Article

Discovery through Machine Learning and Preclinical Validation of Novel Anti-Diabetic Peptides

Rory Casey, Alessandro Adelfio, Martin Connolly, Audrey Wall * , Ian Holyer and Nora Khaldi

Nuritas Ltd., Joshua Dawson House, D02 RY95 Dublin, Ireland; casey.rory@nuritas.com (R.C.);
adelfio.alessandro@nuritas.com (A.A.); connolly.martin@nuritas.com (M.C.); info@nuritas.com (I.H.);
n.khaldi@Nuritas.com (N.K.)
* Correspondence: wall.audrey@nuritas.com; Tel.: +353-1-430-1290

Abstract: While there have been significant advances in drug discovery for diabetes mellitus over the past couple of decades, there is an opportunity and need for improved therapies. While type 2 diabetic patients better manage their illness, many of the therapeutics in this area are peptide hormones with lengthy sequences and a molecular structure that makes them challenging and expensive to produce. Using machine learning, we present novel anti-diabetic peptides which are less than 16 amino acids in length, distinct from human signalling peptides. We validate the capacity of these peptides to stimulate glucose uptake and Glucose transporter type 4 (GLUT4) translocation in vitro. In obese insulin-resistant mice, predicted peptides significantly lower plasma glucose, reduce glycated haemoglobin and even improve hepatic steatosis when compared to treatments currently in use in a clinical setting. These unoptimised, linear peptides represent promising candidates for blood glucose regulation which require further evaluation. Further, this indicates that perhaps we have overlooked the class of natural short linear peptides, which usually come with an excellent safety profile, as therapeutic modalities.

Keywords: drug discovery; peptide; type 2 diabetes; machine learning

1. Introduction

Type 2 diabetes mellitus (T2DM) is a chronic condition which accounts for over 90% of all diabetes mellitus (DM) incidences. In T2DM, cells fail to respond to the hormone insulin, or a relative lack of insulin is produced by the beta cells of the pancreas, which normally allows glucose to enter cells from the blood, reducing blood glucose levels. This condition has long-term implications, affecting several organs in the body, such as nephropathy of the kidney and retina, and hepatic steatosis, all of which contribute to poor quality of life and a high burden on healthcare systems [1]. Major risk factors for T2DM include obesity, lack of exercise and sedentary lifestyle, all of which are increasingly common in the West. Currently, the global incidence of DM continues to grow at an inexorable rate, currently affecting over 450 million people, and is expected to afflict almost 700 million by 2045 [2]. While DM drug discovery has seen some important advances over the last two decades [3,4], in light of such widespread disease prevalence, there is an evident and urgent need for novel, effective anti-diabetic treatments that have improved safety profiles and are well tolerated for chronic use in the DM population [5,6].

Pharmacological interventions for T2DM include metformin, a small-molecule drug known to work via several mechanisms, including AMP-activated protein kinase (AMPK) [7] and mitochondrial activity [8,9]; it is a well-established first line treatment for T2DM, both as a monotherapy and in combination with other medications, and has been routinely shown to have glucose-lowering effects [10]. Thiazolidinediones are cyclic compounds that act as ligands to peroxisome proliferator-activated receptors (PPARs), reducing circulating fatty acids and increasing the expression of glucose transporter Glucose transporter type 4 (GLUT4), allowing cells to take up more glucose from the blood for energy, reducing blood

glucose levels [11,12]. More recently, sodium glucose transporter 2 (SGLT2) inhibitors have been developed, which work to prevent the action of the transporters in the proximal tubule of the kidney from reabsorbing glucose to the body, thereby allowing excretion and reducing blood glucose [13]. Indeed, SGLT2 inhibitors are often used in combinational therapy with other medications to help manage the complex pathophysiology of T2DM [14].

An important and efficacious group of T2DM therapeutics includes a group of peptides such as insulin analogues, and, more recently, incretin mimetics, acting as either agonists (e.g., Glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP) analogues) or antagonists (e.g., GLP-1 receptor antagonists) of endogenous human hormones, with modifications [15]. These therapeutics form a significant number of the modern anti-diabetic drugs and also those in the development pipeline [16]. This group has been shown to suppress glucagon and hepatic glucose production, slow gastric emptying and reduce appetite [17], with two approved members of the class, Exenatide and Liraglutide, having long-term weight loss effects on patients over a 1–2-year period [18]. Both Exenatide and Liraglutide are relatively shorter compared to 51-amino-acid-long insulin, at 39 amino acids and 32 amino acids in length, respectively. Both medications are administered as subcutaneous injectables and havepresented with some adverse effects, including vomiting and nausea [19,20].

An alternative class of peptides which may offer good safety profiles and toleration for chronic use, but have been largely underexplored, are short linear peptides with few modifications. Linear endogenous and synthetic peptides have been shown to be capable of modulating intracellular signalling, without modifications [21]. As such, these advantages position this class as an attractive addition to the diabetes armamentarium. Indeed, peptides can be highly selective, having multiple points of contact with their target, which may result in decreased side effects and toxicity [22]. Furthermore, as they comprise amino acids, peptides are easily metabolised over time, thereby avoiding the tolerance issues that can be associated with chronic administration of many drugs [22,23]. A possible advantage of short linear peptides includes lower manufacturing costs and offers a flexible base for modifications [24,25]. However, presently, a major drawback of these peptides is that they are readily broken down during gastrointestinal digestion (GID); therefore, issues with low bioavailability via oral administration remain problematic [26]. This has resulted in a turn towards developing optimised peptides to enhance therapeutic properties, such as cyclisation [27], although there are instances where anti-cancer linear peptides outperform their cyclic counterparts [25], indicating that this class of peptides should not be so readily dismissed.

Biology is an extremely data-rich discipline owing to various "omics" technologies producing increasingly larger volumes of heterogenous data [28]. In recent years, integration of such data has facilitated a greater understanding of the molecular basis of disease [29], but with continued escalation in terms of scale and complexity, human-directed interpretation is rendered increasingly impractical [27,30]. When considering peptides, deciphering scale and complexity becomes a major hurdle; for example, proteins can be broken down into peptides at a rate of 36 million per minute [21]. However, Artificial Intelligence (AI) and deep learning techniques are perfectly primed to extract previously indecipherable knowledge from disparate biological data streams; as such, machine learning is increasingly seen as a discovery tool in life science, with bioactive peptides being successfully predicted in the areas of inflammation and skin aging [31–34]. Here, similar machine learning methods were employed to identify a short linear novel peptide therapeutics for use in T2DM. The goal of this project was to identify (a) peptide(s) which could modulate an effect on blood glucose levels, GLUT4 expression and/or glycated haemoglobin (HbA1c) levels, while being both non-toxic and showing no off-target effects. The peptide candidates were validated in both in vitro and in vivo assays to ensure these properties.

2. Materials and Methods

2.1. Cell Line

Human skeletal muscle cells (HSkMCs; Cell Applications Inc., San Diego, CA, USA) were cultured at 37 °C, 5% CO_2 in HSkMC growth medium (Cell Applications Inc., San Diego, CA, USA). HSkMCs cultured for, at most, 10 passages were used for all experiments described.

2.2. Animals

All animal procedures were carried out in accordance with Institutional Animal Care and Use (IACUC) guidelines in an Association for Assessment and Accreditation of Laboratory Animal Care International-accredited facility. Ethical approval was granted by the International Association of Religious Freedom (IARF #:MLR-101, IARF #:MLR-115) Studies were performed with 12-week-old male KK.Cg-A^y/J (KK-A^y) mice obtained from the Jackson Laboratory, which were randomly assigned to treatment groups according to baseline fasting blood glucose (IARF #: MLR-101, 1 May 2018). Mice were housed with no more than 4 per cage on a 12-h light/dark cycle with ad libitum access to standard rodent chow and water. Mice were subcutaneously (sc) administered either indicated doses of peptide or saline (untreated vehicle control) once daily for 2 weeks. Bodyweight was measured at baseline and once per week thereafter. Fasting (overnight) blood glucose was measured at baseline, day 5, day 7 and day 13 of dosing (cohort 1). In a second, independent in vivo study (IARF #: MLR-151; 13 January 2020), mice (n = 11/group) were sc administered either indicated dose of peptide, Liraglutide or saline for 6 weeks. Glycated haemoglobin was measured at 6 weeks, and liver sectioning via microtome, staining with haematoxylin and eosin (H&E) and scoring were performed on snap-frozen tissues.

2.3. Prediction Workflow

A similar predictive model to that used by Kennedy et al., 2020 [33] was utilised here; briefly, the model was developed using an ensemble of neural networks. To build the training set for the model, we used structured data from public databases (bioactivity annotations, biological pathways and structural annotations) and unstructured data extracted from peer-reviewed scientific papers and patents (Figure 1). Initial descriptors used to query these data sources were "diabetes", "blood glucose regulation" and "GLUT4". A combination of graph-based techniques and manual curation was used to process the structured data, while, concurrently, Natural Language Processing (NLP) techniques, such as word and sentence embedding and named entity recognition, were applied to the unstructured data. The high-level information extracted was assembled and formatted into a bespoke peptide representation format.

The resulting dataset of peptides with a known effect on blood glucose regulation was used to train our predictive architecture for bioactivity in fold cross-validation. The fully trained model was used to predict novel peptides' glucose uptake efficacy from a large input set of peptide sequences.

