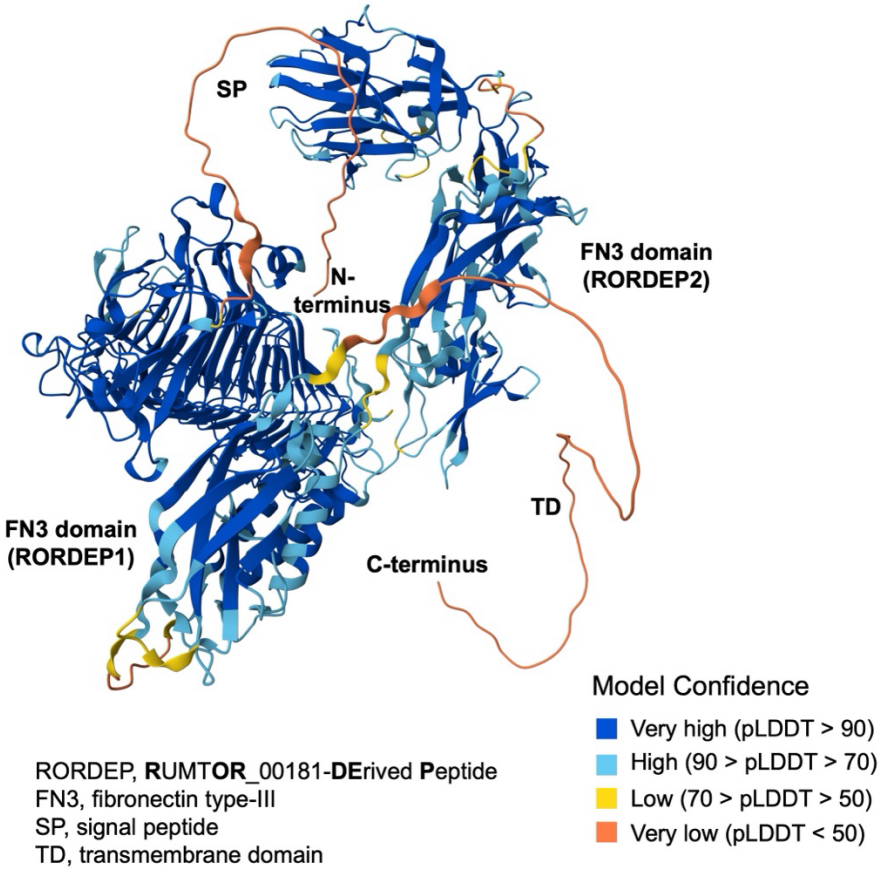
nature microbiology

Supplementary information

https://doi.org/10.1038/s41564-025-02064-x

Polypeptides synthesized by common bacteria in the human gut improve rodent metabolism

In the format provided by the authors and unedited

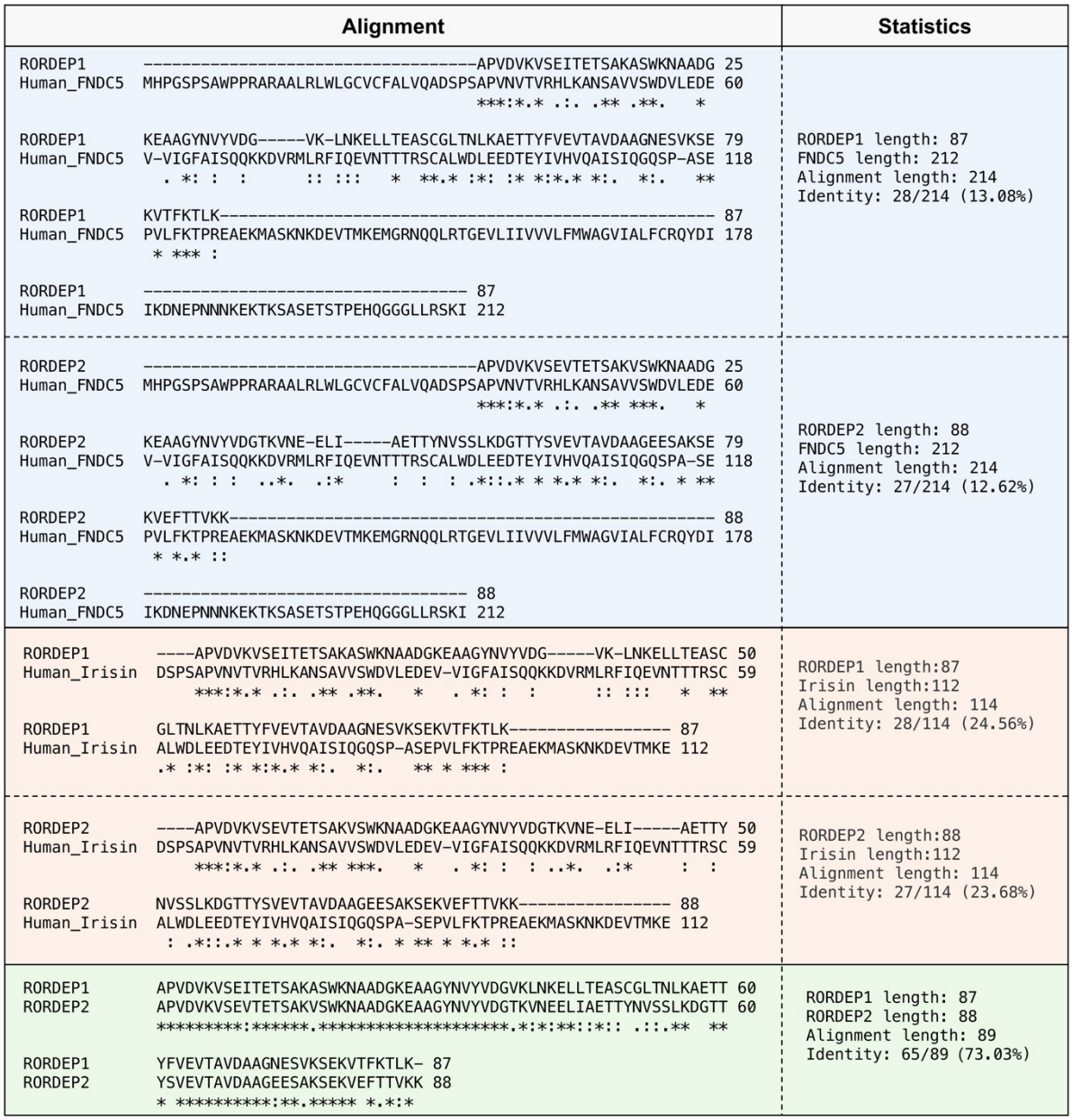


Supplementary Figure 1 | Confidence of three-dimensional structure of RUMTOR_00181 protein downloaded from AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/entry/A5KIY5). The structure is comprehensively colored to represent model confidence, using the per-residue confidence metric known as pLDDT (predicted Local Distance Difference Test). The pLDDT scores, which range from low (1) to high (100) confidence, indicate the reliability of each modeled residue based on the LDDT-Cα metric. The model highlights the specific regions of interest, such as the two RUMTOR_00181-DErived Peptides (RORDEPs), transmembrane domain (TD), and signal peptide (SP).

