Identification and characterisation of peptides from a boarfish (*Capros aper*) protein hydrolysate displaying *in vitro* dipeptidyl peptidase-IV (DPP-IV) inhibitory and insulinotropic activity

Pádraigín A. Harnedy-Rothwella, Chris M. McLaughlinb, Martina B O’Keeffea, Aurélien V. Le Gouica, Philip J. Allsoppb, Emeir M. McSorleyb, Shaun Sharkeyb, Jason Whooleyc, Brian McGovernc, Finbarr P.M. O'Harteb and Richard J. FitzGeralda\*

a Department of Biological Sciences, University of Limerick, Limerick, Ireland

b School of Biomedical Sciences, Ulster University, Coleraine, Co. Derry, Northern Ireland

c Bio-Marine Ingredients Ireland Ltd., Lough Egish Food Park, Castleblaney, Co. Monaghan, Ireland

\*Corresponding author: Tel: +35361202598; Fax: +35361331490

Email address dick.fitzgerald@ul.ie

**Abstract**

Twenty-two novel dipeptidyl peptidase-IV (DPP-IV) inhibitory peptides (with IC50 values <200 µM) and fifteen novel insulinotropic peptides were identified in a boarfish protein hydrolysate generated at semi-pilot scale using Alcalase 2.4L and Flavourzyme 500L. This was achieved by bioassay-driven semi-preparative reverse phase-high performance liquid chromatography fractionation, liquid chromatography-mass spectrometry and confirmatory studies with synthetic peptides. The most potent DPP-IV inhibitory peptide (IPVDM) had a DPP-IV half maximal inhibitory concentration (IC50) values of 21.72 ± 1.08 µM in a conventional *in vitro* and 44.26 ± 0.65 µM in an *in situ* cell-based (Caco-2) DPP-IV inhibition assay. Furthermore, this peptide stimulated potent insulin secretory activity (1.6-fold increase compared to control) from pancreatic BRIN-BD11 cells grown in culture. The tripeptide IPV exhibited potent DPP-IV inhibitory activity (IC50: 5.61 ± 0.20 µM) comparable to that reported for the known DPP-IV inhibitor IPI (IC50: 5.61 µM). Boarfish proteins contain peptide sequences with potential to play a role in glycaemic management *in vivo*.

**Keywords:** bioactive peptide, dipeptidyl peptidase-IV, insulinotropic, boarfish, type 2 diabetes.

1. **Introduction**

Type 2 diabetes mellitus (T2DM) is a complex chronic metabolic condition characterised by insulin resistance and insufficient pancreatic insulin production resulting in hyperglycaemia. Progression of the disease can lead to numerous long-term micro- and macro-vascular complications which increase the risk of premature death (Fowler, 2011). Global data indicates that approximately 415 million people are currently living with T2DM and it is estimated that, if immediate interventions are not taken, this could reach 642 million by 2040 (IDF Atlas, 2015; WHO, 2016). In addition to its effect on the health and well-being of the individual, T2DM and its complications effect national economies through increased direct healthcare and indirect societal costs (WHO, 2016). Therefore, efforts to address this growing challenge have become a major global public health priority.

Incretin hormones such as glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP) stimulate the secretion of insulin in response to nutrient ingestion (Green, Gault, O'Harte, & Flatt, 2004). However, the incretins are rapidly degraded by dipeptidyl peptidase-IV (DPP-IV) resulting in the loss of their insulinotropic properties (Green et al., 2004). Increasing the half-life of endogenous GLP-1 through development of specific DPP-IV inhibitory agents is one approach which has been adopted to aid glycaemic management in T2DM subjects (Li et al., 2018; Deacon, 2018). Numerous synthetic DPP-IV inhibitory drugs such as saxagliptin, sitagliptin, linagliptin and vildagliptin, are currently on the market (Power, Nongonierma, Jakeman, & FitzGerald, 2014; Aroor, Manrique-Acevedo, & DeMarco, 2018). These drugs, which are collectively known as gliptins, are reported to mediate side-effects such as headaches and urinary and upper respiratory tract infections. Therefore, the discovery and application of natural DPP-IV inhibitory compounds with reduced or no side-effects and which can be consumed as part of the normal diet are the subject of much current investigation.

Examination of the literature shows that food-derived proteins/peptides can beneficially modulate biomarkers of T2DM *in vitro* and *in vivo* and they are increasingly being recognised as potential natural biofunctional agents for the prevention and/or management of T2DM (Harnedy et al., 2018a; Harnedy et al., 2018b; Oseguera-Toledo, de Mejía, Reynoso-Camacho, Cardador-Martínez, & Amaya-Llano, 2014; Promintzer & Krebs, 2006). It has been established that proteins, protein hydrolysates, peptides and amino acids can beneficially regulate blood glucose level. The mechanisms involved include direct stimulation of insulin secretion from pancreatic β-cells or by indirect means through inhibition of DPP-IV along with stimulation of the release of the incretin hormones GLP-1 and GIP. However, the exact mechanism(s) and the primary sequences of the proteinaceous components involved have yet to be fully elucidated (Newsholme & Krause, 2012; Oseguera-Toledo et al., 2014). Peptides and amino acids with *in vitro* DPP-IV inhibitory activity have been identified from a variety of food protein sources including those of marine origin (Harnedy, O’Keeffe, & FitzGerald, 2015; Huang, Jao, Ho, & Hsu, 2012; Li-Chan, Hunag, Jao, Ho, & Hsu, 2012; McLaughlin et al., 2018; Neves et al., 2017; Neves, Harnedy, O’Keeffe, & FitzGerald, 2017; Wang et al., 2015; Zhang, Zhang, Wang, Chen, & Luo, 2017). Furthermore, the potential of food protein-derived peptides as DPP-IV inhibitors has more recently been investigated *in situ* using Caco-2 cells which express DPP-IV (Caron, Domenger, Dhulster, Ravallec, & Cudennec, 2017). Studies with these human intestinal epithelial cells help provide information on the fate of peptides in terms of cellular permeability and stability to brush border enzyme(s) and the potential of the peptides to subsequently inhibit soluble DPP-IV found in plasma.

While reports have shown that amino acids such as glutamine, alanine, arginine, leucine, phenylalanine, valine, isoleucine and lysine have strong insulin secretory activity, limited information exists on peptides as insulin secretagogues (Dixon, Nolan, McClenaghan, Flatt, & Newsholme, 2003; Morifuji et al., 2010; Newsholme, Brennan & Bender, 2006; Power-Grant et al. 2015). To date, peptides derived from casein and from *Palmaria palmata* proteins have been shown to stimulate the secretion of insulin from cultured pancreatic BRIN-BD11 cells (Drummond et al., 2018; O’Harte et al., 2016).

In our previous study, a protein hydrolysate from boarfish generated at laboratory scale (≈10 g) using Alcalase 2.4L and Flavourzyme 500L had DPP-IV inhibitory and insulinotropic activity *in vitro* (Parthsarathy et al., 2019). The objective of this study was to fractionate, identify and characterize the peptides potentially associated with the observed DPP-IV inhibitory and insulin secretory activity in the boarfishprotein hydrolysate generated at semi-pilot scale (≈70 kg). DPP-IV inhibition was determined using *in vitro* enzyme inhibition and *in situ* cell-based inhibition assays while insulin secretion was assessed using BRIN-BD11 cells grown in culture.

## 2. Materials and methods

### 2.1. Materials and chemicals

Samples of minced deboned boarfish (*Capros aper*) meat was kindly provided by Bio-Marine Ingredients Ireland, Ltd., Killybegs, Co. Donegal, Ireland. H-Gly-Pro-7-amino-4–methyl coumarin (AMC) and Diprotin A (IPI) were purchased from Bachem Feinchemikalien (Bubendorf, Switzerland). HPLC grade water, acetonitrile and Kjeldahl catalyst tablets were obtained from VWR International (Dublin, Ireland). Electrospray ionisation (ESI) low molecular mass tune mix was from Agilent Technologies (Cork, Ireland). AB Enzymes (Darmstadt, Germany) and Biocatalysts Ltd (Cardiff, Wales, United Kingdom) kindly provided the Corolase® PP and BC pepsin, respectively. Calcium chloride dihydrate (CaCl2.2H20), D-glucose, HEPES, hydrochloric acid (HCl), magnesium sulphate (MgSO4.7H2O), potassium dihydrogen orthophosphate (KH2PO4), potassium chloride (KCl), sodium bicarbonate (NaHCO3) and sodium chloride (NaCl) were purchased from BDH Chemicals Ltd (Poole, Dorset, UK). Foetal bovine serum (FBS), penicillin-streptomycin (0.1g/L) and RPMI-1640 culture media were obtained from Gibco Life Technologies Ltd (Paisley, Strathclyde, UK). Rat insulin standard was purchased from Novo Industria (Copenhagen, Denmark). Low nitrogen (40% (w/v)) sodium hydroxide was obtained from TE Laboratories Ltd. (Carlow, Ireland). Sulphuric acid (low nitrogen) was purchased from Lennox Laboratory Supplies Ltd (Dublin, Ireland). All other reagents including sodium hydroxide (32% w/v), DPP-IV, from porcine kidney (≥10 units/mg protein), Alcalase® 2.4L and Flavourzyme®500L, amphotericin B, glutamine and the human caucasian colon adenocarcinoma (Caco-2) cells were purchased from Sigma Chemical Company Ltd. (Wicklow, Ireland). All synthetic peptides utilised were from GenScript (Piscataway, NJ).