Additional testing and refinement of the predictive architecture was achieved by incorporating a predict–test–refine loop. The predict–test–refine loop is an example of active learning. Uncertainty sampling was performed where peptides that the model was least certain as to what their activity should be were selected for experimental testing. This strategy was mixed with the selection of the best predicted peptides, which were tested concurrently. Uncertainty sampling was prevalent in the first iterations and progressively reduced to be completely replaced by the identification of the most promising peptides in the latest iteration. The results of in vitro testing were additionally integrated into the model to bias it towards the prediction of peptides with specific GLUT4 translocation activity. In this framework, a set of peptides, which the model found most difficult to classify (i.e., those with an efficacy prediction close to 50%), were selected for experimental testing in vitro, with resultant data being fed back to the predictive model. This predict–test–refine loop was completed three times. Across the multiple iterations, in vitro activity

was measured in 74% of cases, where peptides tested in glucose regulation assays would show what was internally classified as "medium to high activity". A similar active learning paradigm was used in Kennedy et al., 2020 [33]. In that case, a lower ratio of in vitro activity was measured across the multiple lab tests, with only 40% of peptides tested in extra cellular matrix development assays displaying "medium to high activity".
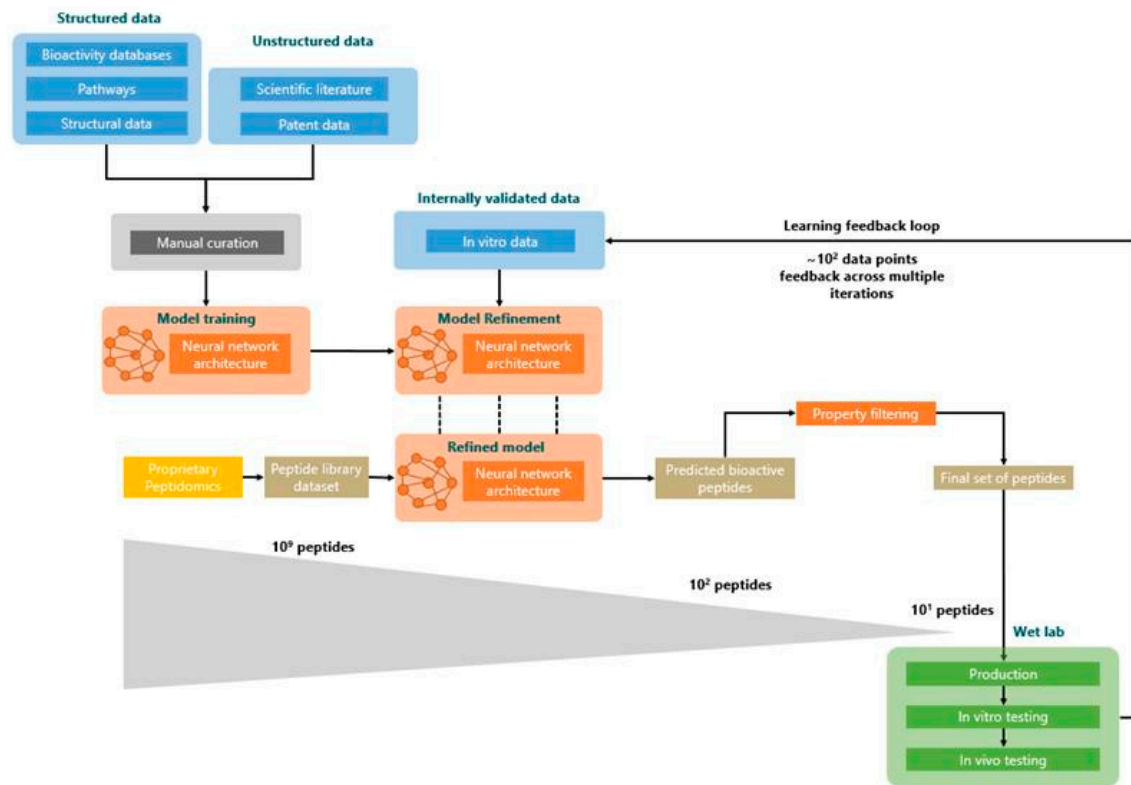

Figure 1. Peptide prediction. Workflow for predictive models adapted from Corrochano et al., 2021 [35].

At the end of the refinement process, a set of 10^9 peptides was fed into the model, which returned an output set of 10^2 peptides classed as "active."

Next, a further collection of internally built tools was used to filter out sequences exhibiting undesirable properties. To narrow down the number of relevant peptides to be tested at each iteration (from hundreds to tens), predicted peptides were ranked using different internal predictive models. Specifically, peptides predicted as cell-penetrant, stable in blood and not toxic were prioritised. Additionally, other filters were applied, removing peptides with an odd number of cysteines or with sequence longer than 30 residues, to facilitate synthesis. Finally, all peptides exhibiting high homology against peptides contained known to have a role in glucose regulation were discarded. This final stage ultimately left a set of 5 distinct, novel peptides suitable for experimental validation.

2.4. Homology Searching and Synthesis of Predicted Peptides

To determine the true novelty of our 5 predicted peptides, we measured their homology to (1) each other and (2) to analogues or antagonists of endogenous human hormones. All searches were performed using BLASTP from the BLAST+ (BLAST+, v2.2.31) suite of programs using the following parameters: word size = 2; matrix = PAM30, E-value = 10,000. Predicted peptides were chemically synthesised by GenScript Corporation (Piscataway, New Jersey, United States). For all peptides screened at initial stages, theoretical molecular weight (MW) was checked, and all were confirmed to have HPLC purity of ≥95.0%. pep_1E99R5 had a theoretical MW of 1270.48 with a HPLC purity of ≥98.0%.

2.5. Glucose Uptake Assay

HSkMCs were plated on collagen-coated 96-well plates (1×10^4/well) and allowed to adhere overnight at 37 °C, 5% CO_2 in 100 μL of HSkMC growth medium. The medium was then changed to HSkMC differentiation medium (Promocell, Heidelberg, Germany) and cells were allowed to differentiate for 7 days, with fresh medium added every 2 days. The day prior to the experimentation, cells were starved overnight in basal medium. Glucose uptake was measured using a glucose uptake assay kit (Abcam, Cambridge, MA, USA) as per the manufacturer's instructions. Briefly, cells were rinsed three times in Dulbecco's phosphate-buffered saline (DPBS; Lonza, Basel, Switzerland) and then starved of glucose by incubating with 100 μL of Krebs-Ringer-Phosphate-Hepes (KRPH) buffer for 40 min at 37 °C. KRPH buffer was made with 20 mM HEPES, 5 mM monopotassium phosphate, 1 mM magnesium sulphate, 1mM calcium chloride, 136 mM sodium chloride, 4.7 mM potassium chloride, adjusted to pH 7.4, all compounds were purchased from Sigma-Aldrich (St. Louis, MO, USA). Cells were treated with HSkMC basal medium containing 0.5 μg/mL (~0.4 μM) of peptide or 1 μM of human insulin solution (Sigma-Aldrich, St. Louis, MO, USA) for 20 min, followed by incubation with 10 μL of 2-deoxyglucose (2-DG; Glucose uptake assay kit; Abcam, Cambridge, UK) for 20 min at 37 °C. Subsequently, cells were washed 3 times with PBS and lysed with 80 μL of extraction buffer, after which cell lysates were freeze–thawed once before heating at 85 °C for 40 min. Following cooling on ice for 5 min, lysates were neutralised by adding 10 μL of neutralisation buffer and then diluted with assay buffer to a total volume of 50 μL (5 μL lysate + 45 μL assay buffer). After two amplification reactions, absorbance of the samples was measured at 412 nm with a microplate spectrophotometer.

2.6. GLUT4 Translocation Assay

HSkMCs aliquots (2×10^5/well) were seeded in muscle growth medium in collagen-coated 6-well plates and differentiated in differentiation medium for 7 days. The day prior to the experimentation, cells were starved overnight in basal medium. The cells were treated with peptide (0.5 μg/mL (~0.4 μM)) or 1 μM human insulin solution for 20 min. Membrane proteins were then solubilised and isolated from cytosolic proteins using a Mem-PER™ Plus Membrane Protein Extraction Kit (ThermoFisher, Waltham, Mass, USA) as per the manufacturer's instructions. A 1-μM treatment with insulin to investigate glucose uptake and the associated downstream signalling is widely used in the literature, in both rat and human cell line models [36,37]. GLUT4 concentration was subsequently measured via a commercially available Human GLUT4 Sandwich ELISA kit (Abbexa, Cambridge, UK). Briefly, 100 μL of standard, blank or sample was loaded into individual wells of a 96-well plate and incubated at 37 °C for 2 h. Liquid was then aspirated and 100 μL of detection reagent A was added for 1 h at 37 °C. Subsequently, liquid was aspirated from each well, which was then washed 3 times using 350 μL of wash buffer before detection reagent B was added for 30 min at 37 °C. Wells were washed 5 times, as before, following which 90 μL of 3,3',5,5'-Tetramethylbenzidine substrate was added to each well and incubated at 37 °C for 10 min while protecting from light. Then, 50 μL of Stop solution was then added, after which the optical density of the sample was determined using a microplate reader (SpectraMax M3, Molecular Devices, Sunnyvale, CA 94089, USA) set to 450 nm.

