347-500, Thermo Fisher Scientific) and Opti-MEM[®]+GlutaMAX[™] (Gibco #51985-034, Thermo Fisher Scientific). Cells were cultivated in FreeStyle[™] 293 medium (Gibco, #12338-018, Thermo Fisher Scientific) under 37°C in 5% CO₂ atmosphere on an orbital shaker set at 140 rpm. The culture supernatant was harvested after 5 days of cultivation. The target protein GIPR(22-138)-avi-HPC4 was first purified by using an anti-HPC4 antibody-coupled affinity chromatography, then biotinylated on avi-tag using BirA biotin ligase prepared in-house. Finally, the biotinylated target protein was purified using gel filtration with Superdex® 75 (GE Healthcare). The produced protein, biotinylated GIPR(22-138)-avi-HPC4, was characterized by SDS-PAGE gel and SEC-HPLC (Superdex 200 10/300 GL column on Agilent HPLC) with >99% purity, identity was confirmed with LC-MS (Agilent PLRP-S column on Waters UPLC Xevo G2 Q-TOF)(impurities detected in the MS: partial deamidation of Asn, and partial deletion of the C-terminal Gly), and endotoxin (Kinetic Turbidimetric LAL Test kit, Charles River Laboratories) level of < 1 EU/mg.





Figure S 2: SEC-HPLC Chromatogram of purified GIP-R (ECD), column: Superdex 200 10/300 GL (GE Healthcare, Catalogue No. 28-9909-44, 300x10mm, 8.6µm partical size), 0.6ml/min isocratic gradient for 35 min (Na2HPO4 (8mM), KH2PO4 (1.5mM), NaCl (137mM), KCl (3mM), pH 7.4)

Preparation of puromycin-fused mRNA library:

A pool of NNK RNA library (5'-TAA TAC GAC TCA CTA TAG GGT TAA CTT TAA GAA GGA GAT ATA CAT ATG TGC (NNK)_m TGC GGC AGC GGC AGC GGC AGC TAG GAC GGG GGG GGG CGG AAA-3', m = 4-12) was prepared by *in vitro* transcription as described¹², and mixed in molar ratio

 $(NNK)_4:(NNK)_5:(NNK)_6:(NNK)_7:(NNK)_8:(NNK)_9:(NNK)_{10}:(NNK)_{11}:(NNK)_{12}=1:2:4:8:16:32:64:64:64.$ The mRNA library was ligated with a puromycin linker (5'-CTCCC GCCCC CCGTC C-(SPC18)_5-CC-puromycin-3') by T4 RNA ligase and purified by phenol–chloroform extraction and ethanol precipitation.

In vitro selection of cyclic peptides binding to hGIP-R ECD:

Translation of the first selection rounds was performed using 25 pmol mRNA-puromycin in 11.5 µl translation mixture (PURExpress delta RF1-3, NEB) at 37 °C for 30 min in the presence of PURExpress disulfide-bond enhancers, followed by incubation at 60 °C for 10 min. CP-mRNA product was then transcribed to CP-mRNA-cDNA using M-MLV reverse transcriptase (Promega), reverse primer P2 (5'-TGC GGC AGC GGC AGC GGC AGC TAG GAC GGG GGG CGG AAA-3') and dNTPs at 42 °C for 1 h. Following this, the contents were diluted to 40 µL with selection buffer (PBS pH7.4 supplemented with 0.05% Tween-20 and 0.01% BSA). Negative selections were performed by incubation of the diluted reverse-transcription solution for 30 min at 4 °C with a suspension of streptavidin-functionalised Dynabeads magnetic beads (M-280, Invitrogen) which were pre-loaded with biotinylated Avi-tag peptide. The beads were precipitated using a magnet, the supernatant collected, and transferred to a fresh tube for another round of negative selection. After three rounds of negative selection, the resulting supernatant was then incubated for 30 min at 4 °C with biotinylated GIP-R loaded magnetic beads. All bead pellets were washed 3 times with milliQ water, and the bound cDNA sequences eluted by incubation at 95 °C for 5 min in 100 uL PCR buffer (1x NH4 buffer, 2.5 mM MgCl₂, 0.25 mM dNTPs, and 0.5 uM each of reverse primer P2 (5'-TGC GGC AGC GGC AGC GGC AGC TAG GAC GGG GGG CGG AAA-3') and forward primer P1 (5'-TAA TAC GAC TCA CTA TAG GGT TAA CTT TAA GAA GGA GAT ATA CAT ATG TGC-3'). Absolute quantities of recovered DNA after each selection step were determined by qPCR. After addition of Taq DNA polymerase to the positive selection eluate, the mixture was used for PCR amplification. The amplified DNA was purified by phenol/chloroform extraction followed by ethanol purification. The resulting DNA library was transcribed