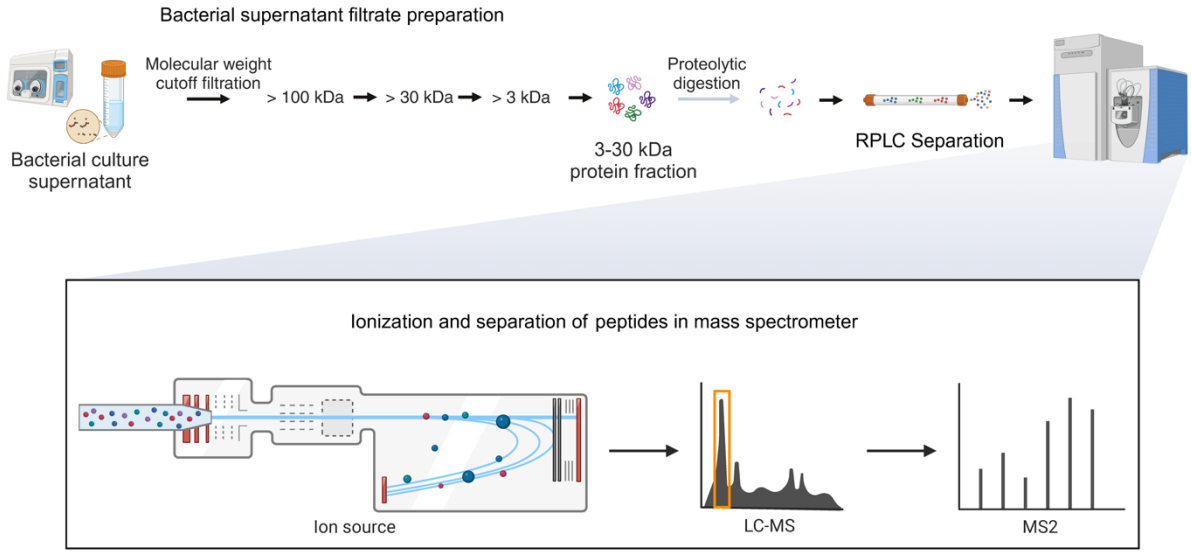
a


sp|Q5VTL7|FNDC7_HUMAN Fibronectin type III domain-containing protein 7 OS=Homo sapiens OX=9606 GN=FNDC7 PE=2 SV=3
sp|Q9H6D8|FNDC4_HUMAN Fibronectin type III domain-containing protein 4 OS=Homo sapiens OX=9606 GN=FNDC4 PE=2 SV=1
tr|A0A3B3ITP1|A0A3B3ITP1_HUMAN Fibronectin type III domain containing 5 OS=Homo sapiens OX=9606 GN=FNDC5 PE=1 SV=2
sp|Q8NAU1|FNDC5_HUMAN Fibronectin type III domain-containing protein 5 OS=Homo sapiens OX=9606 GN=FNDC5 PE=1 SV=3
tr|A0A0A0MRR6|A0A0A0MRR6_HUMAN Fibronectin type III domain-containing protein 5 OS=Homo sapiens OX=9606 GN=FNDC5 PE=1 SV=1
RORDEP1
sp|Q58EX2|SDK2_HUMAN Protein sidekick-2 OS=Homo sapiens OX=9606 GN=SDK2 PE=1 SV=3
tr|H7C2P2|H7C2P2_HUMAN Sidekick cell adhesion molecule 2 (Fragment) OS=Homo sapiens OX=9606 GN=SDK2 PE=1 SV=1
tr|F8W6X9|F8W6X9_HUMAN Sidekick cell adhesion molecule 1 OS=Homo sapiens OX=9606 GN=SDK1 PE=1 SV=1
sp|Q7Z5N4|SDK1_HUMAN Protein sidekick-1 OS=Homo sapiens OX=9606 GN=SDK1 PE=2 SV=3
sp|Q02763|TIE2_HUMAN Angiopoietin-1 receptor OS=Homo sapiens OX=9606 GN=TEK PE=1 SV=2
tr|D6RGG3|D6RGG3_HUMAN Collagen type XII alpha 1 chain OS=Homo sapiens OX=9606 GN=COL12A1 PE=1 SV=1
sp|Q99715|COCA1_HUMAN Collagen alpha-1(XII) chain OS=Homo sapiens OX=9606 GN=COL12A1 PE=1 SV=2
tr|A0A0A0MTS7|A0A0A0MTS7_HUMAN Titin OS=Homo sapiens OX=9606 GN=TTN PE=1 SV=1
tr|C0JYZ2|C0JYZ2_HUMAN Titin OS=Homo sapiens OX=9606 GN=TTN PE=3 SV=1
tr|A0A0A0MRA3|A0A0A0MRA3_HUMAN Titin OS=Homo sapiens OX=9606 GN=TTN PE=1 SV=1
sp|Q8WZ42|TITIN_HUMAN Titin OS=Homo sapiens OX=9606 GN=TTN PE=1 SV=4
tr|Q2HXI4|Q2HXI4_HUMAN protein-tyrosine-phosphatase OS=Homo sapiens OX=9606 GN=PTPRD PE=2 SV=1
sp|P23468|PTPRD_HUMAN Receptor-type tyrosine-protein phosphatase delta OS=Homo sapiens OX=9606 GN=PTPRD PE=1 SV=2
tr|A0A140T9L7|A0A140T9L7_HUMAN Tenascin-X (Fragment) OS=Homo sapiens OX=9606 GN=TNXB PE=1 SV=1
tr|A0A087X0I0|A0A087X0I0_HUMAN Tenascin XB (Fragment) OS=Homo sapiens OX=9606 GN=TNXB PE=1 SV=1
tr|A0A140T8Z8|A0A140T8Z8_HUMAN Tenascin-X (Fragment) OS=Homo sapiens OX=9606 GN=TNXB PE=1 SV=1
tr|A0A3B3ISX9|A0A3B3ISX9_HUMAN Tenascin XB OS=Homo sapiens OX=9606 GN=TNXB PE=1 SV=1
sp|P22105|TENX_HUMAN Tenascin-X OS=Homo sapiens OX=9606 GN=TNXB PE=1 SV=5
tr|A0A140TA33|A0A140TA33_HUMAN Tenascin-X OS=Homo sapiens OX=9606 GN=TNXB PE=1 SV=1
tr|A0A140TA41|A0A140TA41_HUMAN Tenascin-X OS=Homo sapiens OX=9606 GN=TNXB PE=1 SV=1
tr|A0A140TA52|A0A140TA52_HUMAN Tenascin-X OS=Homo sapiens OX=9606 GN=TNXB PE=1 SV=1
tr|A0A140T8Y3|A0A140T8Y3_HUMAN Tenascin-X OS=Homo sapiens OX=9606 GN=TNXB PE=1 SV=1
tr|A0A140T902|A0A140T902_HUMAN Tenascin-X OS=Homo sapiens OX=9606 GN=TNXB PE=1 SV=1
tr|A0A140T9C0|A0A140T9C0_HUMAN Tenascin-X OS=Homo sapiens OX=9606 GN=TNXB PE=1 SV=1
tr|Q6IPK3|Q6IPK3_HUMAN TNXB protein (Fragment) OS=Homo sapiens OX=9606 GN=TNXB PE=2 SV=1
0.5

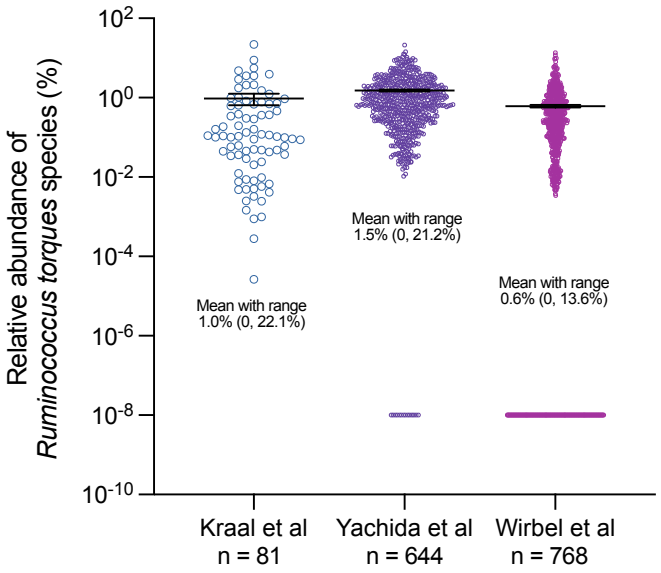
Supplementary Figure 2 | Phylogenetic analyses of protein sequences of RORDEP1 and 2, and known human FN3 domain proteins. The illustration displays phylogenetic trees representing the evolutionary relationships between RORDEP1 (**a**), RORDEP2 (**b**) and known FN3 domain proteins. A total of 1,077 public protein sequences of known FN3 domain proteins were subjected to analyses to generate the phylogenetic trees. Sequences with sequence identity > 10% to either RORDEP1 or RORDEP2 were visualized in this figure. Thirty sequences aligned with RORDEP1 (panel **a**), and 25 sequences aligned with RORDEP2 (panel **b**). The phylogenetic trees are constructed based on sequence homology, with branch lengths correlating to the degree of divergence. Sequences are denoted by their accession numbers and protein IDs, with RORDEPs highlighted in yellow. FNDC5 is highlighted in orange. Scale bars marked with grey background denote tree distances by the Grishin model[1].