2.7. Microarrays

HSkMCs aliquots (2×10^5/well) were seeded in muscle growth medium in collagen-coated 6-well plates and differentiated in differentiation basal medium for 7 days. The day prior to the experimentation, cells were starved overnight in basal medium. The cells were treated, in triplicate, with peptide (0.5 μg/mL (~0.4 μM)) or 1 μM or human insulin solution for 20 min, at which point the treatment medium was removed and the cells were scraped in 1 mL of PBS. The cell suspension was pelleted in a microcentrifuge at 1500 rpm for 5 min, and the supernatant was removed. The cells were immediately flash-frozen in liquid nitrogen and transferred to a −80 °C freezer for storage prior to RNA extraction. RNA

was extracted in accordance with standard operating procedures for RNA extraction from tissue/cell pellets using a RNeasy mini kit (Qiagen, Manchester, UK). RNA quality and integrity were determined via bioanalyser. For each microarray experiment, 100 ng of total RNA was used. To study the whole genome expression with a comprehensive coverage of genes and transcripts, 26,083 Entrez Genes and 30,606 lncRNA, SurePrint G3 Human Gene Expression v3, 8 × 60 K Microarrays (Agilent, Santa Clara, CA, USA) were used. Microarray gene expression experiment was performed according to the manufacturer's protocol (One-Color Microarray-Based Gene Expression Analysis—Low Input Quick Amp Labeling v6.9). After the experiment, the arrays were scanned by SureScan Microarray Scanner (Agilent, Santa Clara, CA, USA) and data were extracted using Feature Extraction Software (Agilent, Santa Clara, CA, USA). The samples were prepared for array hybridisation according to the manufacturer's protocol. Briefly, labelled cRNA was hybridised to the microarray for 16 h before the array slides were washed and scanned using an Agilent G2565CA Microarray Scanner System.

2.8. Transcriptomics and Pathway Enrichment

Microarray data analysis was conducted using the R Bioconductor package limma [38]. Intensities were background corrected using the normexp method with an offset of 50 [39]. Quantile normalisation was then applied. Due to the use of several arrays, subsequent batch effect was identified and removed using ComBat (sva R package) [40]. Each treated group was compared to the negative untreated control using linear modelling, from which an empirical Bayesian analysis was then performed using the function ebayes from limma.

To initiate the pathway enrichment analysis, significant gene lists were filtered using raw p-value < 0.01 and fold-change > 1.3. Enrichment was performed on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways using a hypergeometric test. Pathways were considered enriched at raw p-value < 0.01. Enrichment was assessed for each treatment insulin and pep_1E99R5, respectively, considering up- and downregulated gene lists. All analysis was performed internally and by third-party Fios Genomics Ltd. (Edinburgh, UK).

2.9. Statistical Analyses

All data are presented as mean ± SEM. Replicate numbers for each experiment are indicated in figure legends. Results of in vitro experiments were assessed by one-way ANOVA followed by Dunnett's post-hoc test. Comparisons between different peptide treated and untreated KK-A^y groups were assessed by one-way ANOVA followed by Dunnett's post-hoc test. Statistical significance was defined as $p < 0.05$. Graphs were generated using the "ggplot2" R package [41].

3. Results

3.1. Novel Peptide Prediction and Validation

Using a machine learning approach similar to Kennedy et al., 2020 [33], one hundred peptide candidates were predicted as potentially possessing anti-diabetic functionality via blood glucose regulatory activity. This set of peptides was further refined using a collection of tools to filter out the sequences with undesirable physiochemical properties. Taking into account the significant costs involved in peptide drug manufacturing [42], all predicted peptides were to be less than 20 amino acids in length and linear with no major structures. Ultimately, this resulted in a final set of five peptides that interestingly exhibited no homology to each other or to any known patented or published bioactive peptides. These peptides are referred to hereafter as pep_1E99R5, pep_37MB3O, pep_ANUT7B, pep_RTE62G and pep_QT5XGQ. To validate the bioactivity of these predicted peptides, an in vitro glucose uptake method was employed.

Insulin-stimulated uptake of blood glucose by skeletal muscle plays a fundamental role in the maintenance of glucose homeostasis, accounting for 75% of glucose utilisation in the body [43]. Accordingly, the predicted bioactivity of our predicted peptides was first validated in a cell-based glucose uptake assay, where insulin was used as a positive control.

Three predicted peptides, pep_1E99R5, pep_37MB3O and pep_ANUT7B, demonstrated the ability to significantly increase glucose uptake in human skeletal muscle cells (HSkMCs), with pep_37MB3O and pep_ANUT7B displaying a stronger effect than that of insulin (Figure 2A). No significant glucose uptake activity was reported for pep_RTE62G and pep_QT5XGQ; thus, these peptides were not progressed further in our in vitro and in vivo validation studies. Cumulatively, these results indicate an in vitro validation success rate of 60% for the predictive model and suggest that these sequences should be further examined in relevant models. Of note, a more comprehensive absorption, distribution, metabolism, and excretion (ADMET) prediction could be incorporated for further development of the presented peptides.
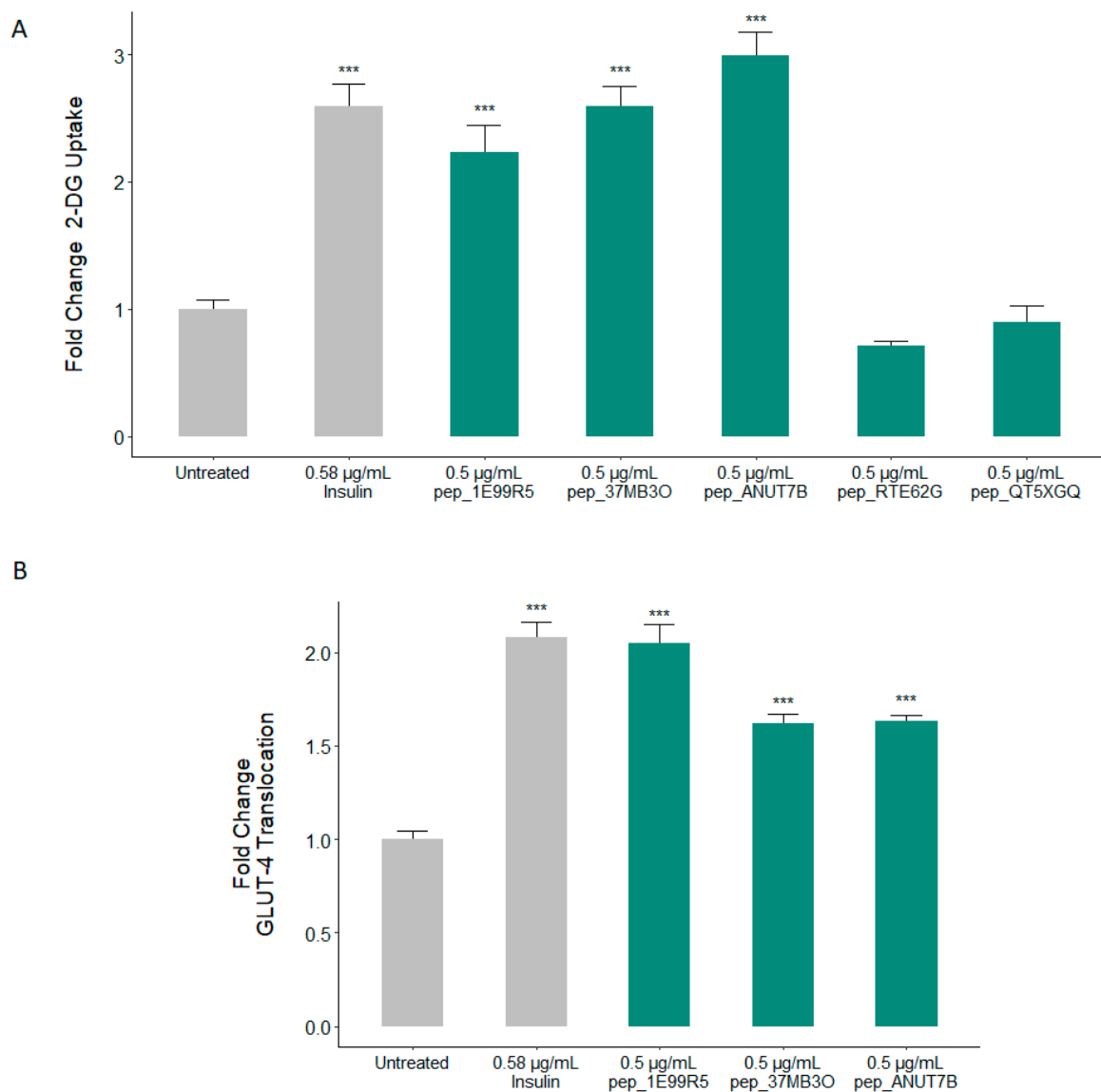

Figure 2. In vitro validation of predicted peptides as potent stimulators of glucose uptake and Glucose transporter type 4 (GLUT4) translocation. Skeletal muscle cells were stimulated with either insulin or predicted natural peptides at the indicated concentrations for 15 min. (**A**) The effect on glucose uptake is measured, while in (**B**), the extent of GLUT4 translocation is assessed (one-way ANOVA analysis with Dunnett's test; *** $p < 0.001$; data presented are the mean $\pm$ SEM of at least 3 independent replicates).