**2.2. Generation of the boarfish protein hydrolysate at semi-pilot scale**

Minced deboned boarfish meat (720 kg) was thawed overnight at room temperature and suspended in 1,200 L H2O heated to 50 oC in a 6 m³ jacketed tank. The boarfish suspension at a protein concentration of 6.0% (w/v) at 50oC was adjusted to pH 7.0 with 32 % (w/v) NaOH. Alcalase 2.4L and Flavourzyme 500L were added at an enzyme-to-substrate ratio (E:S) of 0.67% (v/w) on a protein basis. The pH of the reaction mixture was maintained constant throughout hydrolysis by the addition of 32 % (w/v) NaOH. After 3.75 h incubation at 50oC, the temperature of the suspension was increased to 75 oC over 20 min and maintained at 75oC for a further 20 min. The protein/peptide components were separated from the sediment and oil using a 2 step decanter/centrifugion process. The oil and sediment were first separated from the supernatant using a 3-phase decanter centrifuge (GEA*,* Düsseldorf, Germany). The supernatant was then further clarified using a continuous centrifuge (GEA*,* Düsseldorf, Germany) and the solution was cooled and stored overnight at 7°C. The clarified hydrolysate was pasteurised at 85oC for 85 s, evaporated to ~40% total solids and spray dried using an inlet and outlet temperature of 180 and 90oC, respectively. The spray-dried hydrolysate was stored in airtight containers until used.

**2.3. Protein quantification**

The protein equivalent content of the boarfish protein hydrolysate was determined using the macro-Kjeldahl procedure as described previously by Connolly, Piggott and FitzGerald (2013) using a Nitrogen to protein conversion factor of 6.25 (Kristinsson & Rasco, 2000). All analyses were performed in triplicate.

**2.4. Simulated gastrointestinal digestion (SGID)**

SGID of the boarfish protein hydrolysate was performed as described by Walsh et al.(2004) with modifications. In brief, an aqueous solution of the hydrolysate at 2.0 % (w/v) on a protein equivalent basis was adjusted to pH 2 and incubated for 90 min at 37˚C with 2.5% (w/w) pepsin. The pH was then adjusted to pH 7 and the sample was incubated at 37˚C with 1% (w/w) Corolase PP for 150 min. Thermal inactivation of the enzyme was performed by heating at 80˚C for 20 min.

**2.5. *In vitro* quantification of DPP-IV inhibitory activity**

The DPP-IV inhibitory activity of hydrolysates and peptides was determined as described by Harnedy et al. (2015). Activity was expressed as mean IC50 value (inhibitory concentration that inhibits DPP-IV by 50%) ± SD from three independent replicates (n=3). Diprotin A (IPI) was used as a positive control.

**2.6. *In situ* quantification of DPP-IV inhibitory activity in Caco-2 cell**

The inhibition of DPP-IV in Caco-2 cells was determined according to the method described by Caron et al. (2017) with some modifications. Cells from passage 44 to 46 were used as a source of DPP-IV in this study. Caco-2 cells were grown at 37°C at 5% CO2 in minimum essential Eagle medium supplemented with 10% fetal bovine serum containing 100 U/mL penicillin, 100 μg/mL streptomycin, 0.25 μg/mL amphotericin B and 2 mM glutamine. Following 5 d of growth the cells were trypsinised and seeded at a density of 20,000 cells/well in a 96-well optical black plate (Costar, Corning, NY, USA). After 16 h of contact, the culture medium was removed from the 96-well plate and cells were washed with 100 μL phosphate buffered saline (PBS, 0.01 M, pH 7.4). Fresh PBS (125 μL) was added to the well followed by 25 μL of each peptide solution or PBS (no inhibition control). Peptide solutions were prepared in PBS and assessed at concentrations ranging from 0.025 to 2.500 mM. The DPP-IV inhibition reaction was initiated by the addition of 50 μL of substrate (Gly-Pro-AMC) and the fluorescence was recorded every 2 min for 1 h using a plate reader (Biotek Synergy HT, USA) at 37°C. Excitation and emission wavelengths were at 360 and 460 nm, respectively. DPP-IV inhibition was defined as the percentage of DPP-IV activity inhibited by a given concentration of peptide compared to the control. Activity was expressed as mean IC50 value ± SD obtained from three independent replicates (n=3).

**2.7. Insulin secretion from BRIN‑BD11 cells**

The *in vitro* insulinotropic effects of protein hydrolysates and peptides were determined using clonal pancreatic BRIN-BD11 β-cells as described previously by Harnedy et al. (2018a). All analyses were performed in triplicate. The insulinotropic activity was expressed as the mean fold increase in the concentration of insulin secreted compared to the 5.6 mM glucose control for the hydrolysates analysed and as the mean insulin concentration secreted (ng/ml) for the peptides which were analysed.

**2.8. Fractionation of the SGID sample of the boarfish protein hydrolysate**

The hydrolysate was fractionated by semi-preparative reverse phase-high performance liquid chromatography (SP-RP-HPLC) using the method described by Harnedy et al. (2015) with modifications. An SGID treated sample of the boarfish protein hydrolysate was reconstituted to a concentration of 80 mg/mL in mobile phase A (HPLC-grade H2O) and 500 µL was injected onto a C18 semi-preparative column (250 x 15 mm I.D., 10 µM particle size, Phenomenex, Cheshire, UK) attached to a C18 guard column (Phenomenex). The flow rate was set at 5 mL/min and the sample was separated using the following gradient (mobile phase B: 80% (v/v) ACN): 0–10 min: 0% B; 10-56 min: 0-40% B; 56-61 min: 40-100% B; 61-66 min 100% B; 66-71 min: 100-0% B; 71–76 min 0% B. The absorbance of the eluent was monitored at 214 nm. Fractions eluting between 6-12, 12-15, 15-21, 21-25, 25-29, 29-33 and 33-45 min were collected using a fraction collector (Waters, Milford, MA, USA) and dried using a solvent evaporator (Genevac, EZ-2 Plus, Genevac Ltd., Ipswich, UK). Aliquots of individual fractions with high activity (F26-29 and F13-15) were also collected every min and dried as described above.

**2.9. Peptide identification by UPLC-ESI-MS and MS/MS**

The most active semi-preparative RP-HPLC fractions, F28 and F29, were further separated using an ACQUITY UPLC (Waters, Milford, MA, USA) and analysed with a micrOTOF Q II mass spectrometer (Bruker Daltonics, Bremen, Germany). The peptide fractions were reconstituted to a concentration of 0.1 mg/mL in mobile phase A (0.1% FA in HPLC grade H2O). Mobile phase B was 0.1% FA in 80% (v/v) ACN. The flow rate was set at 0.2 mL/min. A 2 µL aliquot was injected onto an ACQUITY BEH 300 C18 RP column (2.1 x 50 mm, 1.7 μm; Waters, Dublin, Ireland). Peptides were separated using the following gradient: 0–0.28 min: 0% B; 0.28-45 min: 0-80% B; 45-46 min 80-100% B; 46-48 min: 100% B; 48-49 min 100-50% B; 49-50 min 50% B; 50-51 min: 50-0% B; 51-53 min: 0% B and the eluent was monitored at 214 and 280 nm. The MS/MS methods employed for the identification of peptide masses were as previously described (O'Keeffe & FitzGerald, 2015; O'Keeffe, Norris, Alashi, Aluko, & FitzGerald, 2017). All MS/MS spectra were searched against the SwissProt *Chordata* database using PEAKS Studio 6.0 software (Bioinformatics Solution Inc, Waterloo, Canada). *De novo* sequences generated using PEAKS Studio 6.0 software were searched for homology using an in-house database. The database consisted of 5,300 reviewed proteins from *Actinopterygii*, the superclass within which *Capros aper* (boarfish) is classified, along with 132 unreviewed protein sequences from the family *Caproidae*(boarfishes)**.**

**2.10. Statistical analysis**

All statistical analyses were performed using the statistical software program SPSS (Version 22, IBM Inc., Chicago, IL, USA). Statistical significance at a level of p<0.05 was determined using one-way analysis of variance (ANOVA) followed by Tukey’s and Games–Howell post-hoc tests, where applicable. Furthermore, independent-samples T-tests were performed where applicable.