in vitro, the resultant RNA purified by phenol/chloroform extraction, and quantified using an Invitrogen[™] Qubit[™] 4 Fluorometer (Thermo Fisher). The RNA was then ligated with the puromycin linker and used for the next round of selection. Subsequent selection rounds were performed as described, but using 12.5 pmol puromycylated mRNA in 5.45 uL IVTT reactions. After the final rounds of the selection, the resulting cDNA libraries were sequenced by NGS (see below).

Confirmation of cyclisation after in vitro translation

Efficiency of disulfide mediated cyclisation of a test CP sequence (^fMCVWDPRTFYLSRICGSGSGS)^[1] following IVTT with different combinations of RNAse inhibitor and disulphide bond enhancer reagent was investigated. A DNA template encoding for a test CP (5'-



Figure S 3: RNAse inhibitor and disulphide bond enhancer effect on cyclisation

Next Generation sequencing

dsDNA libraries were PCR amplified to contain internal NGS barcodes and stubbed Illumina tails using ACACGACGCTCTTCCGATCTNNNNNXXXAGGGTTAACTTTAAGAAGGAGATATACATATG

GACGTGTGCTCTTCCGATCGCTGCCGCTGCCGCTGCC where XXX denote CAC, GTG, ATC, CCG, AGA or TCA using the KAPA HiFi HotStart Real-Time Library Amp Kit (Roche) following manufactures recommendations on a qPCR machine to limit the number of PCR cycles. Following PCR the products were purified and normalized using SequalPrep[™] Normalization Plate Kit (Thermo) and products with different internal barcodes where pooled as used as template for a second PCR to add outer Illumina barcodes with D501-D508 and D701-D712 again using KAPA HiFi HotStart Real-Time Library Amp Kit (Roche). Following the second PCR reactions the samples were pooled, run on a 1% agarose gel and the fragments of the desired size of 261bp where purified. The library concentrations were measured using Qubit 3.0 fluorometer and Qubit dsDNA HS assay kit (Life Technologies), diluted to 4nM and sequenced on an Illumina NextSeq 75cycle HIGH kit following manufactures recommendations.

Analysis of NGS data

The FASTQC tool (www.bioinformatics.bbsrc.ac.uk/projects/fastqc/) was used to evaluate the quality of the fastq files, an in-house developed R-script was used to deduce the variable region in the data between

CTTTAAGAAGGAGATATACAT and TGCGGCAGCGG and R was furthermore used to merge the sequences and counts between the different samples into one matrix. All sequences that were detected by less than ten sequencing reads across all 35 samples were considered sequencing noise and removed from table.



Figure S 4: a) Recovery (in %) of cDNA in each round (1st negative selection and positive selection shown), b) 10 most enriched sequences after 5 rounds of selection against hGIP-R ECD with number of counts observed in NGS pool from R1 – R5 of selection, all peptides contain an N-terminal Ac group and C-terminal flag tag and are cyclized as disulphides

Sequence analysis:



Figure S 5: Distribution of top 3160 unique peptide sequences by cluster



Figure S 6: Consensus cluster sequence logos

Peptide Synthesis

All chemicals were of analytical grade or higher. Triisopropylsilane (TIPS), N,N'-

diisopropylcarbodiimide (DIC) were from Sigma–Aldrich, Chemie GmBH (Steinheim, Germany). Acetonitrile (ACN) (LiChrosolve), trifluoroacetic acid S3 (TFA), and diethyl ether were purchased from Merck KGaA (Darmstadt, Germany). Water came from a MilliQ equipment (Advantage A10) from Millipore (Molsheim, France). Standard Fmocamino acids, resins and coupling reagents, Oxyma Pure were from Novabiochem (Darmstadt, Germany) or Protein Technologies (Tucson, USA). N-Methyl pyrrolidone (NMP) dimethylformamide (DMF) and piperidine were from Biosolve (Dieuze, France). PS resin, and Fmoc-Rink amide PS resin, (synthesis of peptide 20) were all purchased from Merck Millipore (Novabiochem).