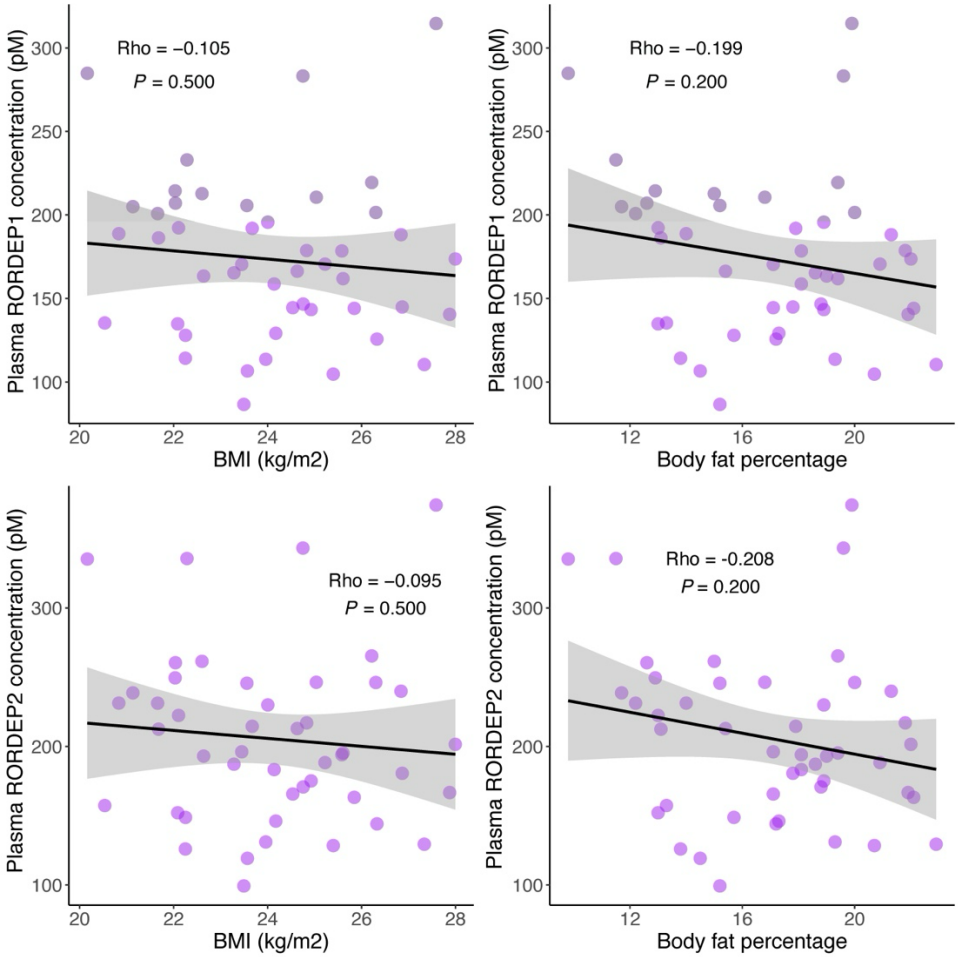
Supplementary Figure 3 | Cross-alignment of amino acid sequences between RORDEPs, human fibronectin type-III domain containing protein 5 (FNDC5), and human irisin. The multiple sequence alignments were performed using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) to determine the positions of the cysteines and the number of the identical and conserved residues. In the left panel, identical residues between the two sequences are denoted by asterisks; low and high degrees of similarity are represented by a period and a colon, respectively. In the right panel, the identity percentage is calculated as the ratio of identical amino acid residues to the total alignment length.



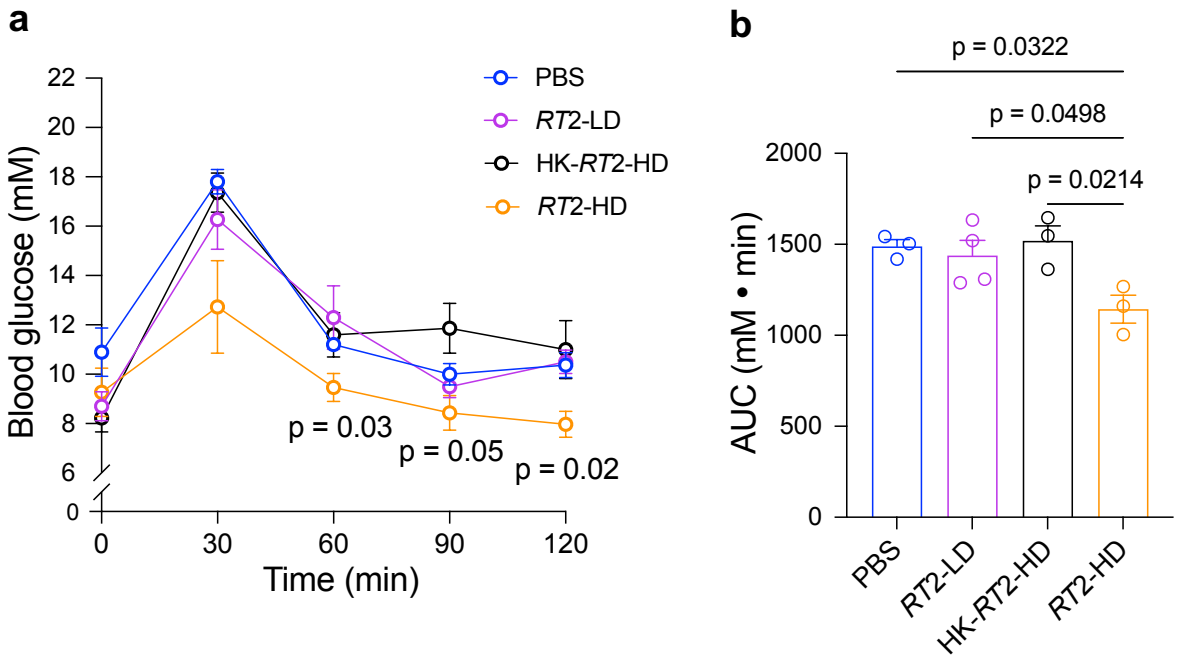
Supplementary Figure 4 | Workflow for preparation and analysis of RORDEPs peptides in bacterial culture supernatants using liquid chromatography tandem mass spectrometry (LC-MS). In the top panel, a series of molecular weight cutoff filtrations were performed. Two unique proteotypic peptides were chosen to unambiguously identify RORDEP1 and RORDEP2, respectively. Given the molecular weights of RORDEPs (approximately 10 kDa), we selected the 3-30 kDa fractions from culture supernatants of *R. torques* strains to develop the assay. These fractions were then subjected to proteolytic digestion by trypsin overnight, followed by reversed-phase liquid chromatography (RPLC) for effective peptide separation. The bottom panel illustrates the data acquisition process in mass spectrometry. This involves ionizing the peptides separated by RPLC for comprehensive analysis in both MS and MS^2 modes. This illustration was created with Biorender (https://BioRender.com/xmbrn69).



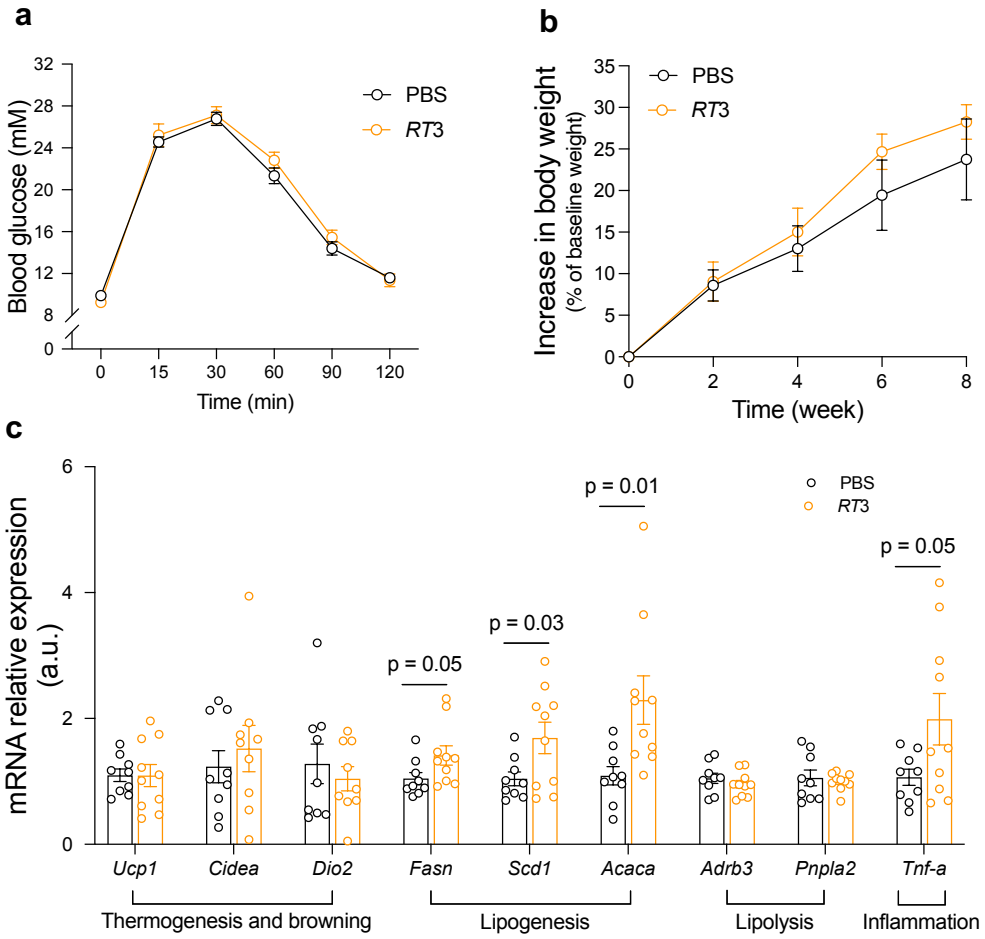
Supplementary Figure 5 | Relative abundance of *R. torques* species in publicly available databases. Relative abundance of *R. torques* at species level as a percentage of total gut microbiota species across three reported independent studies displayed as jittered strip plots with mean values and standard deviations. Each point represents an individual sample with the total number of samples (n) indicated below each study name. Studies are differentiated by color and include data from a cohort from the NIH Human Microbiome Project as reported by Kraal et al[2]. (blue, n = 81), a Japanese cohort as reported by Yachida et al[3]. (purple, n = 644), and a cohort from eight different cohorts in America, Europe, and Asia as reported by Wirbel et al[4]. (pink, n = 768). The Y-axis is log-scaled to visualize data across a range of abundances. Zero abundance was set to 10^{-8} for visualization.