1. **Results and Discussion**

**3.1. Bioassay-guided fractionation of peptides from the boarfish protein hydrolysate subjected to SGID**

As previously outlined, a boarfish protein hydrolysate generated at laboratory scale with Alcalase 2.4L and Flavourzyme 500L was shown to have good DPP-IV inhibitory and insulin stimulatory activity *in vitro* (Parthsarathy et al., 2019). The boarfish protein hydrolysate generated herein under similar conditions at semi-pilot scale using Alcalase 2.4L and Flavourzyme 500L also showed potent DPP-IV inhibitory (IC50: 1.46 ± 0.06 mg/mL) and insulin secretory activity from BRIN-BD11 cells (5.16 ± 0.27-fold increase compared to 5.6 mM glucose control). Simulating *in vitro* gastrointestinal digestion is a commonly used approach to mimic gastric and intestinal digestion and to aid in determining the potential metabolic fate and alteration in bioactivity of functional food ingredients during digestion. The gastric enzyme pepsin and the intestinal enzyme preparation Corolase PP (containing trypsin, chymotrypsin and elastase) were utilised herein to mimic gastric and intestinal digestion, respectively. While static digestion models can provide information on the potential fate of peptides as they pass through the gastrointestinal tract, they have known limitations and cannot mimic the complex dynamics of the digestion process or the physiological interactions with the host *in vivo* (Brodkorb et al., 2019). For example, they do not mimic the variation in stomach pH following meal intake or the varying enzyme to substrate ratio, the gradual addition of gastrointestinal fluids, interactions with minerals within gastrointestinal fluids, gradual gastric emptying, the absorption of digestion products as they are produced and in addition they do not mimic digestion by gut microbiota.

No significant difference (p>0.05) was observed in DPP-IV inhibitory and insulinotropic activities before and after SGID, with post-SGID activities for DPP-IV inhibition equal to an IC50 of 1.42 ± 0.06 mg/mL and an insulin stimulatory activity equivalent to a 4.83 ± 0.49- fold increase versus the control. As shown in Supplementary Figure S1, no significant differences were observed in the overall peptide profile of the hydrolysate pre- and post-SGID. This indicated that the hydrolysate was resistant to gastrointestinal enzyme degradation.

The SGID treated hydrolysate was fractionated by SP-RP-HPLC in order to identify those peptides potentially associated with the observed DPP-IV inhibitory and insulin secretory activities. As shown in Figure 1, the more hydrophobic fractions, F26-29, F30-33 and F34-45 had the most potent DPP-IV inhibitory activity when assessed at 1 mg/mL. F26-29 was shown to mediate significantly (p<0.05) more potent DPP-IV inhibitory activity (IC50: 0.26 ± 0.01 mg/mL) than F30-33 and F34-45 which had IC50 values of 0.32 ± 0.01and 0.39 ± 0.00 mg/mL, respectively. Fractions 7 to 12 (F7-12) were taken at one min intervals (as opposed to a combined fraction) in the initial experiments as the hydrolysate contained a significant quantity of salt (4.99% (w/w)). As F26-29 showed significantly higher levels of DPP-IV inhibitory than all other fractions and had good insulin secretory activity, this region of the chromatogram (25-29 min) was further fractionated into 1 min fractions in order to assess the bioactivities of the individual fractions. F28 showed significantly higher (p<0.01) DPP-IV inhibitory activity (IC50 : 0.23 ± 0.01 mg/mL) than F26, F27 and F29, while F29 stimulated significantly higher (p<0.05) insulin secretory activity from BRIN-BD11 cells compared to F26, F27 and F29 (5.38 ± 0.42-fold increase compared to the control).

**3.2. Identification and *in vitro* characterisation of anti-diabetic boarfish protein-derived peptides**

MS analysis was performed on F28 and F29 in an attempt to identify the peptides associated with the observed DPP-IV inhibitory and insulinotropic activity, respectively. The peptides identified using a database driven search (UniprotKB/SwissProt, limited to *Chordata*) and *de novo* sequencing in F28 are listed in Tables 1 and S1, respectively. The peptides identified in F29 in a similar manner are listed in Tables 2 and S2. Representative mass fragmentation spectra for three boarfish derived peptides are shown in supplementary data (Figure S2-4). A number of peptides identified in F28 and F29 (Tables 1 and S1 and 2 and S2, respectively) were selected for chemical synthesis. The peptides from F28 and F29 which were synthesised for confirmatory DPP-IV inhibition studies from F28 and from F29 (which also showed good DPP-IV activity, Figure 1) were selected based on their homology to an in-house database and on the basis of structural features known to influence DPP-IV inhibitory activity (Nongonierma & FitzGerald, 2017). While reports have shown that amino acids such as glutamine, alanine, arginine, leucine, phenylalanine, valine, isoleucine and lysine have strong insulin secretory activity, limited information is available on the structure-activity relationship of insulinotropic peptides. In addition to homology to the in-house database, peptides containing all amino acids with known insulin secretory activity were selected from F29 for chemical synthesis. All synthetic peptides were initially screened for DPP-IV inhibitory (% inhibition) activity at 1500 µM and peptides showing inhibitory activity above 80% were further assessed to determine their IC50 values. As shown in Table 3, twenty-two novel peptides with potent DPP-IV inhibitory activity (IC50 values <200 µM) were identified. The majority of these peptides contain a proline residue at the penultimate N-terminal position. This is in agreement with previous reports where the presence of proline at position 2 was shown to influence peptide DPP-IV inhibitory activity (Nongonierma & FitzGerald, 2016). Of all the peptides analysed, the highest DPP-IV inhibitory activity was observed with IPVDM, which had an IC50 value of 21.72 ± 1.08 µM. To our knowledge, this peptide is the 3rd most potent DPP-IV inhibitory food protein-derived peptide identified to date, after IPI and VPL, which have IC50 values of 3.2 and 15.8 µM, respectively (Nongonierma & FitzGerald, 2017). IPI was used as a positive control herein and the IC50 value obtained (3.49 ± 0.19 µM) was is in good agreement with that previously reported by Nongonierma et al. (2017). IPVDM contains isoleucine and proline in position 1 and 2, respectively. A number of tripeptides with IP at the N-terminus have also shown potent DPP-IV inhibitory activity. These include IPM, IPY, IPA and IPF which have IC50 values of 69.5, 38.7, 49.0 and 47.3 µM, respectively (Nongonierma & FitzGerald, 2016; Cermeño et al., 2019; Nongonierma & FitzGerald, 2017; Nongonierma et al., 2019). As IPVDM has valine (which has a similar structure to isoleucine) in position 3, the peptides IPV and IP were also synthesised and assessed for DPP-IV inhibitory activity with a view to determining the role of the N-terminal sequence on DPP-IV inhibitory activity as well as the role of leucine/isoleucine on same. IPV was highly potent with a DPP-IV IC50 value of 5.61 ± 0.20 µM. However, this activity was significantly lower (p<0.01) than that obtained for Diprotin A, i.e., IPI (3.49 ± 0.19 µM). IP had a DPP-IV IC50 of 112.40 ± 5.93 µM which was similar to that reported by Nongonierma & FitzGerald (2014). Other peptides identified herein with leucine/alanine in position 1, a proline in position 2 and valine in position 3 also showed potent DPP-IV inhibitory activity. These include LPVYD, LPVDM and APVP which had DPP-IV IC50 values of 51.36, 53.50 and 73.15 µM, respectively (Table 3). Cermeño et al. (2019) have also shown that similar peptides, IPVP and LPVP, derived from a brewers’ spent grain protein hydrolysate, possess good *in vitro* DPP-IV inhibitory activity with IC50 values of 38.67 and 105.25 µM, respectively. Furthermore, for those peptides identified herein with proline at position 2 the presence of either isoleucine or leucine in certain positions was shown to have a significant effect on the DPP-IV inhibitory potency of the peptide. When isoleucine was at position 1 the DPP-IV inhibitory activity was shown to be significantly higher than that mediated by the corresponding peptide with leucine at position 1 (e.g., IPGA/LPGA and IPVDM/LPDVM, p<0.001, (Table 3)). A similar result was observed when either isoleucine or leucine was in position 3 (e.g., APIT/APLT and GPIN/GPLN, p<0.001 (Table 3)). Cermeño et al. (2019) also showed that peptides derived from a brewers’ spent grain protein hydrolysate with proline at position 2 had higher DPP-IV inhibitory activity when isoleucine was at position 1 and 3 compared to when leucine was present at these positions. The presence of either isoleucine or leucine at position 4 for the peptides identified herein appears to have no effect on DPP-IV inhibitory activity (e.g., GPGI/GPGL, TPGI/TPGL and GPSI/GPSL (p>0.05, Table 3)). In general, peptides with proline in position 2, which also contained posttranslational modifications, showed moderate to low DPP-IV inhibitory activity (Table 3). While twenty-two of the peptides identified herein with DPP-IV IC50 values <200 µM were shown to be novel, two of the peptides, GPGL and LPGA, have previously been found within the sequence of marine (rotifer and grass carp)-derived antioxidant peptides: LLGPGLTNHA, GFGPGL and DLGLGLPGAH (Byun, Lee, Park, Jeon, & Kim, 2009; Cai et al., 2015).