Peptides were either synthesised by standard SPPS using PAL-AM or pre-loaded wang resins (loading ~0.3M) on Gyros Protein Technologies Prelude or Symphony X machines (125 μ mol scale) or purchased from Apigenex, Prague, Czech Republic. The resin was washed with dichloromethane (DCM), and peptides were cleaved from the resin by incubating with TFA/TIPS/DTT/water (94:2:2:2) for 3h at room temperature, followed by precipitation with diethyl ether. The precipitated peptides were washed three times with diethyl ether, and diluted to 50 μ M in folding buffer (100 mM HEPES pH 7.4, 20% DMSO) to induce disulfide-mediated macrocyclisation. Purification was performed on Gilson LC systems, equipped with a 322 pump module, 155 UV/vis detector and GX-271 automated sample collector. A typical gradient of 20% - 60% solvent B over 30 min (flow rate of 25 mL/min, Axia Gemini 5uM NX-C18 110 Å columns, 250x30 mm) was employed (Solvent A: 0.1% TFA in water, Solvent B: 0.1 % TFA in Acetonitrile). The final peptides were isolated by lyophilisation.

96 well plate peptide synthesis was performed on an Intavis MultiPep RSi machine on a 5 μmol scale using PAL AM resin already equipped with the flag tag and linker (GSGSDYKDDDDK, final loading 0.30 mmol/g, prepared by standard solid phase synthesis on a Gyros Protein Technologies Symphony X or CS Bio CS136XTmachine). Triple couplings were used to ensure sufficient peptide purity (deprotection: 0.1 M Oxyma in 20% piperidine in DMF, Activator: 3 M DIC, Base: 3M collidine in DMF). Peptide cleavage was performed in 96 well filter plates using 2% DTT, 2% Triisopropylsilane (TIS) and 2% water in TFA for 1h at RT. Subsequently, peptides were precipitated using ether and

redissolved in DMSO after filtration to give a final average concentration of about 1.5 mM. Cyclisation was carried out in 20% DMSO in Hepes buffer (100 mM, pH 7.4) at a concentration of 60 μ M.

The identity and purity of peptides was determined by LC-MS on a Waters Acquity system, equipped with a SQ detector (BA699) and a PDA detector (MIOUPD33A) using an Acquity UPLC BEH C18 column (1.7 μ m, 2.1 mm X 50 mm). For 96 well plates a Waters Acquity system, equipped with a TUV detector (F08UPT321M) and a QDA detector (KDA3351), a plate hotel (G18UP0275H) and an Acquity UPLC BEH C18 column (1.7 μ m, 2.1 mm X 50 mm) was used. Linear gradient: 10 % to 90 % B, Gradient run-time: 3 min; Solvent A: 0.1% TFA in water, Solvent B: 0.1 % TFA in acetonitrile

Peptides were quantified using UHPLC-CAD on a Thermo Fisher Vanquish system equipped with an ACQUITY UPLC CSH C18 Column, (130Å, 1.7 μm, 2.1 mm X 50 mm) and a Charged Aerosol Detector H. (Solvent A: water with 0.1% TFA, solvent B: acetonitrile with 0.1% TFA, gradient 0-80% B 0.0 – 4.0 min flow rate 0.45 mL/min)