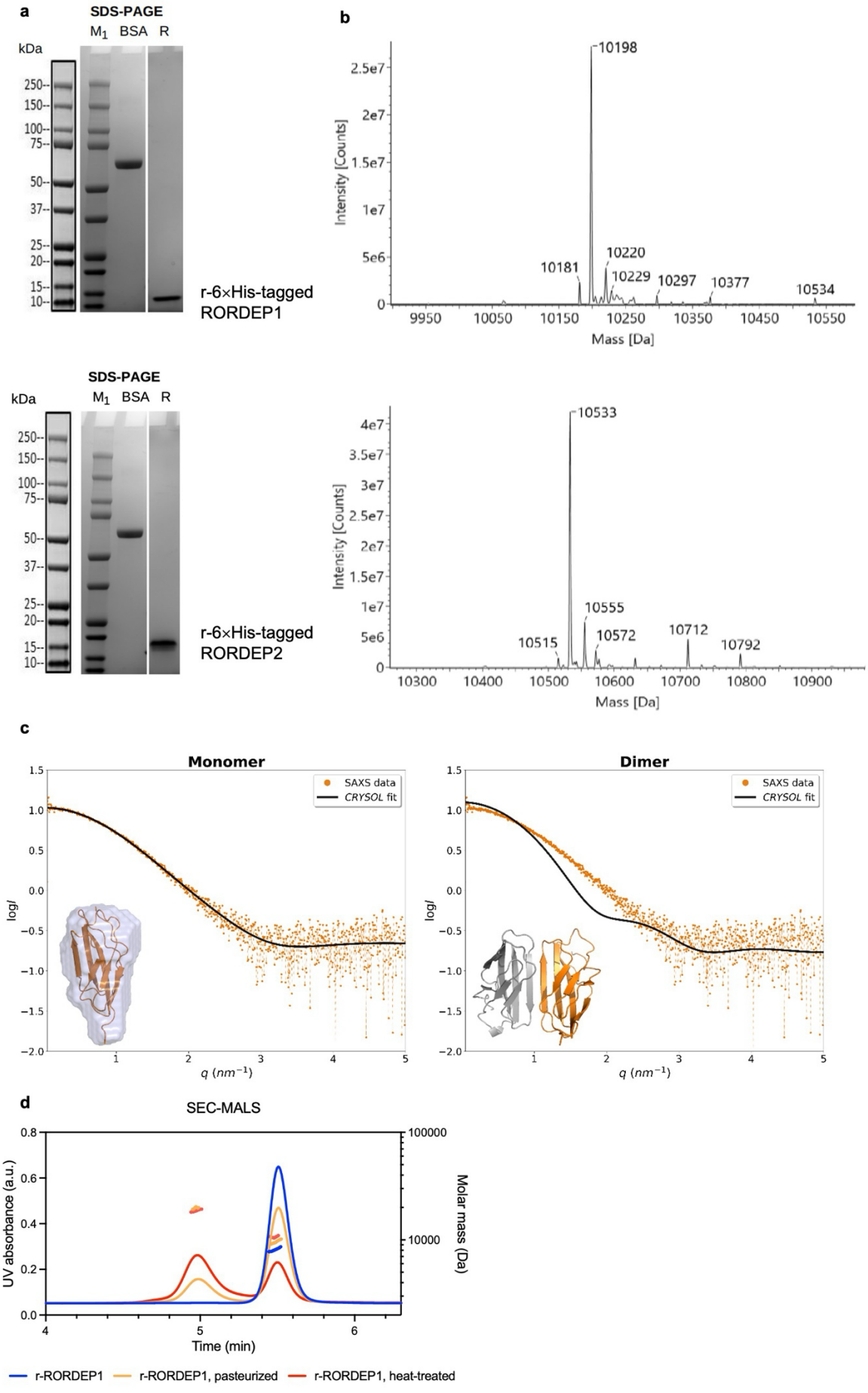
Supplementary Figure 6 | Lack of relationships between plasma concentrations of RORDEPs and body mass index (BMI) or body fat percentage in 46 Danish healthy adult males. This figure displays that there is no significant correlation between plasma RORDEP1 or RORDEP2 and either BMI or body fat percentage in the examined study group. The correlations were assessed by Pearson's two-sided correlation tests. Error bands represent the linear regression line (black solid) and the 95% confidence interval (grey shaded).



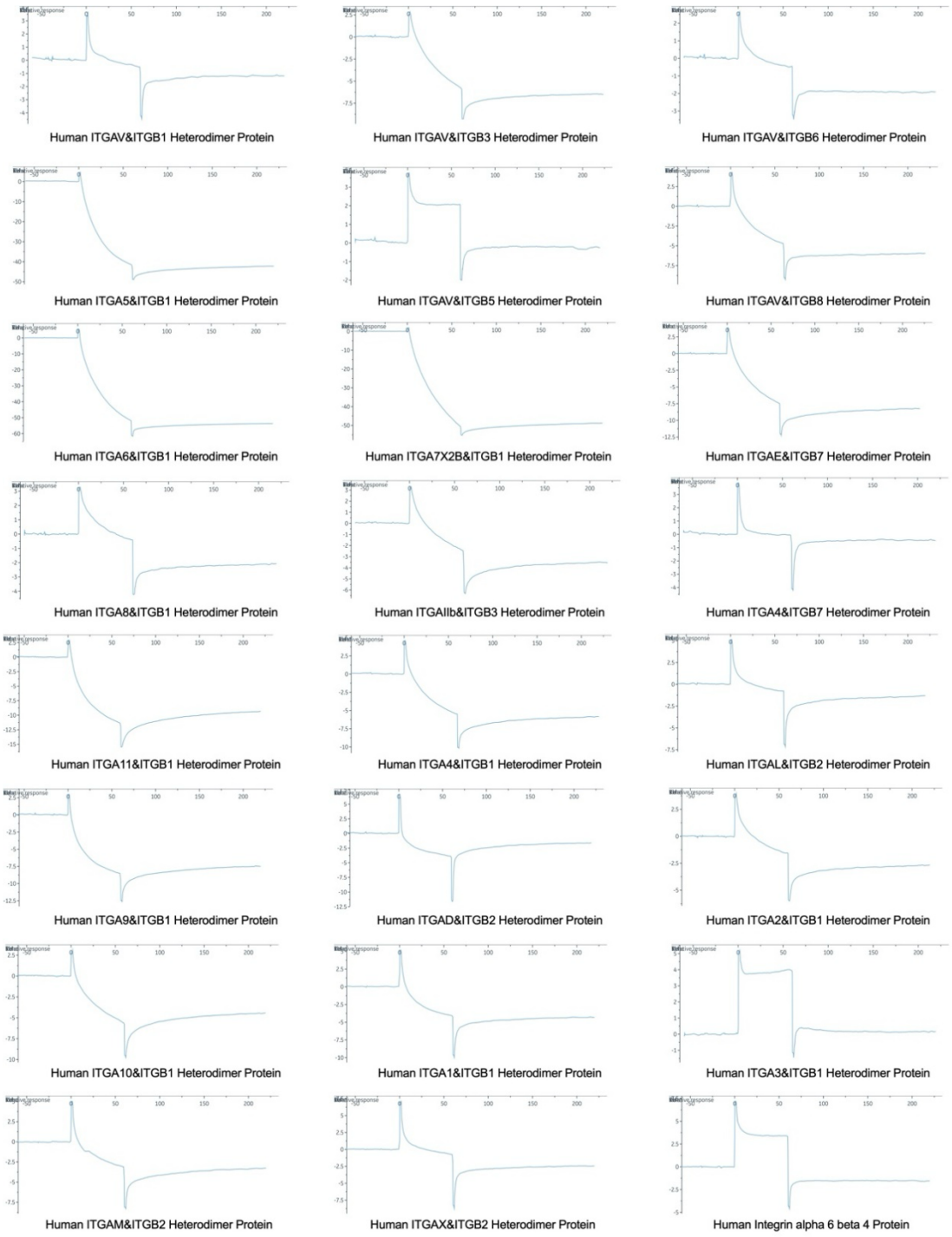
Supplementary Figure 7 | Feasibility study testing effects of two different doses of a live or one dose of heat-killed *R. torques* (RT) strain expressing the RUMTOR_00181 gene on intraperitoneal glucose tolerance in six weeks old lean male C57BL/6N mice. Chow-fed mice that were six weeks old at study start were given bacteria or phosphate-buffered saline (PBS) by oral gavage twice weekly for six weeks. **(a)** Intraperitoneal glucose tolerance test curves at end of interventions. **(b)** Area under the curve (AUC) for glucose tolerance test (n = 3-4 in each group). *RT2*-LD is *RT* ATCC 27756 that carry the gene expressing RUMTOR_00181 at a low dose (LD) of 5×10^7 CFUs/100 µL sterile phosphate-buffered saline (PBS) containing 10% glycerol; *RT2*-HD, is *RT* ATCC 27756 at a higher dose (HD) of 5×10^8 CFUs/100 µL; HK-*RT2*-HD is heat-killed *RT* ATCC 27756 at a dose of 5×10^8 CFUs/100 µL. Data are presented as mean±SEM. Statistical analysis employed two-way ANOVA with Bonferroni *post hoc* correction for panel **(a)**, and one-way ANOVA with Dunnett's *post hoc* correction for panel **(b)**.