The key mediator of glucose uptake into skeletal muscle cells is the protein GLUT4 [44]. Skeletal muscle accounts for the majority of glucose uptake in the body [45]. While glucose uptake efficacy was used as an initial experimental validation of our predicted peptides,

the ultimate objective of the predictor was to identify peptides capable of stimulating GLUT4 translocation to the plasma membrane, which is known to be decreased in type 2 diabetics [46]. Indeed, interventions to promote the expression and translocation of GLUT4 in appropriate cells such as skeletal muscle fibres are of clear benefit to those with T2DM [47]. The HSkMC model used for this study allows for measurement of GLUT4 translocation, which, in turn, allows for a specific molecular mechanism to be identified as the cause of these peptides' ability to modulate glucose levels. Accordingly, we evaluated the ability of these peptides to initiate GLUT4 translocation in HSkMCs. At a test dose of 0.5 µg/mL, we found that all three predicted peptides stimulated a highly significant increase in GLUT4 translocation (Figure 2B). Of the three positively predicted peptides, pep_1E99R5 is seen to demonstrate the most potent effect, eliciting an approximately equivalent response to insulin (Figure 2B).

3.2. Validation of Peptides in a Diabetic Mouse Model

The therapeutic potential of our in-vitro-validated anti-diabetic peptides was evaluated in the KK-A^y model of obese insulin-resistant DM. In KK-A^y mice, fasting blood glucose and HbA1c levels are elevated, and when fed a normal diet, obesity and DM are observed by 12 weeks. Peptides pep_37MB3O, pep_1E99R5 and pep_ANUT7B at 127 mg/kg (100 µM) or vehicle were dosed subcutaneously (sc) once daily for 14 days. All animals tolerated treatment well and body weight did not change per treatment throughout the study (Figure S1). At day 5 post-baseline, all three predicted peptides significantly reduced fasting blood glucose compared to the vehicle control group (Figure 3), with pep_1E99R5 demonstrating the most significant reduction. The effect of pep_ANUT7B was tempered by day 7 ($p = 0.07$) and day 13, with pep_1E99R5 and pep_37MB3O both exhibiting a significant reduction in fasting blood glucose at these time points. However, the effect of pep_1E99R5 was slightly reduced at day 13 (Figure 3), a trend that is in line with what is observed for known anti-diabetic therapies in several animal studies [48,49]. While pep_1E99R5 and pep_37MB3O's effects were maintained throughout the study period, pep_1E99R5 treatment resulted in the most potent effects on fasting blood glucose levels in vivo; therefore, this peptide was chosen for a longer in vivo study to measure HbA1c.
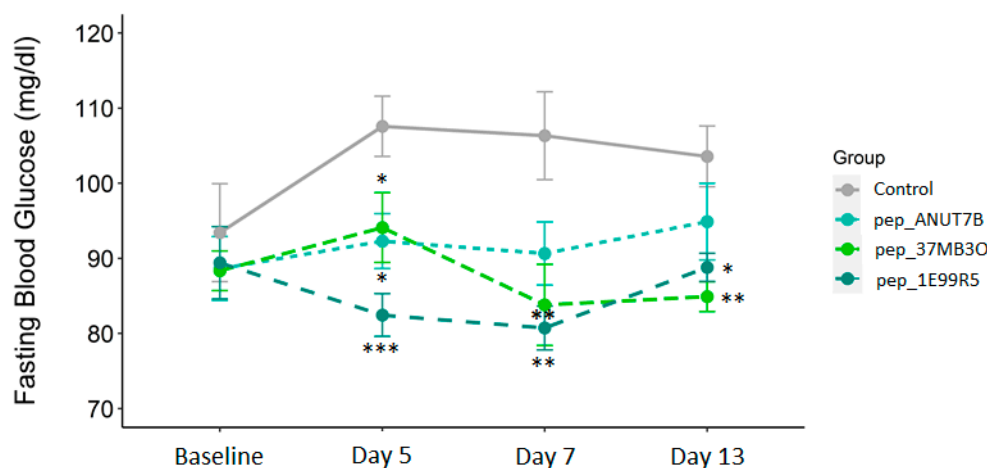

Figure 3. Predicted peptides in a mouse model of type 2 diabetes. Fasting glucose levels (overnight) of all mice were measured as indicated. Data are mean ± SEM ($n = 8$ per group; aged 12 weeks at baseline) and analysed by Dunnett's test to compare the differences between the peptide treatment groups and the vehicle control group (* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$).

As a decrease of >1% HbA1c is considered to be of significant clinical benefit [50], the effects of pep_ 1E99R5 on HbA1c % were measured in a trial over 6 weeks. In KK-A^y mice, pep_1E99R5 (12.7 mg/kg (10 µM) or 63.5 mg/kg (50 µM)) was administered daily via sc injection and compared to control groups of vehicle or Liraglutide (250 µg/kg), previously

shown to have a positive effect on HbA1c [51]. HbA1c levels are a long-term indicator of blood glucose regulation [52]; previous studies have shown significant treatment effects on HbA1c at 8 weeks in similar KK-A^y models [53]. However, Liraglutide treatment did not significantly reduce HbA1c percentage compared to vehicle control following 6 weeks of treatment (Figure 4); it is possible that a significant drop in HbA1c levels would have been recorded if the study duration were extended. In contrast, pep_1E99R5 (63.5 mg/kg)-treated mice showed a significant reduction of approximately 1.3% in HbA1c compared to vehicle control, suggesting a sustained effect of pep_1E99R5 in the animals. No effect was noted in the 10-µM-treated mice, which suggests a dose-dependent effect of the peptide.
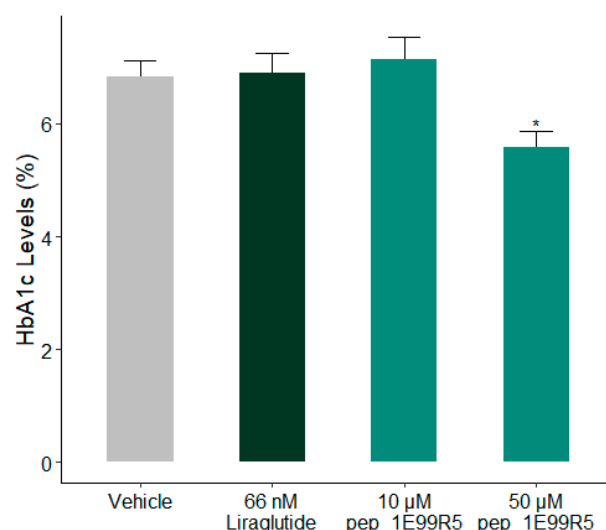
Figure 4. Effect of pep_1E99R5 on HbA1c levels. HbA1c levels (%) of all mice were measured as indicated. Data are mean ± SEM (n = 11 per group; aged 12 weeks at baseline) and analysed by Dunnett's test to compare the differences between peptide treatment groups, Liraglutide group and vehicle control (* $p < 0.05$).

A common complication associated with T2DM is hepatic steatosis (HS), a condition in which fatty deposits accumulate in the liver, affecting up to 75% of T2DM patients, often leading to non-alcohol fatty liver disease (NAFLD) [54]. Moreover, HS among the prediabetic population is considered to be a predictor of conversion to DM [55]. In a study by Fiorentino et al., a subset of HbA1c-defined prediabetic individuals with 1-h postload glucose ≥155 mg/dL were at higher risk of developing HS [56]. Furthermore, the prevalence of prediabetes and DM was found to be six-fold higher in NAFLD patients compared to healthy controls. [57] Therefore, a decrease in HS, alongside the HbA1c decrease, would be of significant clinical interest.