The use of an *in vitro* assay employing purified DPP-IV provides indicative information on the ability of a peptide to inhibit DPP-IV in a highly controlled environment. However, if peptides are to be used as functional food ingredients their efficacies need to be validated *in vivo.* Taking into consideration the financial and ethical costs of *in vivo* studies, Caron et al. (2017) developed a cell-based DPP-IV inhibition assay using human intestinal epithelial Caco-2 cells which express DPP-IV. This assay may provide some information on the potential fate of a peptide *in vivo*. It may also provide some information on the potential of the peptide to reach and thus inhibit soluble DPP-IV found in plasma. In this assay, intestinal mucosal conditions are simulated where peptides are required to pass through human intestinal cell membranes and resist degradation by brush border enzymes in order to mediate their activity. Peptides showing good DPP-IV inhibitory activity (IC50 values <200 µM) in the *in vitro* enzyme inhibition assay were therefore assessed for DPP-IV inhibitory activity in the *in situ* Caco-2 cell based assay. As shown in Table 3, lower, but in the majority of cases comparable, DPP-IV inhibitory activity was observed in the cell-based assay compared to the conventional *in vitro* assay. In contrast to the results reported by Caron et al. (2017), who observed similar DPP-IV inhibitory activity in both assays for the positive control IPI, a significantly higher IC50 value (p<0.001) was observed for IPI herein in the cell-based (9.93 ± 0.63 µM) compared to the *in vitro* assay (3.49 ± 0.19 µM, Table 3). However, it must be noted that the enzyme utilised in the *in vitro* assay herein was porcine-derived DPP-IV. We have previously shown that the DPP-IV inhibitory activity of IPI when assessed using purified human DPP-IV (IC50: 5.00 ± 0.03 µM) was significantly lower (p<0.001) than that with the porcine enzyme (3.49 ± 0.19 µM, Table 3, Harnedy & FitzGerald, 2013). This was also shown with other peptides where Lacroix and Li-Chan (2015) reported that in general, porcine-derived DPP-IV was inhibited to a greater extent by protein-derived peptides than the human enzyme. While the sequence of DPP-IV is known to be highly conserved among mammalian species, the 12% difference in homology across the full porcine and human DPP-IV sequence and more specifically the 8% difference in homology between the sequence of the catalytic site may have an effect on activity (Lacroix, et al., 2015). The peptide which showed the highest DPP-IV inhibitory activity in the *in vitro* assay, IPVDM (IC50: 21.72 ± 1.08 µM), also showed the highest activity in the cell-based assay (44.26 ± 0.65 µM). This indicates that IPVDM is potentially resistant to degradation by membrane associated peptidases. In contrast other peptides, e.g.., GPSI, were shown to mediate significantly lower activity in the cell-based assay (> 312.50 µM) when compared to the *in vitro* assay (72.85 ± 1.66 µM). It is possible that these peptides were not efficiently transported across the cell membrane or were degraded during membrane transit.

Six peptides all containing amino acid residues with known insulinotropic activity (VVVT, AVLQ, AVIQ, QQLA, QQIA and QF) along with peptides identified in F29 with high DPP-IV inhibitory activity in the *in vitro* assay were selected for assessment of their insulinotropic activity. As shown in Figure 2, VPDPR, GPGI, LPVDM, IPVDM, MPAVP, GPSL, GPSI, AVIQ and QQIA (all p<0.001), VPTP, APLER, GPIN, APVP (all p<0.01) and GPLN and VVVT (all p<0.05) stimulated a significantly higher insulinotropic response from BRIN BD11 cells (1.3 to 1.7-fold increase) when assessed at 10-6 M compared to the control (5.6 mM glucose) which secreted 3.93 ± 0.26 ng/mL of insulin. To our knowledge, none of these peptides have previously been reported to mediate insulinotropic activity. Furthermore, three peptides (IPVDM, MPAVP and GPSL) stimulated the secretion of significantly (p<0.001) higher concentrations of insulin from BRIN-BD11 cells compared to the control when assessed at a lower concentration (10-10 M, data not shown). All three peptides (in particular IPVDM) also showed potent DPP-IV inhibitory activity (Table 3). This would suggest that peptides such as IPVDM may contribute to glycaemic control by multiple mechanisms. These include direct or indirect stimulation of insulin secretion due to the presence of higher levels of circulating active GLP-1 arising from peptide mediated inhibition of DPP-IV. As previously stated, amino acids such as glutamine, alanine, arginine, leucine, phenylalanine, valine, isoleucine and lysine have been shown to mediate potent insulin secretory activity. However, of the six peptides tested (which were composed entirely of known insulin secretory amino acids), only three showed significantly higher insulin secretory activity compared to the control when assessed at 10-6 M (Figure 2). These include AVIQ and QQIA (p<0.001), and VVVT (p<0.05). Nevertheless, this activity was lower than the activity mediated by specific peptides containing ≤ 2 amino acids with known insulinotropic activity (e.g., VPDPR, GPGI, IPVDM, MPAVP, GPSL, GPSI). To date there are two reports on peptides identified from food proteins which show insulin secretory activity from BRIN BD11 cells (Drummond et al., 2018; O’Harte et al., 2016). No information is as yet available on the structure-activity relationship of insulinotropic peptides. While no definite interpretations can be drawn on the structure activity relationship of food-derived insulinotropic peptides, some preliminary observations can be made based on the results presented herein. Comparing the activity of IPVDM and LPVDM, it was seen that IPVDM had higher mean insulinotropic activity, albeit not significant, than LPVDM. A similar result was observed with peptides identified from a *Palmaria palmata* protein hydrolysate where ILAP was shown to exhibit higher insulin secretory activity than LLAP (O’Harte et al., 2016). For sequences such as AVLQ/AVIQ and QQLA/QQIA where leucine/isoleucine is in position 3, peptides with isoleucine in that position had significantly higher (p<0.01) insulin secretory activity than corresponding peptides with leucine at this position. When the activity mediated by PAVP was compared to APVP and MPAVP, significantly lower (p<0.001) insulinotropic activity was observed with PAVP compared to that observed with APVP and MPAVP .

**4. Conclusion**

Peptides showing potent *in vitro* DPP-IV inhibitory activity (in a conventional *in vitro* and in a cell-based *in situ* assay) and insulin secretory activity from cultured pancreatic BRIN-BD11 cells were released from boarfish muscle protein following hydrolysis at semi-pilot scale using Alcalase 2.4L and Flavourzyme 500L. Confirmatory studies with synthetic peptides identified 22 novel DPP-IV inhibitory peptides with IC50 values <200 µM along with 15 novel insulinotropic peptides. The DPP-IV inhibitory peptide, IPVDM, which showed potent DPP-IV inhibitory activity in the *in vitro* assay also showed potent activity in a cell-based DPP-IV inhibitory assay. Furthermore, IPVDM, mediated potent insulin secretory activity from cultured pancreatic BRIN-BD11 cells. These results indicate that IPVDM may be transported across the human epithelial cell membrane in an intact form and that it may mediate anti-diabetic effects *in vivo* via multiple mechanisms. However, further studies are required to evaluate the *in vivo* anti-diabetic activity of this peptide and other peptides identified herein.

The successful transfer of the hydrolysis process from laboratory to semi-pilot scale and stabilisation of the peptides through spray-drying would indicate that the boarfish protein-derived hydrolysate herein can be manufactured at industrial scale and may be employed as a functional ingredient for the prevention and management of T2DM. Furthermore, the value of boarfish, a protein rich underutilised fish species, which is available in significant quantities to the marine industry may be significantly enhanced.

Acknowledgments

This research was funded under the National Development Plan, through the Food Institutional Research Measure, administered by the Department of Agriculture, Food and the Marine, Ireland under grant numbers 11/F/063, [11/F/064](http://www.sciencedirect.com/science/article/pii/S1756464617302074), [13/F/467](http://www.sciencedirect.com/science/article/pii/S1756464617302074), [13/F/536](http://www.sciencedirect.com/science/article/pii/S1756464617302074) and [14/F/873](http://www.sciencedirect.com/science/article/pii/S1756464617302074) and a Northern Ireland Department of Employment and Learning PhD scholarship for Chris McLaughlin. The research was also part funded by the Science Foundation Ireland Infrastructure Fund and the Higher Education Authority under the Programme for Research in Third Level Institutions (cycle 4). The authors would like to acknowledge Dr. Thanyaporn Kleekayai for growth of cells.