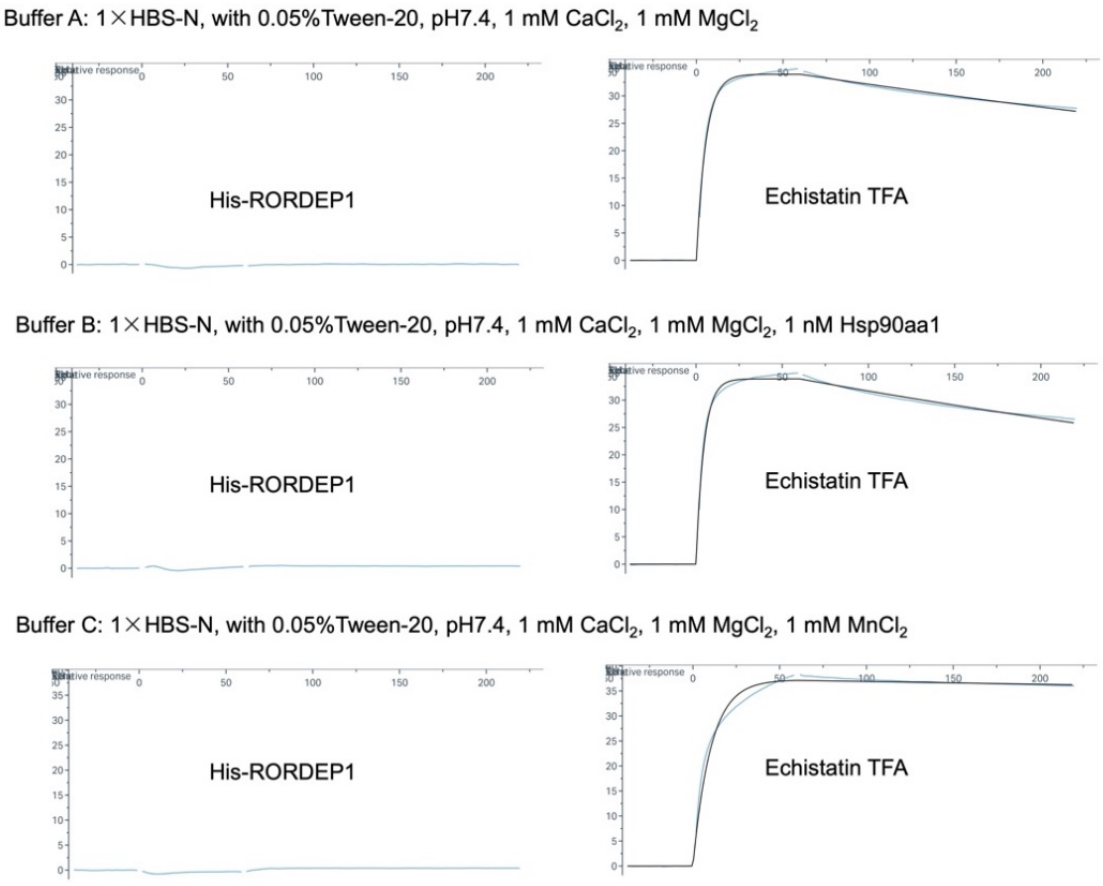
Supplementary Figure 8 | Intervention with *Ruminococcus torques* ATCC 35915 (RT3 strain), lacking the RUMTOR_00181 gene but expressing virulence factors, over eight weeks in high-fat-fed mice. Obese mice were administered live RT3 orally via gavage twice weekly at a dose of 5×10^9 colony-forming units per 100 µl of 10% glycerol in sterile phosphate-buffered saline (PBS), control group received sterile PBS containing 10% glycerol. The intervention had a duration of eight weeks. **(a)** Assessment of glucose tolerance at week six of intervention was done with an intraperitoneal glucose tolerance test. **(b)** Relative body weight changes in mice over the 8-week period. Weight increase was normalized to baseline at week 0. **(c)** mRNA expression levels of genes related to thermogenesis, lipogenesis, lipolysis, and inflammation in inguinal white adipose tissue across groups. Data are presented as mean ± SEM (n = 9 mice per group). Statistical analysis was performed by Student's *t*-test.


a
SDS-PAGE
M1 BSA R
kDa
250--
150--
100--
75--
50--
37--
25--
20--
15--
10--
r-6×His-tagged RORDEP1
b
10198
10181
10220
10229
10297
10377
10534
Intensity [Counts]
Mass [Da]
SDS-PAGE
M1 BSA R
kDa
r-6×His-tagged RORDEP2
10533
10515
10555
10572
10712
10792
Intensity [Counts]
Mass [Da]
c
Monomer
Dimer
SAXS data
CRYSOL fit
log I
q (nm−1)
d
SEC-MALS
UV absorbance (a.u.)
Molar mass (Da)
Time (min)
r-RORDEP1
r-RORDEP1, pasteurized
r-RORDEP1, heat-treated