Consequently, levels of hepatic steatosis (HS) in the mice were measured via sectioning and H&E staining. Histological NAFLD scoring was performed by an independent third-party reviewer. Features were scored according to a murine liver scoring system devised by [58] (Figure 5A). Vehicle control and Liraglutide-treated mice exhibited signs of NAFLD; however, these were reduced significantly in 50-µM pep_1E99R5 and trended to a decrease in 10-µM pep_1E99R5 treatment, suggesting a dose-dependent effect (Figure 5B).
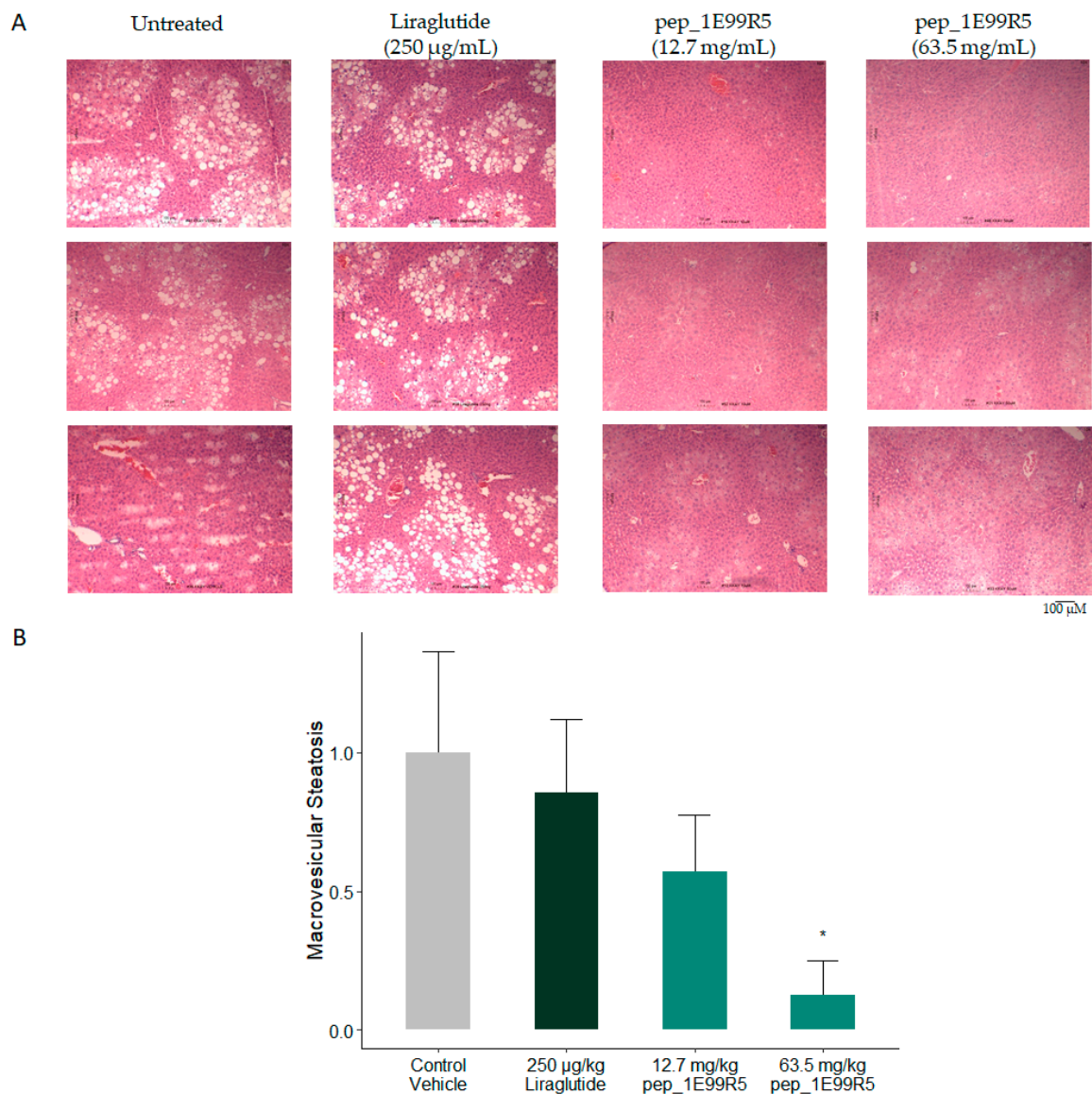
Figure 5. Effect of pep_1E99R5 on hepatic steatosis. (**A**) Livers of all mice were excised and sectioned before staining with H&E. Images were taken using a Motic BA310E trinocular compound microscope at 10x magnification. Histological non-alcohol fatty liver disease (NAFLD) scoring was performed by an independent third-party reviewer. (**B**) Features were scored according to a murine liver scoring system devised by Liang et al., 2014. Data are mean $\pm$ SEM (n = 6 per group; aged 12 weeks at baseline) and analysed by Dunnett's test to compare the differences between the two peptide treatment groups and vehicle control and Liraglutide groups (* $p < 0.05$).

### 3.3. Characteristics of Predicted Peptides

As short linear peptides offer an intriguing therapeutic option due to decreased side effects and toxicity [22], the predictive model focused on peptides <20 amino acids (AA) in length. Of note, pep_1E99R5 consists of 11 AA, WKDEAGKPLVK, with no major structures (Table 1; Figure 6A). While this specific sequence of amino acids was predicted due to desirable features identified through the predictive model, further work is being carried out on derivatives of pep_1E99R5 to assess bioactivity retention. Apart from structure and length, an important consideration often overlooked when describing predictions of functional compounds is the level of true novelty offered by the de-novo-discovered compounds. Accordingly, focusing our most promising candidate peptide, pep_1E99R5, presented in Table 1, we assessed the extent to which it was truly novel in light of (1) key

human peptide hormones involved in glucose regulation and (2) the larger set of 1550 known endogenous human peptides (Figure 6B). To achieve this, we calculated three major peptide properties: charge, molecular weight and hydrophobicity. Here, pep_1E99R5's charge (+1) and hydrophobicity (45.5%) are no different to most human endogenous peptides, with an average charge of +2.9 (50% of the peptides between −1 and +7) and an average hydrophobicity of 46.4 (50% of the peptides between 42.9% and 50.4%) (Figure S2). However, pep_1E99R5 (1.3 kDa) is substantially smaller than 99.6% of human endogenous peptides and smaller than both insulin (5.8 kDa) and GLP-1 (3.3 kDa), as well as on-the-market DM-treatment peptides Exenatide (4.1 kDa) and Liraglutide (3.7 kDa), which offers a considerable advantage in terms of production cost. pep_1E99R5 was further assessed against the natural glucose-regulating hormones in terms of homology, where, even at a very lenient e-value threshold, no sequence similarity was reported.

Table 1. pep_1E99R5 characteristics.

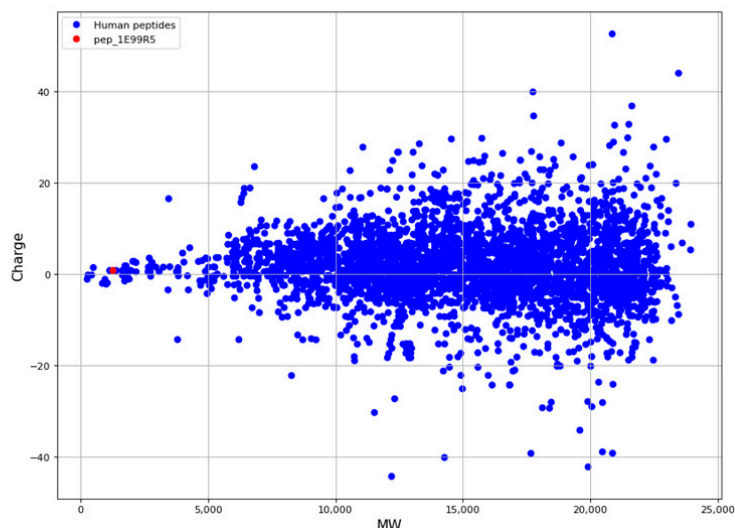| Sequence	Length (Amino Acid)	Molecular Weight (Da)	Charge	Isoelectric Point	Hydrophobicity
WKDEAGKPLVK	11	1270.48	1	8.5	45.5%

A
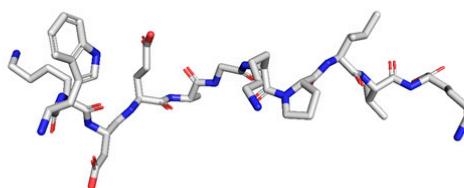
B

Figure 6. Linear representation of pep_1E99R5 and molecular weight and charge dispersion of pep_1E99R5 compared to endogenous human peptides. (**A**) Representation of the linear structure for pep_1E99R5 generated using PyMol, Version 2.3.5, Schrödinger, LLC (stick visualisation). (**B**) Dispersion of molecular weight and charge of human endogenous peptides (blue) and pep_1E99R5 (red). The human peptides were retrieved from UniprotKB (https://www.uniprot.org/statistics/Swiss-Prot (Accession date: 8 March 2021)) and filtered with a threshold of 200 amino acids. The average molecular weight of the human peptides is 14.4 kDa (Q1 = 11.3 kDa, Q3 = 18.1 kDa) while pep_1E99R5 is 1.3 kDa. The average charge of the human peptides is +2.9 (Q1 = −1, Q3 = +7) while pep_1E99R5's charge is +1.

3.4. Molecular Mechanisms Modulated by pep_1E99R5

To elucidate the molecular pathways modulated by pep_1E99R5 in vitro, HSkMCs treated with the peptide underwent a full transcriptomic screen. HSkMCs treated with insulin were screened in parallel. To determine gene expression, RNA was extracted from the cells, fluorescently labelled and run in triplicate on SurePrint G3 Human Gene Expression v3 8 × 60 K Microarrays. Microarray technology simultaneously measures the expression of large numbers of transcripts in treated and untreated samples.

A 20-min stimulation of HSkMCs with 0.5 μg/mL (0.4 μM) pep_1E99R5 changed the expression, up or down (>1.3 fold; $p < 0.01$) of 625 transcripts, compared to untreated cells. A 1-μM insulin treatment of HSkMCs was run in parallel and changed the expression, up or down (>1.3 fold; $p < 0.01$) of 540 transcripts, compared to untreated cells (data not shown). KEGG analysis was used to obtain a biochemical overview of the pathways differentially expressed under the influence of pep_1E99R5 and insulin [59].

Pathways related to glycolysis, oxidative phosphorylation and the citrate cycle were among the most highly ranked of the 21 KEGG pathways enriched by our peptide (all $p < 0.01$) (Figure 7). These pathways are typically downregulated in T2DM patients [60]. The pentose phosphate pathway, an alternative pathway to glycolysis, and the citrate cycle for oxidation of glucose also showed a trend toward enrichment when treated with pep_1E99R5, as did the TGF-β pathway, shown to stimulate the glucose uptake through GLUT1 [61]. Stimulation of these key glucose metabolism pathways with pep_1E99R5 may be key to its efficacy in enhancing glucose uptake in skeletal muscle cells. Nineteen KEGG pathways were downregulated by pep_1E99R5 ($p < 0.01$), including the PI3k-Akt pathway, the primary pathway for insulin-stimulated glucose uptake. This is further evidence that our peptide stimulates glucose uptake independent of the insulin pathway. The p53 signalling pathway was also downregulated by pep_1E99R5; typically, a tumour supressing pathway, this mechanism has the added ability to mediate metabolic changes in cells through the regulation of energy metabolism and has been shown to disrupt glucose uptake into cells [62]; hence, downregulation of this pathway will promote glucose uptake into cells.