Table S1: sequences and identity of purified peptides

Name	Sequence	N-terminus	C-terminus	retention time UPLC [min]	purity UPLC [%]	exact mass found	exact mass calculated [M+xH]/x	charge state MS x =
A_0	MCLILPIFPYFICGSGSDYKDDDDK	acetyl	amide	5,23	96	1447,5880	1447,6507	2
B_1275	MCFTHFHMLWPFCGSGSDYKDDDDK	acetyl	amide	4,53	95	1511,0409	1511,1013	2
B_1275.1	MCFTHFHMLWPFCGSGSDYKDDDDK	gGlu 2xOEG	amide	5,33	93	1232,1588	1232,2059	3
B_1275.2	MSFTHFHMLWPFSGSGSDYKDDDDK	acetyl	amide	4,31	97	1496,0591	1496,1320	2
B_1275.3	MCFTHFHMLWPFC	gGlu 2xOEG	amide	6,06	96	1206,5192	1206,5613	3
B_1275.4	ACFTHFDLLWPFCGSGSDYKDDDDK	acetyl	amide	4,69	97	1461,0425	1461,1055	2
B_1275.5	ACFTHFDLLWPFC	acetyl	acid	5,27	95	1639,6091	1639,6999	1
B_1275.6	ACFTHFDLLWPFC	gGlu 2xOEG	acid	6,41	98	1157,0229	1157,0575	2
B_1275.7	ACWTEFELLWPFCGSGSDYKDDDDK	acetyl	amide	4,94	99	1483,5472	1483,6106	2
B_1275.8	SCFRAFQLLWPFCGSGSDYKDDDDK	acetyl	amide	4,69	96	1470,0677	1470,1346	2
B_1275.9	ACFRHFELLWPFCGSGSDYKDDDDK	acetyl	amide	4,46	98	1495,5684	1495,6400	2
B_1275.10	ACFTHFHLLWPPCGSGSDYKDDDDK	acetyl	amide	4,31*	56	1447,0586	1447,1136	2
B_1275.11	SCWEEFELVWPFCGSGSDYKDDDDK	acetyl	amide	4,67	100	1498,5457	1498,5977	2
B_1275.12	MCFTHFHMLWPFC	acetyl	acid	5,11	100	1739,5823	1739,6917	1
B_1275.13	ACFRHELLWPFC	acetyl	acid	5,12	97	1708,6688	1708,7690	1
B_3	MCFQYFHILWPFCGSGSDYKDDDDK	acetyl	amide	4,86	83	1528,5607	1528,6308	2
B_3.1	MCFQYFHILWPFCGSGSDYKDDDDK	gGlu 2xOEG	amide	5,61	95	1243,8380	1243,8922	3
B_3.2	MSFQYFHILWPFSGSGSDYKDDDDK	acetyl	amide	4,65	95	1513,5875	1513,6615	2
B_5	MCFIHWDMLWPFCGSGSDYKDDDDK	acetyl	amide	4,91	96	1525,5377	1525,6090	2
B_5.1	MCFIHWDMLWPFCGSGSDYKDDDDK	gGlu 2xOEG	amide	5,66	96	1241,8324	1241,8776	3
B_5.2	MSFIHWDMLWPFSGSGSDYKDDDDK	acetyl	amide	4,72	97	1510,5636	1510,6397	2
B_5.3	MCFIHWDMLWPFC	acetyl	acid	5,83	97	1768,6033	1768,7070	1
B_68	MCFRFFSLLWPFCGSGSDYKDDDDK	acetyl	amide	5,09	95	1509,5775	1509,6412	2
B_68.1	MCFRFFSLLWPFCGSGSDYKDDDDK	gGlu 2xOEG	amide	5,85	91	1231,1820	1231,2324	3
B_68.2	MSFRFFSLLWPFSGSGSDYKDDDDK	acetyl	amide	4,93	99	1494,6030	1494,6718	2
M_46	MCLPWFILRSVVCGSGSDYKDDDDK	acetyl	amide	4,91	97	1444,5841	1444,6490	2
M_46.1	MCLPWFILRSVVC	acetyl	amide	5,68	98	1605,7119	1605,8030	2

*double peak with same exact mass, possibly cis-trans proline isomers

Binding measurements

For single concentration binding experiments, 50 μ M peptide stocks in folding buffer were diluted 1:50 in mRNA display selection buffer (PBS pH7.4 supplemented with 0.05% Tween-20 and 0.01% BSA) in a 384 well tilted bottom microplate (ForteBio). Biotinylated hGIP-R protein was loaded onto Streptavidin-functionalised biosensors (ForteBio), and immersed into 1 μ M peptide stocks to obtain association kinetics curves. Dissociation kinetics curves were obtained by subsequent immersion of ligand-associated loaded sensors into a 1:50 dilution of folding buffer in mRNA display selection buffer.