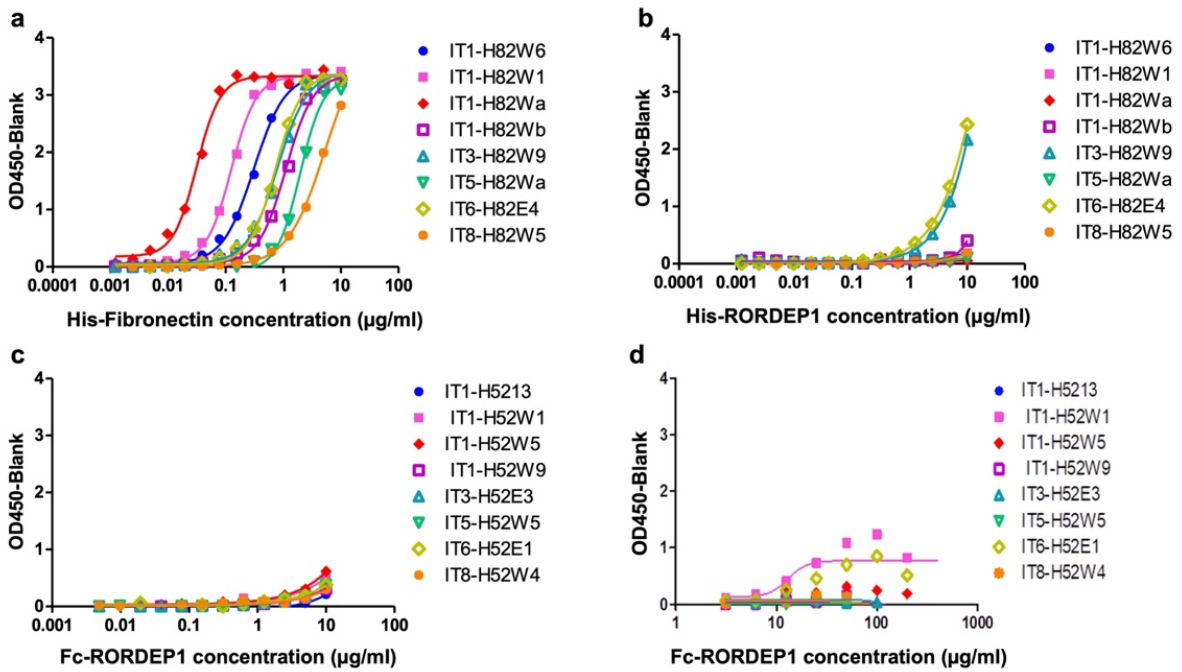
Supplementary Figure 9 | Expression and biophysical characterization of recombinant RORDEP1 and recombinant RORDEP2. (a) Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis displaying the purity and molecular weight of 6×His-tagged r-RORDEP1 and 6×His-r-tagged RORDEP2 polypeptides. Lane M1 represents the protein molecular weight marker, with sizes indicated in kilodaltons (kDa). Lane BSA contains pure bovine serum albumin as a concentration and purity reference. Lane R shows r-6×His-tagged RORDEP1 protein under reducing conditions, indicating its relative molecular mass. The lower part of panel **(a)** shows a similar SDS-PAGE analysis for r-6×His-tagged RORDEP2 protein. The experiment was repeated for n > 3 times independently with similar results. **(b)** Deconvoluted mass spectra of 6×His-tagged r-RORDEP1 (top) and 6×His-tagged r-RORDEP2 (bottom) obtained from high-resolution mass spectrometry. The spectra display the observed mass-to-charge (m/z) ratios for the proteins, with the corresponding calculated molecular mass (Da) labeled for prominent peaks. These results confirm the successful expression and purity of the recombinant proteins, with masses aligning closely with the expected theoretical values (10,198 Da for RORDEP1, and 10,534 for RORDEP2), confirming the identity and integrity of the 6×His-tagged r-RORDEP1 and 6×His-tagged r-RORDEP2 polypeptides. **(c)** Characterization of the polymerization of 6×His-tagged r-RORDEP1 was conducted by fitting the theoretical scattering data of the monomer (based on PDB code: 4LSD, chain A) and dimer (derived from chains A and B) by CRYSOL 2 modeling to the experimental data obtained from Size Exclusion Chromatography coupled with Small Angle X-ray Scattering (SEC-SAXS). In the graphical representation, the theoretical scattering curves are depicted in black, while the experimental scattering curves are shown in orange. **(d)** Thermostability characterization of 6×His-tagged r-RORDEP1 by multi-angle light scattering (SEC-MALS) analysis. The left Y-axis corresponds to UV absorption, while the right Y-axis indicates the molecular weight detected.



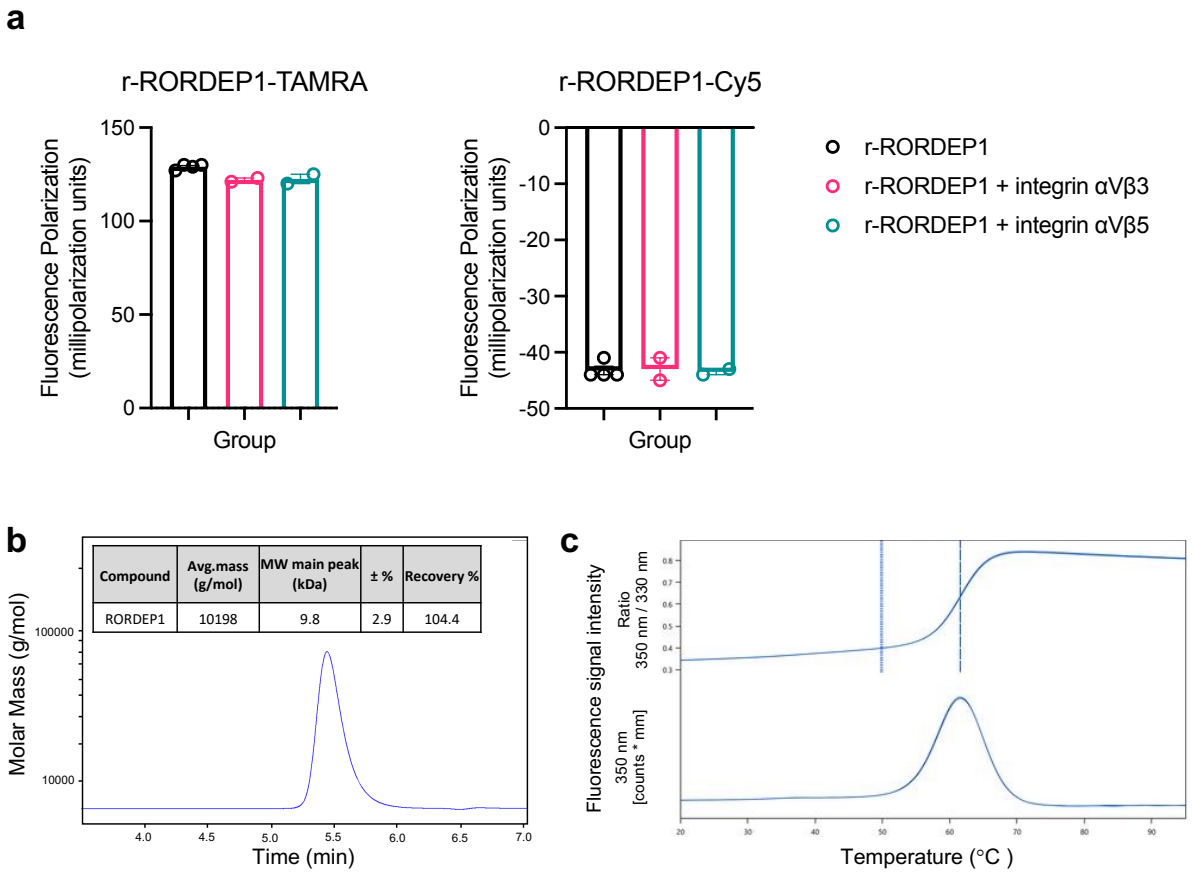
Supplementary Figure 10 | Use of Surface Plasmon Resonance (SPR) assay testing single-concentration affinities between Fc-r-RORDEP1, and 24 human integrin heterodimers showing no evidence of binding between Fc-r-RORDEP1 and integrins. Sensograms are shown. In this experimental setting, the SPR chip captured and immobilized Fc-r-RORDEP1 (10 μg/mL), and 24 human integrin candidates (1000 nM) were added to the system, followed by plate reading by SPECTROstarNano (BMG LABTECH) plate reader. X-axis indicates experimental running time, and y-axis indicates the intensity of signals. Fc denotes fragment crystallizable region.



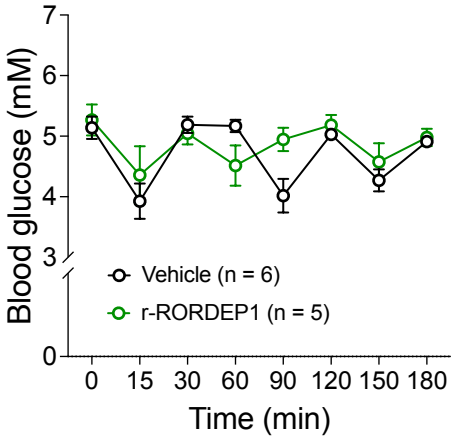
Supplementary Figure 11 | Sensorgrams of affinity testing by Surface Plasmon Resonance (SPR) assay between His-tagged recombinant RORDEP1, Echistatin TFA and biotinylated human ITGAV&ITGB5 heterodimer protein in the presence of three different buffer solutions shows no significant binding. The SPR chip captured and immobilized biotinylated human ITGAV&ITGB5 (20 μg/mL), and r-His-RORDEP1 (1000 nM) or Echistatin TFA (1000 nM) were added to the system, followed by plate reading by SPECTROstarNano (BMG LABTECH) plate reader. X-axis indicates experimental running time, and y-axis indicates the intensity of signals.