The transcriptomic profile of cells treated with 1 μM insulin showed minimal KEGG pathway enrichment related to T2DM. This may be due to the relatively high insulin dose that the cells were exposed to, promoting upregulation in genes related to insulin resistance, including SOCS3, which was significantly upregulated by the insulin treatment and has been shown to be key to the physiological regulation of insulin signalling [63]. Conversely the KEGG pathway related to type I DM (T1DM) was among the pathways decreased by insulin treatment ($p < 0.01$). This finding is easier to interpret as the deficiency of bioavailable insulin is the primary mechanism by which T1DM arises. Furthermore, while the oxidative phosphorylation pathway showed a trend towards enrichment in the insulin-treated cells, this enrichment was less than that observed in the pep_1E99R5-treated samples. Comparison of significant gene enrichment in this pathway when the cells are stimulated with insulin ($n = 3$ genes) versus pep_1E99R5 ($n = 9$ genes) suggests that pep_1E99R5 has a greater effect in activating this key glucose metabolism pathway. While these data indicate different transcriptomic profiles for pep_1E99R5 and insulin, further work is underway to elucidate specific targets of pep_1E99R5.

Peptides represent the largest class of signalling molecules in animals, acting as hormones, neurotransmitters and growth factors to perform many critical physiological functions. Given that peptides have evolved to interact with specific biological targets, they offer great promise as selective, potent therapeutics that are less likely to suffer from issues of tolerability and toxicity [22,23]. Although beyond the scope of the current study, it is expected that natural peptides in their current form, due to their linear nature and absence of modifications, can circumvent issues with safety [64]. Indeed, no tolerability concerns were noted in any animals during both in vivo studies in this project with doses as high as 127 mg/kg.
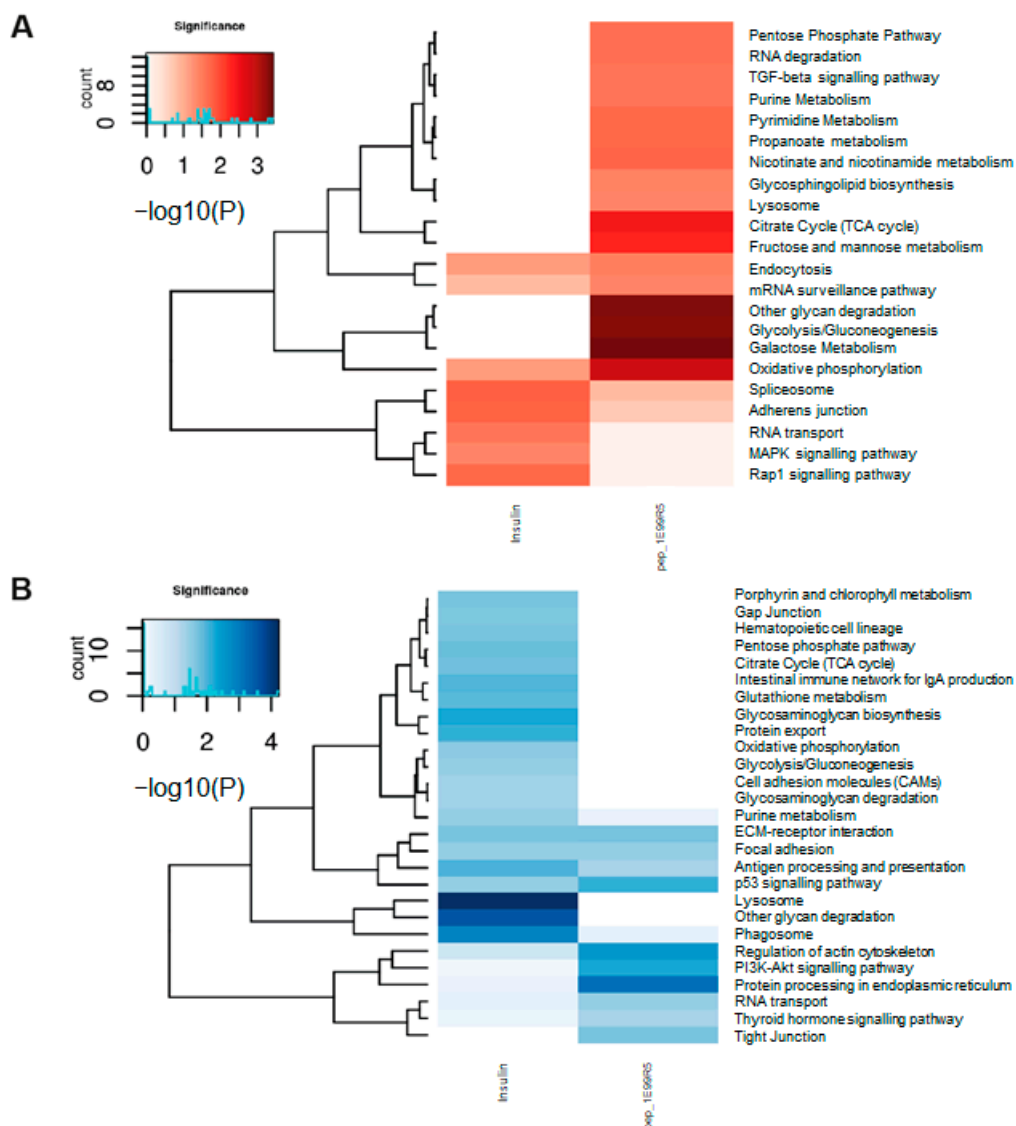
Figure 7. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of pep_1E99R5 and insulin. Representation of the KEGG pathway enrichment of significantly differentially expressed genes for insulin and pep_1E99R5 compared to untreated. Gene lists were filtered using raw *p*-value < 0.05 and fold change ≥ 1.3. The significance of a given KEGG pathway is assessed with raw *p*-value < 0.05. (**A**) Heatmap showing significantly downregulated (red) enrichment in KEGG pathways and (**B**) heatmap showing significantly upregulated (blue) enrichment in KEGG pathways (deeper colour indicates increased significance). Heatmaps present log-transformed raw *p*-values. Enrichment analyses were performed for each treatment (insulin and pep_1E99R5) separately and then brought together for comparison. Colour gradients indicate significance; the darker the colour, the more significant the result is.

To date, linear peptides as therapeutics have been largely underexplored. By their nature, they usually have an excellent safety profile and are easier to manufacture, with reduced loss of yield during synthesis [65]. Given their simple structure, a linear peptide can, in many cases, be optimised for bioavailability and stability more easily [42] than more structured, "difficult" peptides where optimisation alters efficacy [27,66,67]. Our aim here was to find a novel short natural linear peptide that can improve glucose modulation in vivo. When taking molecule size into account in current T2DM therapeutics, pep_1E99R5 is demonstrably smaller than current peptide therapeutics, which creates the potential for improved precision at the target site and reduced manufacture scale-up costs [68]. This coupled with the reduced economic burden observed following improved glycaemic control in T2DM of up to 13% [69], indicates further potential for such a peptide.

Understanding the mechanism of action (MOA) of pep_1E99R5 might reveal new mechanisms of glucose regulation in vivo; our results indicate that its MOA is different to that of insulin and further work is underway to elucidate this. Our initial study is a first step in investigating the world of natural linear peptides which, combined with good stability and bioavailability profiles, could become a repository for future therapeutic development.

The exploration of a diverse class of linear peptides and their association with glucose metabolism could not have been possible without the use of AI and machine learning. The predicted peptide, pep_1E99R5, is capable of modulating GLUT4 translocation, thereby affecting glucose uptake in vitro. Preclinical studies suggest that this peptide is biologically functional, leading to potential clinically relevant changes in both blood glucose and glycated haemoglobin, as well as a concomitant reduction in hepatic steatosis. Furthermore, analysis of the peptide itself, along with KEGG pathway analysis compared to insulin, suggests a unique, novel function of pep_1E99R5 in modulating blood glucose metabolism. An interesting application would be to integrate these machine learning approaches to explore the bioavailability and stability of linear peptides, which could give rise to candidates with not only good safety and efficacy profiles but also desirable pharmacokinetic properties for future therapeutic development in metabolic disorders such as T2D or others.

4. Conclusions

In undertaking this study, we aimed to explore short linear peptides with glucose-regulating activity and present experimental evidence that machine learning methods can reveal truly novel molecules capable of demonstrating meaningful and clinically relevant biological effects—in this case, in the context of T2DM. Of note, efficacious short linear peptides with good tolerability in vivo also present an opportunity for the pharmaceutical industry, with reduced manufacturing costs. Although further work is required to elucidate bioavailability, mechanism of action and clinical efficacy, we show initial evidence that unoptimised predicted peptides can display enhanced bioactivity in vitro than insulin and outperform Liraglutide in a hyperinsulinemic in vivo model. Ultimately, we highlight the capabilities of AI in discovery and present pep_1E99R5 as a short, linear bioactive peptide capable of affecting blood glucose metabolism in vitro and in vivo via robust modulation of a unique network of several key signalling pathways.

Molloy., Tim Considine, Therese Holton, Hansel Gomez and Brian Keogh in the planning and structuring of the data and manuscript. We also appreciate the guidance of Vivian Cong (Melior Discovery Inc., Exton, PA, USA) for animal study design and execution.