**References**

Aroor, A. R., Manrique-Acevedo, C., & DeMarco, V. G. (2018). The role of dipeptidyl peptidase-4 inhibitors in management of cardiovascular disease in diabetes; focus on linagliptin. *Cardiovascular Diabetology*, *17*, 59.

Brodkorb, A., Egger, L., Alminger, M., Alvito, P., Assunção, R., Balance, S., Bohn, T., Bourlieu-Lacanal, C., Boutrou, R., Carrière, F., Clemente, A., Corredig, M., Dupont, D., Dufour, C., Edwards, C., Golding, M., Karakaya, S., Kirkhus, B., Le Feunteun, S., Lesmes, U., Macierzanka, A., Mackie, A. R., Martins, C., Marze, S., McClements, D. J., Ménard, O., Minekus, M., Portmann, R., Santos, C. N., Souchon, I., Singh, R. P., Vegarud, G. E., Wickham, M. S. J., Weitschies, W. & Recio I. (2019) INFOGEST static *in vitro* simulation of gastrointestinal food digestion, [*Nature Protocols*](https://www.ncbi.nlm.nih.gov/pubmed/30886367), *14*, 991-1014.

Byun, H. G., Lee, J. K., Park, H. G., Jeon, J. K., & Kim, S. K. (2009). Antioxidant peptides isolated from the marine rotifer, *Brachionus rotundiformis*. *Process Biochemistry, 44*, 842–846.

Cai, L., Wu, X., Zhang, Y., Li, X., Ma, S., & Li, J. (2015). Purification and characterization of three antioxidant peptides from protein hydrolysate of grass carp (*Ctenopharyngodon idella*) skin. *Journal of Functional Foods, 16*, 234-242.

Caron, J., Domenger, D., Dhulster, P., Ravallec, R., & Cudennec, B. (2017). Using Caco-2 cells as novel identification tool for food-derived DPP-IV inhibitors. *Food Research International, 92*, 113-118.

Cermeño, M., Connolly, A., O'Keeffe, M. B., Flynn, C., Alashi, A. M., Aluko, R. E. & FitzGerald, R. J. (2019). Identification of bioactive peptides from brewers’ spent grain and contribution of Leu/Ile to bioactive potency. *Journal of Functional Foods*, *60,* 103455.

Connolly, A., Piggott, C. O., & FitzGerald, R. J. (2013). Characterisation of protein-rich isolates and antioxidative phenolic extracts from pale and black brewers’ spent grain. *International Journal of Food Science and Technology, 48*, 1670-1681.

Deacon, C. F. (2018) Peptide degradation and the role of DPP-4 inhibitors in the treatment of type 2 diabetes. *Peptides, 100*, 150-157.

Dixon, G., Nolan, J., McClenaghan, N., Flatt, P. R., & Newsholme, P. (2003). A comparative study of amino acid consumption by rat islet cells and the clonal beta-cell line BRIN-BD11 - the functional significance of L-alanine. *Journal of Endocrinology, 179*(3), 447-454.

Drummond, E., Flynn, S., Whelan, H., Nongonierma, A. B., Holton, T. A., Robinson, A., Egan, T., Cagney, G., Shields, D. C., Gibney, E. R., Newsholme, P., Gaudel, C., Jacquier, J.-C., Noronha, N., FitzGerald, R. J., & Brennan, L. (2018). Casein hydrolysate with glycemic control properties: evidence from cells, animal models, and humans. *Journal of Agricultural and Food Chemistry, 66*(17), 4352-4363.

Fowler, M. J. (2011). Microvascular and macrovascular complications of diabetes. *Clinical Diabetes, 29*, 116-122.

Green, B. D., Gault, V. A., O'Harte, F. P., & Flatt, P. F. (2004). Structurally modified analogues of glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP) as future antidiabetic agents. *Current Pharmaceutical Design, 10*(29), 3651-3662.

Harnedy, P. A., & FitzGerald, R. J. (2013). *In vitro* assessment of the cardioprotective, anti-diabetic and antioxidant potential of *Palmaria palmata* protein hydrolysates. *Journal of Applied Phycology, 25*(6), 1793-1803.

Harnedy, P. A., O’Keeffe, M. B., & FitzGerald, R. J. (2015). Purification and identification of dipeptidyl peptidase (DPP) IV inhibitory peptides from the macroalga *Palmaria palmata*. *Food Chemistry, 172*, 400-406.

Harnedy, P. A., Parthsarathy, V., McLaughlin, C. M., O'Keeffe, M. B., Allsopp, P. J., McSorley, E. M., O'Harte, F. P. M., & FitzGerald, R. J. (2018a). Atlantic salmon (*Salmo salar*) co-product-derived protein hydrolysates: A source of antidiabetic peptides. *Food Research International, 106*, 598-606.

Harnedy, P. A., Parthsarathy, V., McLaughlin, C. M., O'Keeffe, M. B., Allsopp, P. J., McSorley, E. M., O'Harte, F. P. M., & FitzGerald, R. J. (2018b). Blue whiting (*Micromesistius poutassou*) muscle protein hydrolysate with *in vitro* and *in vivo* antidiabetic properties. *Jounal of Functional Foods, 40*, 137-145.

Huang, S. L., Jao, C. L., Ho, K. P., & Hsu, K. C. (2012). Dipeptidyl-peptidase IV inhibitory activity of peptides derived from tuna cooking juice hydrolysates. *Peptides, 35*, 114-121.

IDF Atlas. (2015). International Diabetes Federation.

Kristinsson, H. G., & Rasco, B. A. (2000). Biochemical and functional properties of Atlantic salmon (*Salmo salar*) muscle proteins hydrolyzed with various alkaline proteases. *Journal of Agricultural and Food Chemistry, 48*(3), 657-666.

Lacroix, I. M. E., & Li-Chan, E. C. Y. (2015). Comparison of the susceptibility of porcine and human dipeptidyl-peptidase IV to inhibition by protein-derived peptides. *Peptides, 69*, 19-25.

Li, N., Wang, L. J., Jiang, B., Li, X. Q., Guo, C. L., Guo, S. J., & Shi, D. Y. (2018) [Recent progress of the development of dipeptidyl peptidase-4 inhibitors for the treatment of type 2 diabetes mellitus.](https://www.ncbi.nlm.nih.gov/pubmed/29609120) *European Journal of Medicinal Chemistry, 151,* 145-157.

Li-Chan, E. C., Hunag, S. L., Jao, C. L., Ho, K. P., & Hsu, K. C. (2012). Peptides derived from Atlantic salmon skin gelatin as dipeptidyl-peptidase IV inhibitors. *Journal of Agricultural and Food Chemistry, 60*(4), 973-978.McLaughlin, C. M., Harnedy, P. A., Parthsarathy, V., Allsopp, P. J., McSorley, E. M., FitzGerald R. J., & O’Harte F. P. M. (2018). Synthetic peptides identified from *Palmaria palmata* enhance glucagon-like peptide-1 stability *in vitro* and show acute anti-hyperglycaemic and insulinotropic actions in mice. *Proceedings of the Nutrition Society*, *77* (OCE3), E113.

Morifuji, M., Ishizaka, M., Baba, S., Fukuda, K., Matsumoto, H., Koga, J., Kanegae, M., & Higuchi, M. (2010). Comparison of different sources and degrees of hydrolysis of dietary protein: Effect on plasma amino acids, dipeptides, and insulin responses in human subjects. *Journal of Agricultural and Food Chemistry, 58*(15), 8788-8797.

Neves, A. C., Harnedy, P. A., O’Keeffe, M. B., Alashi, M. A., Aluko, R. E., & FitzGerald, R. J. (2017). Peptide identification in a salmon gelatin hydrolysate with antihypertensive, dipeptidyl peptidase IV inhibitory and antioxidant activities. *Food Research International, 100*(1), 112-120.

Neves, A. C., Harnedy, P. A., O’Keeffe, M. B., & FitzGerald, R. J. (2017). Bioactive peptides from Atlantic salmon (*Salmo salar*) with angiotensin converting enzyme and dipeptidyl peptidase IV inhibitory, and antioxidant activities. *Food Chemistry, 218*, 396-405.

Newsholme, P., Brennan, L., & Bender, K. (2006). Amino acid metabolism, β-cell function, and diabetes. *Diabetes, 55*(2), 39-47.

Newsholme, P., & Krause, M. (2012). Nutritional regulation of insulin secretion: implications for diabetes. *The Clinical Biochemist Reviews, 33*, 35-47.

Nongonierma, A. B., & FitzGerald, R. J. (2014). Susceptibility of milk protein-derived peptides to dipeptidyl peptidase IV (DPP-IV) hydrolysis. *Food Chemistry, 14*, 845-852.

Nongonierma, A. B., & FitzGerald, R. J. (2016). Structure activity relationship modelling of milk protein-derived peptides with dipeptidyl peptidase IV (DPP-IV) inhibitory activity. *Peptides, 79*, 1-7.