For accurate K_d determination, dilution series of each peptide were performed, and association and dissociation kinetics curves were obtained for each concentration in order to fit a multi-point K_d.

Data collection and analysis was performed using the Octet Instrument Control and Data Analysis software packages (ForteBio, Ver 9.0.0.10) and Prism (GraphPad Software, Ver. 7.05).



Figure S 7: Single-concentration association and dissociation kinetics curves for CP binding to hGIP-R ECD



Figure S 8: Single concentration hGIP-R ECD binding K_d values and dissociation rates of synthesized crude peptides, and corresponding abundances/reads of peptide sequences in the final round of selection against GIP-R ECD. (N.D: not determined due to failure of parallel peptide synthesis)





Figure S 9: Multi-concentration association and dissociation kinetics curves for purified CP binding to hGIP-R ECD

GIP displacement assays

BHK cells stably expressing the human GIP receptor were seeded into a 96 well plate (10000 cells per well, Poly-Dlysine plate, Corning, cat# 354651) the day before the assay. Cells were washed with HBSS (2 x 140 µL, Gibco, 14025) at room temperature, and incubated for 16 h at 4 °C with a dilution series of the cyclic peptides in assay buffer (10 mM HEPES in HBSS, 0.1% pluronic F-68 (Gibco, 2404), 0.1% ovalbumin (SigmaAldrich, A5503), pH 7.4) and 60 pM [¹²⁵I]hGIP (produced at Novo Nordisk A/S). The cells were then washed with ice-cold PBS (2 x140µL), and lysed with NaOH (0.1 mM, 50 µl/well) with shaking for 5 min. Following this, microscint-40 (100 µl/well, Perkin Elmer, 6013641) was added, and the plates incubated at RT for 30 min before reading in scintillation counter (Topcount® NXT™ HTS from Packard). Nonlinear regression analysis on the output files was performed in the Windows program GraphPad Prism 7 (GraphPad software, USA) using the equation "log(inhibitor) vs response (three parameters)" to get IC50 values. These chosen peptides were shown to not displace radiolabelled GCG or GLP-1 from GCG-R or GLP-1R expressing cells, showing that they are GIP-receptor selective



Figure S 10: Competitive binding measurements of CPs to the GIP-R



Figure S 11: Competitive binding measurements of CPs to the GLP-1 and GCG receptors

Modelling

A selected parent peptide (seq MCFTHFHMLWPFCG) was folded inside the GIP receptor's binding site using an inhouse Rosetta script^[2] (Fleishman et al., 2011) see attached supplementary code. Briefly, the peptide was first initiated as a residue (stub) located inside the receptor's binding site; to place the stub we used the coordinates of Phe22 of the GIP and Phe103 of the Fab heavy chain from the 2qkh and 6dkj structures^[3] (Parthier, et al. 2007; Killion et al., 2018), respectively. The stub was further mutated in one of the 4 most crucial according to SAR residues, in particular, Leu, Trp, Pro or Phe. The rest of the peptide was then grown, cyclized through the N- and C- terminal Cys residues using the generalized kinematic closure protocol^[4] (Coutsias et al., 2004; Mandell et al., 2009), and relaxed. The obtained poses were further filtered based on the total energy, the peptide energy, and the disulfide bridge's quality.

The final poses (24,000 in total) were clustered using hierarchical clustering and LWPF_RMSD as a measure of distance, where LWPF_RMSD is the Root Mean Square Deviation of the LWPF ligand atoms between two given structures provided that the structures are aligned on the receptor atoms. Using 4 Å as a threshold, 4 major clusters have been identified (674, 682, 1503 and 1519); note that 3 of them (674, 682 and 1503) were also characterized by low values of the binding energy (Figure CP1_SI A and B).