References

1. Trikkalinou, A.; Papazafiropoulou, A.K.; Melidonis, A. Type 2 diabetes and quality of life. *World J. Diabetes* **2017**, *8*, 120–129. [CrossRef]
2. Cho, N.; Shaw, J.; Karuranga, S.; Huang, Y.; Fernandes, J.D.R.; Ohlrogge, A.; Malanda, B. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes Res. Clin. Pract.* **2018**, *138*, 271–281. [CrossRef]
3. Tahrani, A.A.; Bailey, C.J.; Del Prato, S.; Barnett, A.H. Management of type 2 diabetes: New and future developments in treatment. *Lancet* **2011**, *378*, 182–197. [CrossRef]
4. Nicholson, G.; Hall, G.M. Diabetes mellitus: New drugs for a new epidemic. *Br. J. Anaesth.* **2011**, *107*, 65–73. [CrossRef] [PubMed]
5. Bhatt, H.B. Thoughts on the progression of type 2 diabetes drug discovery. *Expert Opin. Drug Discov.* **2014**, *10*, 107–110. [CrossRef]
6. Turner, N.; Zeng, X.-Y.; Osborne, B.; Rogers, S.; Ye, J.-M. Repurposing Drugs to Target the Diabetes Epidemic. *Trends Pharmacol. Sci.* **2016**, *37*, 379–389. [CrossRef]
7. Zhou, G.; Myers, R.; Li, Y.; Chen, Y.; Shen, X.; Fenyk-Melody, J.; Wu, M.; Ventre, J.; Doebber, T.; Fujii, N.; et al. Role of AMP-activated protein kinase in mechanism of metformin action. *J. Clin. Investig.* **2001**, *108*, 1167–1174. [CrossRef] [PubMed]
8. Fontaine, E. Metformin-Induced Mitochondrial Complex I Inhibition: Facts, Uncertainties, and Consequences. *Front. Endocrinol.* **2018**, *9*, 753. [CrossRef]
9. Madiraju, A.K.; Erion, D.M.; Rahimi, Y.; Zhang, X.-M.; Braddock, D.T.; Albright, R.A.; Prigaro, B.J.; Wood, J.L.; Bhanot, S.; Macdonald, M.J.; et al. Metformin suppresses gluconeogenesis by inhibiting mitochondrial glycerophosphate dehydrogenase. *Nat. Cell Biol.* **2014**, *510*, 542–546. [CrossRef] [PubMed]
10. Nasri, H.; Rafieian-Kopaei, M. Metformin: Current knowledge. *J. Res. Med. Sci.* **2014**, *19*, 658–664. [PubMed]
11. Hauner, H. The mode of action of thiazolidinediones. *Diabetes/Metab. Res. Rev.* **2002**, *18*, S10–S15. [CrossRef] [PubMed]
12. Smith, U. Pioglitazone: Mechanism of action. *Int. J. Clin. Pract. Suppl.* **2001**, *121*, 13–18.
13. Wright, E.M.; Loo, D.D.F.; Hirayama, B.A. Biology of Human Sodium Glucose Transporters. *Physiol. Rev.* **2011**, *91*, 733–794. [CrossRef] [PubMed]
14. Clar, C.; Gill, J.A.; Court, R.; Waugh, N. Systematic review of SGLT2 receptor inhibitors in dual or triple therapy in type 2 diabetes. *BMJ Open* **2012**, *2*, e001007. [CrossRef] [PubMed]
15. Gupta, V. Glucagon-like peptide-1 analogues: An overview. *Indian J. Endocrinol. Metab.* **2013**, *17*, 413–421. [CrossRef] [PubMed]
16. Kahn, S.E.; Cooper, M.E.; Del Prato, S. Pathophysiology and treatment of type 2 diabetes: Perspectives on the past, present, and future. *Lancet* **2014**, *383*, 1068–1083. [CrossRef]
17. Tasyurek, H.M.; Altunbas, H.A.; Balci, M.K.; Sanlioglu, S. Incretins: Their physiology and application in the treatment of diabetes mellitus. *Diabetes/Metab. Res. Rev.* **2014**, *30*, 354–371. [CrossRef] [PubMed]
18. Astrup, A.; Carraro, R.; Finer, N.N.; Harper, A.; Kunesova, M.; Lean, M.E.J.; Niskanen, L.; Rasmussen, M.F.; Rissanen, A.; Rössner, S.; et al. Safety, tolerability and sustained weight loss over 2 years with the once-daily human GLP-1 analog, liraglutide. *Int. J. Obes.* **2011**, *36*, 843–854. [CrossRef] [PubMed]
19. Astrup, A.; Rössner, S.; Van Gaal, L.; Rissanen, A.; Niskanen, L.; Al Hakim, M.; Madsen, J.; Rasmussen, M.F.; Lean, M.E. Effects of liraglutide in the treatment of obesity: A randomised, double-blind, placebo-controlled study. *Lancet* **2009**, *374*, 1606–1616. [CrossRef]
20. Rosenstock, J.; Klaff, L.J.; Schwartz, S.; Northrup, J.; Holcombe, J.H.; Wilhelm, K.; Trautmann, M. Effects of Exenatide and Lifestyle Modification on Body Weight and Glucose Tolerance in Obese Subjects With and Without Pre-Diabetes. *Diabetes Care* **2010**, *33*, 1173–1175. [CrossRef]
21. Zuconelli, C.R.; Brock, R.; Adjobo-Hermans, M.J. Linear Peptides in Intracellular Applications. *Curr. Med. Chem.* **2017**, *24*. [CrossRef] [PubMed]
22. Craik, D.J.; Fairlie, D.P.; Liras, S.; Price, D. The Future of Peptide-based Drugs. *Chem. Biol. Drug Des.* **2013**, *81*, 136–147. [CrossRef] [PubMed]
23. Fosgerau, K.; Hoffmann, T. Peptide therapeutics: Current status and future directions. *Drug Discov. Today* **2015**, *20*, 122–128. [CrossRef] [PubMed]
24. Radermecker, R.P.; Scheen, A.J. Continuous subcutaneous insulin infusion with short-acting insulin analogues or human regular insulin: Efficacy, safety, quality of life, and cost-effectiveness. *Diabetes/Metab. Res. Rev.* **2004**, *20*, 178–188. [CrossRef] [PubMed]
25. Roxin, A.; Zheng, G. Flexible or fixed: A comparative review of linear and cyclic cancer-targeting peptides. *Futur. Med. Chem.* **2012**, *4*, 1601–1618. [CrossRef]
26. Aungst, B.J.; Saitoh, H.; Burcham, D.L.; Huang, S.-M.; Mousa, S.A.; Hussain, M.A. Enhancement of the intestinal absorption of peptides and nonpeptides. *J. Control. Release* **1996**, *41*, 19–31. [CrossRef]
27. Lee, A.C.-L.; Harris, J.L.; Khanna, K.K.; Hong, J.-H. A Comprehensive Review on Current Advances in Peptide Drug Development and Design. *Int. J. Mol. Sci.* **2019**, *20*, 2383. [CrossRef]
28. Marx, V. The big challenges of big data. *Nature* **2013**, *498*, 255. [CrossRef]

29.   Schadt, E.E.; Friend, S.H.; Shaywitz, D.A. A network view of disease and compound screening. *Nat. Rev. Drug Discov.* **2009**, *8*, 286–295. [CrossRef]

30.   Mamoshina, P.; Vieira, A.; Putin, E.; Zhavoronkov, A. Applications of Deep Learning in Biomedicine. *Mol. Pharm.* **2016**, *13*, 1445–1454. [CrossRef]

31.   Rein, D.; Ternes, P.; Demin, R.; Gierke, J.; Helgason, T.; Schön, C.; Gierke, J. Artificial intelligence identified peptides modulate inflammation in healthy adults. *Food Funct.* **2019**, *10*, 6030–6041. [CrossRef]

32.   Wall, A.; Kennedy, K.; Cal, R.; Casey, R.; Holton, T.; Adelfio, A.; Khaldi, N. pep_35E7UW, a natural peptide with cutaneous anti-ageing effects discovered within the Oryza sativa proteome through machine learning. *J. Dermatol. Cosmetol.* **2020**, *4*, 109–116. [CrossRef]

33.   Kennedy, K.; Keogh, B.; Lopez, C.; Adelfio, A.; Molloy, B.; Kerr, A.; Wall, A.M.; Jalowicki, G.; Holton, T.A.; Khaldi, N. An Artificial Intelligence Characterised Functional Ingredient, Derived from Rice, Inhibits TNF-α and Significantly Improves Physical Strength in an Inflammaging Population. *Foods* **2020**, *9*, 1147. [CrossRef] [PubMed]

34.   Kennedy, K.; Cal, R.; Casey, R.; Lopez, C.; Adelfio, A.; Molloy, B.; Wall, A.M.; Holton, T.A.; Khaldi, N. The anti-ageing effects of a natural peptide discovered by artificial intelligence. *Int. J. Cosmet. Sci.* **2020**, *42*, 388–398. [CrossRef] [PubMed]

35.   Corrochano, A.R.; Cal, R.; Kennedy, K.; Wall, A.M.; Murphy, N.; Trajkovic, S.; O'Callaghan, S.; Adelfio, A.; Khaldi, N. Characterising the efficacy and bioavailability of bioactive peptides identified for attenuating muscle atrophy within a Vicia faba-derived functional ingredient. *Curr. Res. Food Sci.* **2021**, in press.