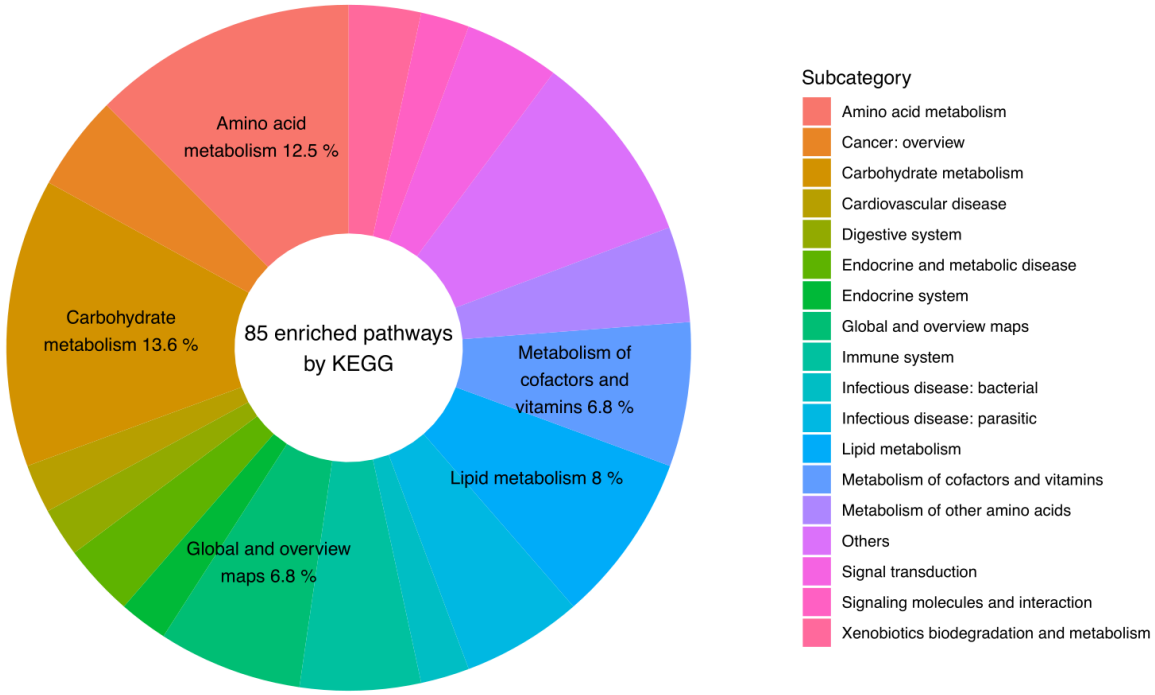
Supplementary Figure 12 | Use of ELISA-based affinity tests show no binding of recombinant RORDEP1 to selected integrin heterodimers. (a) Binding curves of His-Fibronectin (positive control) towards eight selected integrin receptors (10 μg/mL). **(b)** Binding curves of His-r-RORDEP1 towards eight selected integrin receptors (10 μg/mL) showing negligible affinity. **(c)** Binding curves of eight selected integrin receptors (coating protein, 5 μg/mL) towards Fc-r-RORDEP1 (running protein) with increased concentrations, showing minimal binding. **(d)** Binding curves of Fc-r-RORDEP1 (coating protein, 5 μg/mL) with various concentrations towards eight selected integrin receptors (running protein), showing no noteworthy binding. Fc denotes fragment crystallizable region.



Supplementary Figure 13 | Use of a fluorescence polarization (FP) assay shows no binding between recombinant RORDEP1 and integrin αVβ3 or αVβ5 integrin heterodimers in the presence of $MnCl_2$. In addition, tests of recombinant RORDEP1 folding and thermostability. **(a)** FP signal intensity in the binding between integrin αVβ3 (340 nM, n = 2) or αVβ5 (500 nM, n = 2) receptors and r-RORDEP1 when labelled with 5-carboxytetramethylrhodamine (5-TAMRA, 70 nM) and Cyanine-5 (Cy5, 70 nM), respectively. **(b)** Size exclusion chromatography with multi-angle static light scattering confirms proper folding of r-RORDEP1. **(c)** Thermostability test of r-RORDEP1 by minituarized differential scanning fluorimetry (nanoDSF) showing a melting temperature of 60°C. The y-axis displays fluorescence intensity at 350 nm and the 350/330 ratio, while the x-axis represents the temperature of the nanoDSF system.

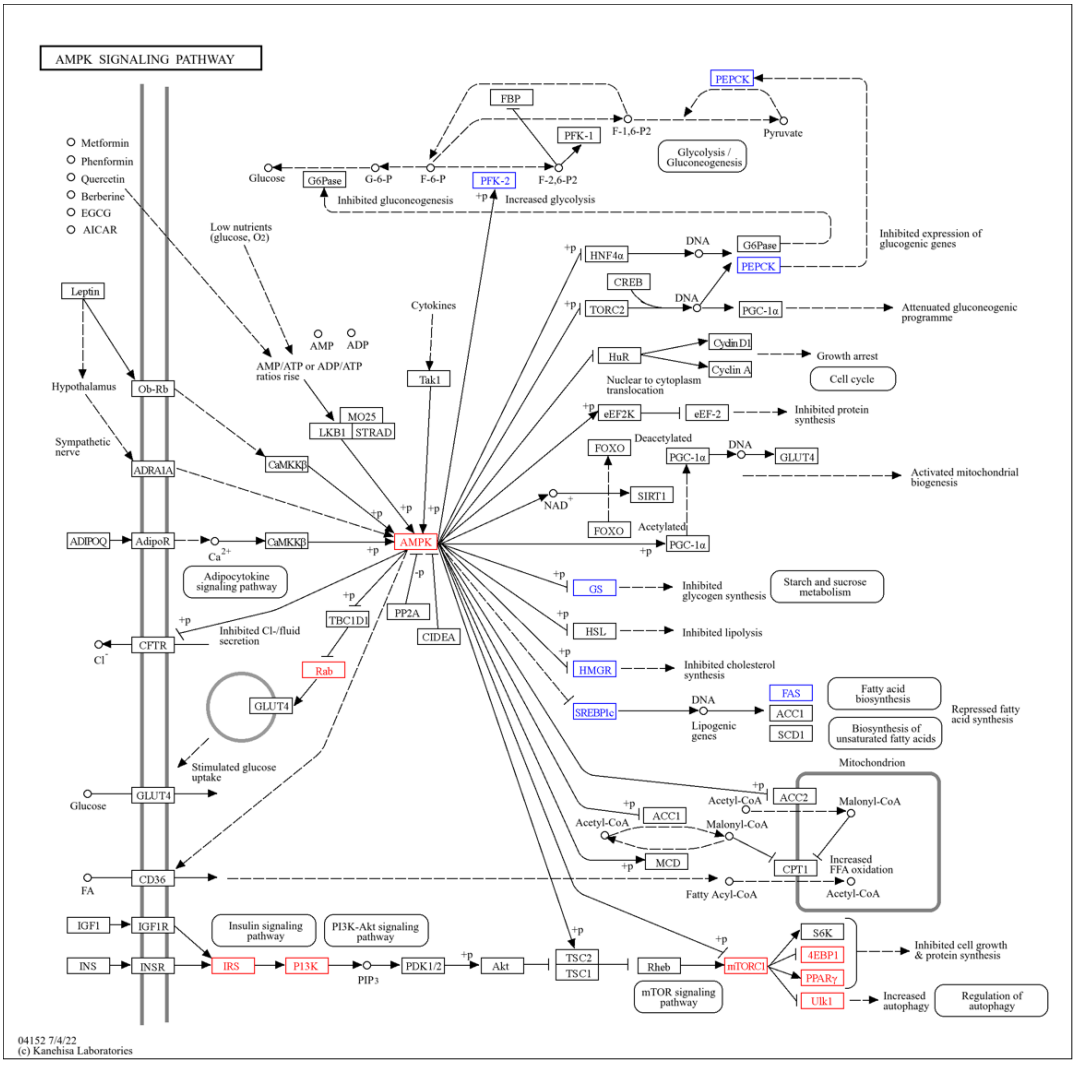


Supplementary Figure 14 | Intravenous infusion of recombinant RORDEP1 for three hours has no effect on blood glucose concentration in rats. The figure demonstrates the lack of impact of continuous intravenous infusion of r-RORDEP1 (n = 5) at a dose of 200 pmol/kg/min or infusion of phosphate-buffered saline (vehicle, n = 6) at a similar infusion rate for three hours on blood glucose in rats. The data points represent mean blood glucose values ± standard error of the mean (SEM) for each time point.