Nongonierma, A. B., & FitzGerald, R. J. (2017). Features of dipeptidyl peptidase IV (DPP-IV) inhibitory peptides from dietary proteins. *Journal of Food Biochemistry, e12451*, 1-11.

Nongonierma, A. B., Dellafiora, L., Paolella, S., Galaverna, G., Cozzini, P., & FitzGerald, R. J. (2018) *In silico* approaches applied to the study of peptide analogs of Ile-Pro-Ile in relation to their dipeptidyl peptidase IV inhibitory properties. *Frontiers in Endocrinology*, 9, 329.

O'Keeffe, M. B., & FitzGerald, R. J. (2015). Identification of short peptide sequences in complex milk protein hydrolysates. *Food Chemistry, 184*, 140-146.

O'Keeffe, M. B., Norris, R., Alashi, M. A., Aluko, R. E., & FitzGerald, R. J. (2017). Peptide identification in a porcine gelatin prolyl endoproteinase hydrolysate with angiotensin converting enzyme (ACE) inhibitory and hypotensive activity. *Journal of Functional Foods, 34*, 77-88.

O’Harte, F. P. M., Mullan, C., Harnedy, P. A., McLaughlin, C. M., Parthsarathy, V., Allsopp, P. J., McSorley, E. M., & FitzGerald, R. J. (2016). Synthetic peptides from marine origin enhance both incretin hormone stability and insulin secretion *in vitro*. In Proceedings *Diabetes UK Professional Conference*, vol. 33 (pp. 59). Glasgow, Scotland: Diabetic Medicine.

Oseguera-Toledo, M. E., de Mejía, E. G., Reynoso-Camacho, R., Cardador-Martínez, A., & Amaya-Llano, S. L. (2014). Proteins and bioactive peptides: Mechanisms of action on diabetes management. *Nutrafoods, 13*, 147-157.

Parthsarathy, V., McLaughlin, C. M., Harnedy, P. A., Allsopp, P. J., Crowe, W., McSorley, E. M., FitzGerald, R. J., & O’Harte, F. P. M. (2019). Boarfish (*Capros aper*) protein hydrolysate has potent insulinotropic and GLP-1 secretory activity *in vitro* and acute glucose lowering effects in mice. *International Journal of Food Science and Technology*, *54*, 271-281.

Power-Grant, O., Bruen, C., Brennan, L., Giblin, L., Jakeman, P., & FitzGerald, R. J. (2015). *In vitro* bioactive properties of intact and enzymatically hydrolysed whey protein: targeting the enteroinsular axis. *Food & Function, 6*(3), 972-980.

Power, O., Nongonierma, A. B., Jakeman, P., & FitzGerald, R. J. (2014). Food protein hydrolysates as a source of dipeptidyl peptidase IV inhibitory peptides for the management of type 2 diabetes. *Proceedings of the Nutrition Society, 73*(1), 34-36.

Promintzer, M., & Krebs, M. (2006). Effects of dietary protein on glucose homeostasis. *Current Opinion in Clinical Nutrition and Metabolic Care, 9*(4), 463-468.

Walsh, D. J., Bernard, H., Murray, B. A., MacDonald, J., Pentzien, A. K., Wright, G. A., Wal, J. M., Struthers, A. D., Meisel, H., & FitzGerald, R. J. (2004). *In vitro* generation and stability of the lactokinin beta-lactoglobulin fragment (142-148). *Journal of Dairy Science, 87*(11), 3845-3857.

Wang, T. Y., Hsieh, C. H., Hung, C. C., Jao, C. L., Chen, M. C., & Hsu, K. C. (2015). Fish skin gelatin hydrolysates as dipeptidyl peptidase IV inhibitors and glucagon-like peptide-1 stimulators improve glycaemic control in diabetic rats: A comparison between warm- and cold-water fish. *Journal of Functional Foods, 19*, 330-340.

WHO. (2016). World Health Organization: Global report on diabetes.

Zhang, C., Zhang, Y., Wang, Z., Chen, S., & Luo, Y. (2017). Production and identification of antioxidant and angiotensin-converting enzyme inhibition and dipeptidyl peptidase IV inhibitory peptides from bighead carp (*Hypophthalmichthys nobilis*) muscle hydrolysate. *Journal of Functional Foods, 35*, 224-235.

**Table 1** Sequences of the peptides (with flanking N and C terminal residue) identified by liquid chromatography-mass spectrometry/mass spectrometry in boarfish (*Capros aper*) protein hydrolysate fraction 28 (F28) using a database (UniprotKB/SwissProt) search limited to *Chordata*.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Peptide sequence\*** | **Protein: *Species*****(Accession number)** | **Mass /Charge (*m/z*)** | ***z*** | **Error (ppm)** | **Retention time (min)** |
| T.GVDNPGHP.F | Creatine kinase U-type, mitochondrial: Sus scrofa (Q29577) | 396.6825 | 2 | -7.2 | 6.23 |
| L.TEAPLNPK.A | Actin, gamma-enteric smooth muscle: Gallus gallus (P63270) | 435.2393 | 2 | -1.6 | 7.89 |
| T.GIVLD.S | 516.2890 | 1 | -26.7 | 12.4 |
| E.KSYELPDGQ.V | 518.7515 | 2 | 1.2 | 10.32 |
| W.LPVYD.A | Myosin heavy chain, fast skeletal muscle: Cyprinus carpio (Q90339) | 303.6599 | 2 | -1.4 | 11.51 |
| Y.DYPMIS.Q | 725.3174 | 1 | 0.0 | 13.22 |
| A.DSKPGSI.R | Nucleoside diphosphate kinase B: Merluccius senegalensis (P85282) | 352.1855 | 2 | 2.4 | 6.85 |
| T.NPADSKPGSI.R | 493.2503 | 2 | -1.6 | 8.29 |

ppm: mass error, calculated as 106 × (experimental mass - theoretical mass) / theoretical mass. **\***Peptides displayed using one letter code for the amino acids.

**Table 2** Sequences of the peptides (with flanking N and C terminal residue) identified by liquid chromatography-mass spectrometry/mass spectrometry in boarfish (*Capros aper*) protein hydrolysate fraction 29 (F29) using a database (UniprotKB/SwissProt( search limited to *Chordata*.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Peptide sequence\*** | **Protein: *Species*****(Accession number)** | **Mass/ Charge (*m/z*)** | ***z*** | **Error (ppm)** | **Retention time (min)** |
| L.TEAPLNPK.A | Actin, gamma-enteric smooth muscle: Gallus gallus (P63270) | 435.2436 | 2 | 8.3 | 7.72 |
| K.YPIEH(+14.02).G | 336.6758 | 2 | 13.6 | 8.07 |
| S.YELPDGQVIT. I | 567.7916 | 2 | 7.3 | 14.99 |
| K.SYELPDGQVIT. I |  | 611.3044 | 2 | 1.5 | 15.1 |
| A.AFPPDVAGN.V | Myosin regulatory light chain 2, skeletal muscle isoform: Gallus gallus (P02609) | 444.2211 | 2 | 10.4 | 14.63 |
| W.AAFPPDVAGN.V | 958.4749 | 1 | 12.5 | 15.15 |
| T.NPADSKPGSI.R | Nucleoside diphosphate kinase B: Merluccius senegalensis (P85282) | 493.2527 | 2 | 3.3 | 8.35 |
| I.YERPDFGGQ.M | Beta-crystallin S-2: Chiloscyllium indicum (P48647) | 534.7458 | 2 | 9.4 | 10.4 |

ppm: mass error, calculated as 106 × (experimental mass - theoretical mass) / theoretical mass. **\***Peptides displayed using one letter code for the amino acids.