For each cluster a top pose was selected based on the following metrics: the binding energy, the shape complementarity between the peptide and the receptor, and the area of the interface surface. For every pose the molecular dynamics (MD) simulations using Amber 2018 suite have been performed (Case et al., 2018). Briefly, the structures of the complexes have been immersed in a water box; the systems were further neutralized with Na⁺ and Cl⁻ ions. All the atoms were described using the Amber ff14SB force field^[5] (Maier et al., 2015), and TIP3P was used as a water model^[6] (Jorgensen et al., 1983). The relaxation protocol consisted of three steps: i) minimization with protein atoms constrained, ii) heating from 100 to 300 K with protein atoms constrained (100 ps), and iii) production with protein atoms released (600 ns). The heating step was performed in the NVT ensemble, while the production step in the NPT ensemble. Langevin dynamics was used to maintain the temperature and the pressure (where applied) constant at 300 K and 1 bar, respectively. The timestep was set to 2 fs.

To estimate the stability of each pose during the MD simulations, we computed L_RMSD, i.e. the Root Mean Square Deviation of the ligand backbone atoms of all MD frames with respect to the initial frame provided that they are aligned on the receptor atoms (Figure CP2_SI). Based on this analysis, the top pose of the 674 cluster was shown to be

the most stable along the MD simulations and was further selected as an atomistic model of the cyclic peptide with the GIP receptor complex (Figure CP1_MT A).



Figure S 12: Clustering of the cyclic peptide conformations folded in the GIP receptor's binding site. Only 50 top most populated clusters are shown. A. Clusters' populations. The most populated clusters (674, 682, 1503 and 1519) are highlighted with red arrows. B. Clusters' binding energies, i.e. the energy of binding between the cyclic peptide and the GIP receptor.



Figure S 13: L_RMSD estimated for the four molecular dynamics trajectories of the top poses extracted from the most populated clusters. Here, L_RMSD is the Root Mean Square Deviation of the ligand backbone of all MD frames with respect to the initial frame provided that they are aligned on the receptor.

Analysis of solubility



Figure S 14: Solubility of peptides as measured by Nephelometry at different PEG concentrations. The flag tag provides solubility (see B_5 vs B_5.3 and B_1275 vs B_1275.12), but low solubility of non tagged peptides can be overcome by carefully chosen substitutions in the peptide sequence (B_1275.12 vs B_1275.5, but not B_1275.13). Attachment of an albumin binder has no apparent influence on the solubility (B_5 vs B_5.1, B_68 vs B_68.1, and B_1275.5 vs B_1275.6)

Peptide solubility by PEG assay

Polyethylene glycol has been extensively used as a crowding agent to mimic the intracellular environment and as an agent to promote aggregation of biomolecules^[7] (Akabayov et al., 2013). Assays in which proteins and peptides are tested against different concentrations of PEG are routinely used to rank solubility of compounds of interest^[8] (Toprani et al., 2016). In this assay 0.5 mM of each peptides was dissolved in a wellplate at 50mM phosphate buffer at pH 7.4, 70 mM NaCl with 0, 1.25, 2.5, 5, 10 and 20 % (V/V) PEG 1000 (Sigma–Aldrich, USA). Turbidity of each well on the plate was read by a Nephelostar reader (BMG Labtech, Germany) with standard settings. Values represent the turbidity of a solution, with the lowest ones indicating a clear solution and the highest a milky one due to compound aggregation. Compound B_1275.3 did not resuspend in the buffer at all.

In vitro plasma stability and quantification of in vivo samples by LC-MS:

In vitro plasma stability of the peptides was assessed by incubation of 1 μ M peptide in 80% pooled human Li-heparin plasma (pooled and blinded, mixed male and female from BioIVT, Westbury, NY, US) and 20% PBS buffer, pH 7.4 at 37°C under shaking. At selected time points samples were taken: 5, 15, 30, 60, 120, 210, 300 mins by taking one volume of the incubations and subjected to protein precipitation using three volumes of ice-cold ethanol, followed by centrifugation and dilution of one volume of the supernatant with one volume of water containing 1% FA before LC-MS analysis.