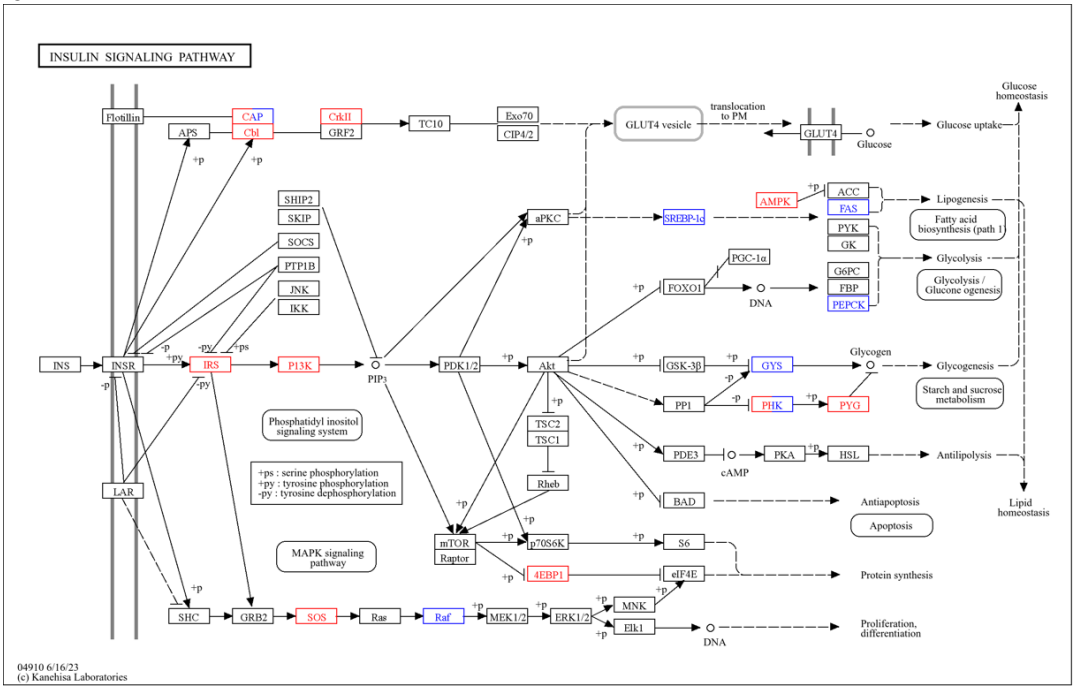


Supplementary Figure 15 | Categorical composition of Kyoto Encyclopedia of Genes and Genomes (KEGG)-enriched pathways based on differentially expressed genes in liver from rats following duodenal infusion of recombinant RORDEP1 or phosphate buffered saline for three hours. Twelve 8-week-old lean male Sprague-Dawley rats (six after r-RORDEP1 and six after infusion of phosphate buffered saline (PBS)) were included. r-RORDEP1 was infused into duodenum with a rate of 200 pmol/kg/min for three hours. A total of 2,145 differentially expressed genes were significantly enriched (*padj* (Benjamini-Hochberg) < 0.05) annotated to 85 distinct KEGG pathways, further categorized into 18 KEGG subcategories, all depicted in the donut chart with the color showing different subcategories; and area of each color showing the proportion of pathways within a specific subcategory relative to the total number of enriched pathways. The top five predominant subcategories are highlighted in the visualization.

a


AMPK SIGNALING PATHWAY
Metformin
Phenformin
Quercetin
Berberine
EGCG
AICAR
Leptin
Hypothalamus
Ob-Rb
Sympathetic nerve
ADRA1A
ADIPOQ
AdipoR
Adipocytokine signaling pathway
CPTR
Cl-
Inhibited Cl-/fluid secretion
Glucose
G6Pase
G-6-P
F-6-P
PFK-2
F-2,6-P2
FBP
PFK-1
F-1,6-P2
PEPCK
Pyruvate
Glycolysis / Gluconeogenesis
Inhibited gluconeogenesis
Increased glycolysis
Low nutrients (glucose, O2)
AMP
ADP
AMP/ATP or ADP/ATP ratios rise
Cytokines
Tak1
MO25
LKB1
STRAD
CaMKKβ
CaMKKβ
Ca2+
AMPK
PP2A
TBC1D1
CIDEA
Rab
GLUT4
Stimulated glucose uptake
Glucose
GLUT4
FA
CD36
IGF1
IGF1R
INS
INSR
Insulin signaling pathway
PI3K-Akt signaling pathway
IRS
PI3K
PIP3
PDK1/2
Akt
TSC2
TSC1
Rheb
mTOR signaling pathway
mTORC1
S6K
4EBP1
PPARγ
Ulk1
Inhibited cell growth & protein synthesis
Increased autophagy
Regulation of autophagy
HNF4α
DNA
G6Pase
PEPCK
Inhibited expression of glucogenic genes
CREB
TORC2
PGC-1α
Attenuated gluconeogenic programme
HuR
Cyclin D1
Cyclin A
Growth arrest
Cell cycle
Nuclear to cytoplasm translocation
eEF2K
eEF-2
Inhibited protein synthesis
FOXO
Deacetylated
PGC-1α
GLUT4
Activated mitochondrial biogenesis
NAD+
SIRT1
FOXO
Acetylated
PGC-1α
GS
Inhibited glycogen synthesis
Starch and sucrose metabolism
HSL
Inhibited lipolysis
HMGR
Inhibited cholesterol synthesis
SREBP1c
Lipogenic genes
FAS
ACC1
SCD1
Fatty acid biosynthesis
Biosynthesis of unsaturated fatty acids
Repressed fatty acid synthesis
Mitochondrion
Acetyl-CoA
ACC2
Malonyl-CoA
ACC1
MCD
CPT1
Increased FFA oxidation
Fatty Acyl-CoA
04152 7/4/22
(c) Kanehisa Laboratories

b


INSULIN SIGNALING PATHWAY
Flotillin
APS
CAP
Cbl
CrkII
GRF2
TC10
Exo70
CIP4/2
GLUT4 vesicle
translocation to PM
GLUT4
Glucose
Glucose homeostasis
Glucose uptake
SHIP2
SKIP
SOCS
PTP1B
JNK
IKK
aPKC
SREBP-1c
AMPK
ACC
FAS
Lipogenesis
Fatty acid biosynthesis (path 1)
PYK
GK
Glycolysis
PGC-1α
FOXO1
G6PC
FBP
PEPCK
Glycolysis / Glucone ogenesis
DNA
INS
INSR
IRS
PI3K
PIP3
PDK1/2
Akt
GSK-3β
GYS
Glycogen
Glycogenesis
PP1
PHK
PYG
Starch and sucrose metabolism
TSC2
TSC1
Rheb
PDE3
cAMP
PKA
HSL
Antilipolysis
Lipid homeostasis
LAR
Phosphatidyl inositol signaling system
+ps : serine phosphorylation
+py : tyrosine phosphorylation
-py : tyrosine dephosphorylation
BAD
Antiapoptosis
Apoptosis
mTOR
Raptor
p70S6K
S6
4EBP1
eIF4E
Protein synthesis
MAPK signaling pathway
SHC
GRB2
SOS
Ras
Raf
MEK1/2
ERK1/2
MNK
Elk1
DNA
Proliferation, differentiation
04910 6/16/23
(c) Kanehisa Laboratories

Supplementary Figure 16 | Differential pathway enrichment analysis showing that phosphorylation of AMPK and insulin signaling pathways were activated following *in vivo* exposure to r-RORDEP1. (a) Proteins within the AMPK pathway modulated by r-RORDEP1 treatment. **(b)** Proteins within the insulin signaling pathway affected by r-RORDEP1 treatment. Up- and down-regulated phosphoryl proteins are marked with red and blue boxes, respectively.

Reference:

1. Grishin, N. V. Estimation of the number of amino acid substitutions per site when the substitution rate varies among sites. *J Mol Evol* **41**, 675–679 (1995).
2. Kraal, L., Abubucker, S., Kota, K., Fischbach, M. A. & Mitreva, M. The Prevalence of Species and Strains in the Human Microbiome: A Resource for Experimental Efforts. *PLOS ONE* **9**, e97279 (2014).
3. Yachida, S. *et al.* Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat Med* **25**, 968–976 (2019).
4. Wirbel, J. *et al.* Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat Med* **25**, 679–689 (2019).