**Table 3** Dipeptidyl peptidase-IV (DPP-IV) inhibitory activities of synthetic boarfish (*Capros aper*) protein-derived peptides obtained using *in vitro* and *in situ* (Caco-2 cell) DPP-IV inhibitory assays.

|  |  |  |
| --- | --- | --- |
| Peptide\* | *In vitro* assay | *In situ* assay |
| Inhibition (%) at 1500 µM | IC50 (µM) | IC50 (µM) |
| IPI (+ve control) | n.d. | 3.49 ± 0.19 | 9.93 ± 0.63 |
| IPVDM | 99.47 ± 0.46 | 21.72 ± 1.08 | 44.26 ± 0.65 |
| APIT | 96.52 ± 0.78 | 34.73 ± 4.66 | 123.40 ± 7.50 |
| VPTP | 100 ± 1.00 | 38.93 ± 2.65 | 66.68 ± 2.18 |
| GPIN | 97.97 ± 1.34 | 48.96 ± 3.93 | 111.40 ± 7.21 |
| LPVYD | 97.76 ± 1.89 | 51.36 ± 3.61 | 60.63± 1.84 |
| LPVDM | 97.79 ± 0.88 | 53.50 ± 6.31 | 150.80 ± 5.44 |
| APLER | 98.66 ± 1.23 | 63.67 ± 5.40 | 153.10 ± 2.69 |
| IPGA | 98.66 ± 0.67 | 66.37 ± 3.84 | 96.32 ± 5.93 |
| GPSL | 95.73 ± 0.47 | 68.13 ± 2.79 | 227.63 ± 6.90 |
| GPSI | 94.86 ± 1.09 | 72.85 ± 1.66 | > 312.5  |
| APVP | 96.39 ± 0.53 | 73.15 ± 3.09 | 284.65 ± 3.32 |
| VPDPR | 99.85 ± 1.46 | 79.10 ± 2.67 | 128.90 ± 4.95 |
| APLDK | 94.87 ± 1.06 | 90.37 ± 1.70 | 223.90 ± 4.38 |
| APLT | 95.61 ± 3.81 | 91.10 ± 11.11 | 165.55 ± 4.17 |
| MPAVP | 93.67 ± 2.36 | 115.27 ± 4.21 | 149.70 ± 13.44 |
| GPGI | 90.89 ± 1.93 | 116.27 ± 8.60 | 307.30 ± 2.69 |
| GPLN | 92.34 ± 1.84 | 116.37 ± 9.97 | > 312.5 |
| PAVP | 88.91 ± 1.65 | 126.51 ± 2.54 | > 312.5 |
| GPGL | 92.59 ± 0.89 | 131.90 ± 8.44 | 245.50 ± 1.98 |
| LPGA | 95.03 ± 0.77 | 154.12 ± 7.04 | 245.45 ± 13.65 |
| AALP | 80.00 ± 1.11 | 164.37 ± 18.78 | > 312.5 |
| TPTV | 91.05 ± 2.11 | 164.52 ± 19.75 | 177.00 ± 10.61 |
| AAIP | 82.40 ± 1.80 | 251.58 ± 24.47 | n.d. |
| YPL{pSER}L | 89.77 ± 4.35 | 261.46 ± 1.17 | n.d. |
| TPGI | 87.41 ± 3.06 | 282.16 ± 12.89 | n.d. |
| TPGL | 85.00 ± 1.97 | 297.41 ± 13.02 | n.d. |
| YPII{pSER} | 83.48 ± 2.50 | 300.67 ± 11.29 | n.d. |
| YPIL{pSER} | 84.28 ± 2.72 | 302.47 ± 18.19 | n.d. |
| ISAP | 80.88 ± 0.90 | 393.88 ± 22.16 | n.d. |
| YPL{pTHR}V | 81.17 ± 1.21 | 511.47 ± 6.36 | n.d. |
| YPLV{pTHR} | 80.55 ± 3.86 | 508.37 ± 33.43 | n.d. |
| ISGP | 66.63 ± 0.92 | n.d. | n.d. |
| LSAP | 57.88 ± 3.34 | n.d. | n.d. |
| GALP | 44.59 ± 1.90 | >1500 | n.d. |
| GAIP | 43.46 ± 1.42 | >1500 | n.d. |
| IGGP | 43.25 ± 2.03 | >1500 | n.d. |
| LGGP | 39.58 ± 4.06 | >1500 | n.d. |
| LSGP | 38.11 ± 2.28 | >1500 | n.d. |
| GLAGQ | 18.04 ± 3.42 | >1500 | n.d. |
| SAGIH | 8.30 ± 0.93 | >1500 | n.d. |
| SAGLH | 6.05 ± 0.79 | >1500 | n.d. |
| AAPF | 4.54 ± 4.42 | >1500 | n.d. |
| FDLP | n.a. | >1500 | n.d. |
| FDIP | n.a. | >1500 | n.d. |
| LDFP | n.a. | >1500 | n.d. |
| IDFP | n.a. | >1500 | n.d. |
| GGLP | n.a. | >1500 | n.d. |
| GGIP | n.a. | >1500 | n.d. |
| VVVT | n.a. | >1500 | n.d. |
| AVLQ | n.a. | >1500 | n.d. |
| AVIQ | n.a. | >1500 | n.d. |
| QQLA | n.a. | >1500 | n.d. |
| QQIA | n.a. | >1500 | n.d. |
| QF | n.a. | >1500 | n.d. |

Mean ± SD (n=3). % inhibition studies were performed at 1500 µM. IC50: inhibitory concentration that inhibits enzyme activity by 50 %. n.d. not determined. n.a. no activity detected. **\***Peptides displayed using one letter code for the amino acids.

**Figure 1.** Semi-preparative reversed phase-high performance liquid chromatography profile of the boarfish (*Capros aper*) protein hydrolysate subjected to simulated gastrointestinal digestion along with the dipeptidyl peptidase-IV (DPP-IV) inhibitory and insulin secretory activity of the associated pooled and individual fractions (F7-F45). Samples were tested for *in vitro* DPP-IV inhibition and insulin secretory activity by pancreatic BRIN-BD11 cells at a concentration of 1.00 and 1.25 mg/mL, respectively. DPP-IV inhibitory data is displayed as % inhibition and as an IC50 value (inhibitory concentration that inhibits enzyme activity by 50 %) for selected fractions. Insulin secretory activity was expressed as fold increase compared to the control (5.6 mM glucose). Values (mean ± SD, n=3) with different superscript letters are significantly different at p<0.05.



|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Fraction** | **DPP-IV inhbition at 1 mg/ml (%)** | **DPP-IV IC50 (mg/ml)** | **Insulin secretion (fold increase versus control)** |  |  |  |  |  |
| F7 | 15.11 ± 2.00g | - | 4.83 ± 0.82 a |  |  |  |  |  |
| F8 | 13.75 ± 1.17g | - | 3.83 ± 0.70 b |  |  |  |  |  |
| F9 | 29.18 ± 1.43f | - | 5.04 ± 0.49 a |  |  |  |  |  |
| F10 | 42.35 ± 2.05e | - | 2.33 ± 0.34 c |  |  |  |  |  |
| F11 | 57.74 ± 2.68cd | - | 1.25 ± 0.15 d |  |  |  |  |  |
| F12 | 56.33 ± 1.36cd | - | 1.00 ± 0.10 d |  |  |  |  |  |
| F13-15 | 54.17 ± 1.57d | - | 4.71 ± 0.43 a |  | **Fraction** | **DPP-IV inhibition at 1 mg/ml (%)** | **DPP-IV IC50 (mg/ml)** | **Insulin secretion (fold increase v control)** |
| F16-21 | 58.95 ± 1.82cd | - | 4.27 ± 0.34 a |  |
| F22-25 | 61.47 ± 2.54c | - | 4.71 ± 0.38 a |  | F26 | 76.10 ± 0.32 | 0.40 ± 0.01c | -0.50 ± 0.05 d |
| F26-29 | 86.21 ± 3.03a | 0.26 ± 0.01a | 4.71 ± 0.32 a |  | F27 | 82.35 ± 1.04 | 0.27 ± 0.01b | 4.08 ± 0.48 b |
| F30-33 | 80.08 ± 3.06ab | 0.32 ± 0.01b | 4.07 ± 0.27 a |  | F28 | 83.09 ± 2.08 | 0.23 ± 0.01a | 1.50 ± 0.12 c |
| F34-45 | 77.48 ± 3.87b | 0.39 ± 0.00c | 1.07 ± 0.11 d |  | F29 | 83.09 ± 1.04 | 0.27 ± 0.00b | 5.38 ± 0.42 a |

**Figure 2:** Insulin secretion from pancreatic BRIN-BD11 cells for selected synthetic boarfish (*Capros aper*) protein hydrolysate-derived peptides when tested at 10-6 M. Values are expressed as mean ± SD. (n=6). \**p*< 0.05, \*\**p*< 0.01 and \*\*\**p*<0.001 compared to the 5.6 mM glucose control. Positive control: glucagon-like peptide-1 (GLP-1, 10-6 M).



**SUPPLEMENTARY DATA**

Figure S1: Reverse phase ultra-performance liquid chromatography profiles of the boarfish (*Capros aper*) protein hydrolysate (BFPH) before and after simulated gastrointestinal digestion (BFPH:SGID)



Figure S2: Mass fragmentation shows spectrum for the boarfish (*Capros aper*) derived peptide Leu/Ile-Pro-Val-Asp-Met (*m*/*z* 574.2983). The *x*-axis represents the *m*/*z* at which the precursor (pre) and fragment (*a*, *b*, *y*) ions were detected. The *y*-axis



**Relative Intensity (%)**

**m/z**

Figure S3: Mass fragmentation spectrum for the boarfish (*Capros aper*) derived peptide Pro-Ala-Val-Pro (*m*/*z* 383.2328). The *x*-axis represents the *m*/*z* at which the precursor (pre) and fragment (*a*, *b*, *y*) ions were detected. The *y*-axis shows the relative intensity of ions.