The LC-MS analysis was carried out using a TurboFlow HPLC system from Thermo Fisher Scientific (Bremen, Germany) coupled to a Q Exactive HF Orbitrap Mass Spectrometer. The LC mobile phases consisted of A: MQ water with 5% organic solvent (50% methanol / 50% acetonitrile) and 1% formic acid and B: MQ water with 95% organic solvent (50% methanol / 50% acetonitrile) and 1% formic acid. For quantification of the *in vivo* samples a TurboFlow Cyclone 0,5 x 100 mm column from Thermo Fischer Scientific (Bremen, Germany) was used for extraction, before analytical elution on a Aeris Peptide XB-C18, 3.6µm, 2.1 x 50 mm column (controlled at 60°C) from Phenomenex (Torrence, CA) using a

flow rate of 400μ /min and a linear gradient of 30% buffer B increase over 2.5 minutes (55-85%B for B_3.1 and B_1275.1, 45-75%B for B_1275.4 and B_1275.5 and 65-95%B for B_1275.6). The samples from the *in vitro* experiments was directly loaded onto the analytical column and eluted using a linear gradient of 10-90 %B over 5 mins. The Orbitrap mass spectrometer was operating in positive ionization mode with a spray voltage of 4 kV, with a resolution of 30K using a *m*/*z* 300-1500 full scan mode for the *in vitro* samples and either SIM (single ion monitoring) or PRM (parallel reaction monitoring) scan modes using 5 *m*/*z* isolation windows and 25 NCE HCD fragmentation (PRM only) of the most abundant charge state of the individual peptides from the *in vivo* samples.

The LC-MS data was processed and quantified using the Quan Browser in the Xcalibur software from Thermo Fisher Scientific (Bremen, Germany). Plotting of data and calculation of t¹/₂ was calculated using Prism (version 8.02, GraphPad Software, Inc.).

Compound	T ¹ / ₂ - <i>In vitro</i> plasma stability						
B_3.1	Stable						
B_3.2	~2h						
B_1275	~3.5h						
B_1275.1	Stable						
B_1275.2	~2h						
B_1275.3	Stable						
B_1275.4	Stable						
B_1275.5	Stable						
B_1275.6	Stable						
Native GIP	45 min						
Native GLP-1	35 min						

Table S2: In vitro plasma stability in human plasma

Pharmacokinetic studies

The pharmacokinetic studies conducted in rats were all approved by the Danish Animal Experiments Inspectorate in accordance with European Union Directive 2010/63/EU.

Male, nonfasted awake Sprague-Dawley rats were dosed intravenously in the tail vein with the peptides of interest (2 nmol/kg; 2 µM peptide, 5 mM phosphate, 140 mM sodium chloride, 70 ppm polysorbate 20, pH 7.4) and plasma concentration-time profiles were followed for 2 days after dosing with frequent blood sampling. Plasma concentrations of the peptides were analysed by quantitative LC-MS. For quantification of the *in vivo* rat samples, selected plasma standards in the range from 0.5 to 2000 nM of the peptides were prepared. The standards were prepared by spiking blank plasma from Sprague-Dawley rats. Prior to LC-MS analysis, the plasma samples (blank plasma, standards and study samples) were prepared by plasma protein precipitation. Protein precipitations were conducted by adding three volumes of ethanol to one volume of plasma. The samples were centrifuged, and one volume of supernatant was mixed with three volumes of water containing 1% FA. LC-MS analysis was conducted as described above with a typical lower limit of quantification of 1-2 nM for the *in vivo* samples.

Pharmacokinetic analysis

Plasma concentration-time profiles were analysed by non-compartmental PK analysis (NCA) using the WinNonlin software (Certara, Ca). Calculations were performed using individual concentration-time values from each animal. Uniform weighting was used for estimation of the terminal rate constant (λ_z) and the elimination half-life was calculated as $t_{\lambda} = \ln 2/\lambda_z$.