36.   Guigas, B.; Sakamoto, K.; Taleux, N.; Reyna, S.M.; Musi, N.; Viollet, B.; Hue, L. Beyond AICA riboside: In search of new specific AMP-activated protein kinase activators. *IUBMB Life* **2009**, *61*, 18–26. [CrossRef] [PubMed]

37.   Freymond, D.; Guignet, R.; Lhote, P.; Passaquin, A.-C.; Rüegg, U.T. Calcium homeostasis and glucose uptake of murine myotubes exposed to insulin, caffeine and 4-chloro-m -cresol. *Acta Physiol. Scand.* **2002**, *176*, 283–292. [CrossRef] [PubMed]

38.   Smyth, G.K. Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Stat. Appl. Genet. Mol. Biol.* **2004**, *3*, 1–25. [CrossRef] [PubMed]

39.   Silver, J.D.; Ritchie, M.E.; Smyth, G.K. Microarray background correction: Maximum likelihood estimation for the normal-exponential convolution. *Biostatistics* **2008**, *10*, 352–363. [CrossRef] [PubMed]

40.   Leek, J.T.; Johnson, W.E.; Parker, H.S.; Jaffe, A.E.; Storey, J.D. *sva: Surrogate Variable Analysis*, R package version 3.10; Bioconductor: New York, NY, USA, 2014.

41.   Wickham, H. *ggplot2*; Springer: New York, NY, USA, 2009.

42.   Di, L. Strategic Approaches to Optimizing Peptide ADME Properties. *AAPS J.* **2015**, *17*, 134–143. [CrossRef]

43.   Ehrenborg, E.; Krook, A. Regulation of Skeletal Muscle Physiology and Metabolism by Peroxisome Proliferator-Activated Receptor δ. *Pharmacol. Rev.* **2009**, *61*, 373–393. [CrossRef]

44.   Zisman, A.; Peroni, O.D.; Abel, E.D.; Michael, M.D.; Mauvais-Jarvis, F.; Lowell, B.B.; Wojtaszewski, J.F.; Hirshman, M.F.; Virkamaki, A.; Goodyear, L.J.; et al. Targeted disruption of the glucose transporter 4 selectively in muscle causes insulin resistance and glucose intolerance. *Nat. Med.* **2000**, *6*, 924–928. [CrossRef]

45.   Vargas, E.; Podder, V.; Sepulveda, M.A.C. Physiology, Glucose Transporter Type 4 (GLUT4). In *StatPearls*; StatPearls Publishing: Treasure Island, FL, USA, 2020.

46.   Kennedy, J.W.; Hirshman, M.F.; Gervino, E.V.; Ocel, J.V.; Forse, R.A.; Hoenig, S.J.; Aronson, D.; Goodyear, L.J.; Horton, E.S. Acute exercise induces GLUT4 translocation in skeletal muscle of normal human subjects and subjects with type 2 diabetes. *Diabetes* **1999**, *48*, 1192–1197. [CrossRef] [PubMed]

47.   Alam, F.; Islam, A.; Khalil, I.; Gan, S.H. Metabolic Control of Type 2 Diabetes by Targeting the GLUT4 Glucose Transporter: Intervention Approaches. *Curr. Pharm. Des.* **2016**, *22*, 3034–3049. [CrossRef] [PubMed]

48.   Lotfy, M.; Singh, J.; Rashed, H.; Tariq, S.; Zilahi, E.; Adeghate, E. Mechanism of the beneficial and protective effects of exenatide in diabetic rats. *J. Endocrinol.* **2013**, *220*, 291–304. [CrossRef] [PubMed]

49.   Chen, L.-N.; Lyu, J.; Yang, X.-F.; Ji, W.-J.; Yuan, B.-X.; Chen, M.-X.; Ma, X.; Wang, B. Liraglutide ameliorates glycometabolism and insulin resistance through the upregulation of GLUT4 in diabetic KKAy mice. *Int. J. Mol. Med.* **2013**, *32*, 892–900. [CrossRef] [PubMed]

50.   Stratton, I.M.; Adler, A.I.; Neil, H.A.W.; Matthews, D.R.; Manley, S.E.; Cull, C.A.; Hadden, D.; Turner, R.C.; Holman, R.R. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): Prospective observational study. *BMJ* **2000**, *321*, 405–412. [CrossRef] [PubMed]

51.   Ostawal, A.; Mocevic, E.; Kragh, N.; Xu, W. Clinical Effectiveness of Liraglutide in Type 2 Diabetes Treatment in the Real-World Setting: A Systematic Literature Review. *Diabetes Ther.* **2016**, *7*, 411–438. [CrossRef] [PubMed]

52.   Sherwani, S.I.; Khan, H.A.; Ekhzaimy, A.; Masood, A.; Sakharkar, M.K. Significance of HbA1c Test in Diagnosis and Prognosis of Diabetic Patients. *Biomark. Insights* **2016**, *11*, 95–104. [CrossRef] [PubMed]

53.   Liu, Q.; Liu, S.; Gao, L.; Sun, S.; Huan, Y.; Li, C.; Wang, Y.; Guo, N.; Shen, Z. Anti-diabetic effects and mechanisms of action of a Chinese herbal medicine preparation JQ-R in vitro and in diabetic KK Ay mice. *Acta Pharm. Sin. B* **2017**, *7*, 461–469. [CrossRef]

54.   Angulo, P. Nonalcoholic Fatty Liver Disease. *N. Engl. J. Med.* **2002**, *346*, 1221–1231. [CrossRef]

55.   Franch-Nadal, J.; Caballeria, L.; Mata-Cases, M.; Mauricio, D.; Giraldez-García, C.; Mancera, J.; Goday, A.; Mundet-Tudurí, X.; Regidor, E. For the PREDAPS Study Group Fatty liver index is a predictor of incident diabetes in patients with prediabetes: The PREDAPS study. *PLoS ONE* **2018**, *13*, e0198327. [CrossRef]

56. Fiorentino, T.V.; Andreozzi, F.; Mannino, G.C.; Pedace, E.; Perticone, M.; Sciacqua, A.; Perticone, F.; Sesti, G. One-Hour Postload Hyperglycemia Confers Higher Risk of Hepatic Steatosis to HbA1c-Defined Prediabetic Subjects. *J. Clin. Endocrinol. Metab.* **2016**, *101*, 4030–4038. [CrossRef]

57. Singh, S.P.; Pati, G.K.; Misra, B.; Kar, S.K.; Panigrahi, M.K.; Meher, C.; Agrawal, O.; Rout, N.; Pattnaik, K.; Bhuyan, P.; et al. A Study of Prevalence of Diabetes and Prediabetes in Patients of Non-Alcoholic Fatty Liver Disease and the Impact of Diabetes on Liver Histology in Coastal Eastern India. *J. Diabetes Mellit.* **2014**, *4*, 290–296. [CrossRef]

58. Liang, W.; Menke, A.L.; Driessen, A.; Koek, G.H.; Lindeman, J.H.; Stoop, R.; Havekes, L.M.; Kleemann, R.; Hoek, A.M.V.D. Establishment of a General NAFLD Scoring System for Rodent Models and Comparison to Human Liver Pathology. *PLoS ONE* **2014**, *9*, e115922. [CrossRef] [PubMed]

59. Kanehisa, M. From genomics to chemical genomics: New developments in KEGG. *Nucleic Acids Res.* **2006**, *34*, D354–D357. [CrossRef] [PubMed]

60. Bouché, C.; Serdy, S.; Kahn, C.R.; Goldfine, A.B. The Cellular Fate of Glucose and Its Relevance in Type 2 Diabetes. *Endocr. Rev.* **2004**, *25*, 807–830. [CrossRef]

61. Kitagawa, T.; Masumi, A.; Akamatsu, Y. Transforming growth factor-beta 1 stimulates glucose uptake and the expression of glucose transporter mRNA in quiescent Swiss mouse 3T3 cells. *J. Biol. Chem.* **1991**, *266*, 18066–18071. [CrossRef]

62. Schwartzenberg-Bar-Yoseph, F.; Armoni, M.; Karnieli, E. The Tumor Suppressor p53 Down-Regulates Glucose Transporters GLUT1 and GLUT4 Gene Expression. *Cancer Res.* **2004**, *64*, 2627–2633. [CrossRef]

63. Shi, H.; Tzameli, I.; Bjørbaek, C.; Flier, J.S. Suppressor of Cytokine Signaling 3 Is a Physiological Regulator of Adipocyte Insulin Signaling. *J. Biol. Chem.* **2004**, *279*, 34733–34740. [CrossRef]

64. Uhlig, T.; Kyprianou, T.; Martinelli, F.G.; Oppici, C.A.; Heiligers, D.; Hills, D.; Calvo, X.R.; Verhaert, P. The emergence of peptides in the pharmaceutical business: From exploration to exploitation. *EuPA Open Proteom.* **2014**, *4*, 58–69. [CrossRef]

65. Tovi, A.; Eidelman, C.; Shushan, S.; Elster, S.; Hagi, A.; Ivchenko, A.; Butilca, G.M.; Zaoui, G.; Alterman, E.; Bar-Oz, L. High Purity Peptides. U.S. Patent PCT/US2008/002869, 12 September 2008.

66. Paradís-Bas, M.; Tulla-Puche, J.; Albericio, F. The road to the synthesis of "difficult peptides. " *Chem. Soc. Rev.* **2015**, *45*, 631–654. [CrossRef] [PubMed]

67. Werner, H.M.; Cabalteja, C.C.; Horne, W.S. Peptide Backbone Composition and Protease Susceptibility: Impact of Modification Type, Position, and Tandem Substitution. *ChemBioChem* **2016**, *17*, 712–718. [CrossRef] [PubMed]

68. Ghosh, S. Peptide therapeutics market. *Chim. Oggi-Chem. Today* **2016**, *34*, 3.

69. Lage, M.J.; Boye, K.S. The relationship between HbA1c reduction and healthcare costs among patients with type 2 diabetes: Evidence from a U.S. claims database. *Curr. Med. Res. Opin.* **2020**, *36*, 1–7. [CrossRef] [PubMed]