Compound		t%	t½	MRT	MRT	Dose (pmol	Weight	C₀ (pmol	AUC (min*	AUC/Dose (min*	AUCextrap	CI	Vz
		(min)	(h)	(min)	(h)	/kg)	(kg)	/L)	pmol/L)	kg/L)	(%)	(L/min/kg)	(L/kg)
													0,14288
B3.1	mean	585	9,75	738	12,30	2017	0,422	39517	12056069	5977	15,8	0,000169	3
		[561-	[9,35 -	[713-	[11,88 -								0,02116
	SD	651]	10,85]	797]	13,28]	20	0,017	3716	1200826	589	1,3	0,000016	6
													0,11652
B1275.1	mean	643	10,72	821	13,68	1993	0,435	38820	15863409	7962	18,6	0,000126	9
		[553-	[9,22 -	[743-	[12,38 -								0,00961
	SD	724]	12,07]	906]	15,10]	15	0,023	4659	683139	360	2,4	0,000006	9
													0.22734
B1275.4	mean	3,46	0,06	4,45	0,07	9975	0,439	58510	228078	23	6,7	0,045705	6
		[2,87-	[0.05 -	[3.52-	[0.06 -								0.06582
	SD	4.31]	0,07]	5.80]	0,10]	121	0,036	22040	51410	5	5,6	0,010866	3
				-									0.26352
B1275.5	mean	1.82	0.03	2.62	0.04	9979	0.423	40597	104496	10	6.5	0.100370	4
		[1.54-	[0.03 -	[2.23-	[0.04 -								0.08068
	SD	2.22]	0.041	3.201	0.051	121	0.044	11062	24204	2	2.6	0.025407	4
		,	.,		.,,		.,			_	,-	.,	0.06475
B1275.6	mean	250	4.17	325	5.42	2013	0.441	54317	11157944	5544	14.2	0.000184	1
212/3.0	cun	[182-	[3 03 -	[238-	[3 97 -	_515	0,741	5.517		2244	_ ,	0,000104	0.01168
	SD	352	5 871	412]	6 871	9	0.024	2602	1773929	889	84	0.000028	4
	55	552]	3,37]	.12]	5,07]	5	3,324	2302	1,,5525	555	5,7	0,000020	-

Supplementary Info References

- T. E. McAllister, T. L. Yeh, M. I. Abboud, I. K. H. Leung, E. S. Hookway, O. N. F. King,
 B. Bhushan, S. T. Williams, R. J. Hopkinson, M. Munzel, N. D. Loik, R. Chowdhury, U.
 Oppermann, T. D. W. Claridge, Y. Goto, H. Suga, C. J. Schofield, A. Kawamura, *Chem Sci* 2018, 9, 4569-4578.
- [2] S. J. Fleishman, A. Leaver-Fay, J. E. Corn, E. M. Strauch, S. D. Khare, N. Koga, J. Ashworth, P. Murphy, F. Richter, G. Lemmon, J. Meiler, D. Baker, *PLoS One* 2011, 6, e20161.
- [3] aE. A. Killion, J. Wang, J. Yie, S. D. Shi, D. Bates, X. Min, R. Komorowski, T. Hager, L. Deng, L. Atangan, S. C. Lu, R. J. M. Kurzeja, G. Sivits, J. Lin, Q. Chen, Z. Wang, S. A. Thibault, C. M. Abbott, T. Meng, B. Clavette, C. M. Murawsky, I. N. Foltz, J. B. Rottman, C. Hale, M. M. Veniant, D. J. Lloyd, *Sci Transl Med* 2018, *10*; bC. Parthier, M. Kleinschmidt, P. Neumann, R. Rudolph, S. Manhart, D. Schlenzig, J. Fanghanel, J. U. Rahfeld, H. U. Demuth, M. T. Stubbs, *Proc Natl Acad Sci U S A* 2007, *104*, 13942-13947.
- [4] aE. A. Coutsias, C. Seok, M. P. Jacobson, K. A. Dill, *J Comput Chem* 2004, 25, 510-528;
 bD. J. Mandell, E. A. Coutsias, T. Kortemme, *Nat Methods* 2009, 6, 551-552.
- [5] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J Chem Theory Comput* **2015**, *11*, 3696-3713.
- [6] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J Chem Phys* 1983, 79, 926-935.
- [7] B. Akabayov, S. R. Akabayov, S. J. Lee, G. Wagner, C. C. Richardson, *Nat Commun* **2013**, *4*, 1615.
- [8] V. M. Toprani, S. B. Joshi, L. A. Kueltzo, R. M. Schwartz, C. R. Middaugh, D. B. Volkin, J Pharm Sci-Us 2016, 105, 2319-2327.

Analytical data on purified peptides











H₃C H₃ H₃C H₃C D D D K-NH₂



























H₃C H₃ C H₃





SFQYFHILW PF SGSGSDYK DDDDK_NH₂













Report_v01/Integration

on (c) Version 7.2.10.24543