**Relative Intensity (%)**

**m/z**

Figure S4: Mass fragmentation spectrum for the boarfish (*Capros aper*) derived peptide Ala-Pro-Leu/Ile-Thr (*m*/*z* 383.2328). The *x*-axis represents the *m*/*z* at which the precursor (pre) and fragment (*a*, *b*, *y*) ions were detected. The *y*-axis shows the relative intensity of ions.



**Relative Intensity (%)**

**m/z**

**Table S1:** Sequences of the peptides identified by liquid chromatography-mass spectrometry/mass spectrometry in boarfish (*Capros aper*) protein hydrolysate fraction 28 (F28) using the *de novo* sequencing function of PEAKS software.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Peptide sequence\*** | **ALC (%)** | **Mass (m/z)** | **z** | **Retention time (min)** | **Error (ppm)** | **Homologous Protein: Species** |
| LELP | 96 | 471.2733 | 1 | 14.57 | -17.1 | numerous |
| LEIP |  |  |  |  |  | numerous |
| IELP |  |  |  |  |  | numerous |
| IEIP |  |  |  |  |  | numerous |
| VGGM | 94 | 363.1657 | 1 | 6.48 | -10.8 | numerous |
| AGLT | 94 | 361.2066 | 1 | 5.58 | -4.2 | numerous |
| AGIT |  |  |  |  |  | numerous |
| EAPPHL | 94 | 332.1736 | 2 | 9.46 | -9.2 | - |
| EAPPHI |  |  |  |  |  | myosin heavy chain, fast skeletal muscle: *Cyprinus carpio*  |
| VGVT | 94 | 375.2198 | 1 | 5.99 | -10.7 | numerous |
| DLAPK | 94 | 272.1596 | 2 | 5.47 | -3.1 | - |
| DIAPK |  |  |  |  |  | kynurenine formamidase: *Salmo salar* |
| DPTNLK | 93 | 344.1844 | 2 | 5.98 | -8.1 | - |
| DPTNIK |  |  |  |  |  | - |
| NSLM | 93 | 464.2145 | 1 | 9.26 | -6.2 | numerous |
| NSIM |  |  |  |  |  | actin, cytoplasmic 1: numerous species |
| WYPR | 93 | 311.1507 | 2 | 5.5 | -32.6 |  - |
| VDLK | 93 | 237.6465 | 2 | 5.96 | -13.5 | numerous |
| VDIK |  |  |  |  |  | numerous |
| LTLLP | 92 | 556.3253 | 1 | 15.09 | -81.2 | G-protein coupled receptor: *Danio rerio* / Homeobox protein SEBOX: *Danio rerio*  |
| ITIIP |  |  |  |  |  | - |
| LTILP |  |  |  |  |  | numerous |
| LTLIP |  |  |  |  |  | Cytoplasmic tRNA 2-thiolation protein 2: *Danio rerio*  |
| ITLIP |  |  |  |  |  | Leucine-rich repeat-containing G-protein coupled receptor 6: *Danio rerio* /Protein Jumonji: *Danio rerio* |
| ITILP |  |  |  |  |  | G-protein coupled receptor 98: *Danio rerio* |
| ITLLP |  |  |  |  |  | - |
| ITLLP |  |  |  |  |  | - |
| LTLLP |  |  |  |  |  | G-protein coupled receptor 98: *Danio rerio*/ homeobox protein SEBOX: *Danio rerio* |
| LDLL | 92 | 473.2923 | 1 | 16.07 | -9.8 | numerous |
| IDII |  |  |  |  |  | numerous |
| LDIL |  |  |  |  |  | numerous |
| LDLI |  |  |  |  |  | numerous |
| IDLI |  |  |  |  |  | numerous |
| IDIL |  |  |  |  |  | numerous |
| IDLL |  |  |  |  |  | numerous |
| IDLL |  |  |  |  |  | numerous |
| EAPPLH | 91 | 332.1935 | 2 | 9.35 | 51.0 | - |
| EAPPIH |  |  |  |  |  | - |
| ASLP | 91 | 387.2196 | 1 | 8.97 | -10.9 | numerous |
| ASIP |  |  |  |  |  | numerous |
| ANPL | 91 | 414.2303 | 1 | 8.46 | -10.6 | numerous |
| ANPI |  |  |  |  |  | numerous |
| NQLP | 91 | 471.2747 | 1 | 14.92 | 39.5 | numerous |
| NQIP |  |  |  |  |  | numerous |
| GSLP | 90 | 373.2065 | 1 | 8.64 | -4.5 | numerous |
| GSIP |  |  |  |  |  | numerous |
| KDVLA | 90 | 273.1659 | 2 | 6.63 | -8.9 | numerous |
| KDVIA |  |  |  |  |  | multidrug and toxin extrusion protein 1: *Danio rerio*/dihydropyrimidine dehydrogenase [NADP(+)]: *Danio rerio* |
| MNTPGVP | 89 | 358.1877 | 2 | 7.8 | 33.3 | - |
| DAGPGPPAA | 89 | 376.7006 | 2 | 6.55 | 48.7 | - |
| APLT | 89 | 401.235 | 1 | 6.05 | -11.2 | numerous |
| APIT |  |  |  |  |  | numerous |
| RDNPGHP | 89 | 396.6825 | 2 | 6.23 | -21.4 | - |
| LEKML | 88 | 317.1674 | 2 | 5.25 | -57.7 | numerous |
| IEKML |  |  |  |  |  | coatomer subunit beta: *Danio rerio*/ hypoxia-inducible factor 1-alpha inhibitor: *Danio rerio* |
| LEKMI |  |  |  |  |  |  - |
| IEKMI |  |  |  |  |  |  - |
| VGSL | 88 | 375.2209 | 1 | 7.92 | -7.8 | numerous |
| VGSI |  |  |  |  |  | numerous |
| AGGF | 86 | 351.1627 | 1 | 7.53 | -10.1 | numerous |
| DPTKAR | 86 | 344.1836 | 2 | 6.1 | -27.0 | - |
| TVPPLH | 86 | 332.175 | 2 | 9.38 | -60.0 | - |
| TVPPIH |  |  |  |  |  | - |
| VPTP | 86 | 413.2347 | 1 | 8.08 | -11.5 | numerous |
| EAPFSL | 86 | 332.1767 | 2 | 9.72 | 17.1 | - |
| EAPFSI |  |  |  |  |  | - |
| NAGKQP | 86 | 307.6763 | 2 | 8.47 | 32.0 | - |
| VDYKN | 85 | 319.6578 | 2 | 4.64 | -9.5 |  - |
| SLGP | 85 | 373.2069 | 1 | 8.17 | -3.4 | numerous |
| SIGP |  |  |  |  |  | numerous |
| DYSPM | 85 | 306.6379 | 2 | 6.45 | 57.4 |  - |
| FVSQV | 85 | 290.1575 | 2 | 7.94 | -10.3 | numerous |
| LANDPF | 85 | 338.6689 | 2 | 14.91 | 0.6 | - |
| IANDPF |  |  |  |  |  | - |
| LVLD | 84 | 459.2757 | 1 | 11.4 | -12.3 | numerous |
| LVID |  |  |  |  |  | numerous |
| IVID |  |  |  |  |  | numerous |
| IVLD |  |  |  |  |  | numerous |
| NNLVP | 84 | 556.3298 | 1 | 15.13 | 37.7 |  - |
| NNIVP |  |  |  |  |  |  - |
| LSGP | 83 | 373.2057 | 1 | 8.24 | -6.5 | numerous |
| ISGP |  |  |  |  |  | numerous |
| EGYPR | 83 | 311.1491 | 2 | 5.46 | -13.2 |  - |
| AVSP | 83 | 373.205 | 1 | 7.44 | -8.4 | numerous |
| DNLAPM(+15.99) | 83 | 338.6661 | 2 | 14.8 | 41.3 | - |
| DNIAPM(+15.99) |  |  |  |  |  | - |
| DGYPR | 83 | 304.1445 | 2 | 5.42 | -2.9 | adenylate kinase isoenzyme 1: *Cyprinus carpio* |
| LSAP | 83 | 387.219 | 1 | 9.05 | -12.4 | numerous |
| ISAP |  |  |  |  |  | numerous |
| M(+15.99)NCCPF | 83 | 365.6281 | 2 | 15.13 | 18.0 | cornifelin homolog: *Danio rerio* |
| SGLP | 83 | 373.2023 | 1 | 8.71 | -15.7 | numerous |
| SGIP |  |  |  |  |  | numerous |
| VAAGQ | 82 | 445.2364 | 1 | 5.22 | -9.3 |  - |
| ILGSPT | 82 | 294.1599 | 2 | 5.21 | -46.5 | pigment-dispersing hormone: *Faxonius limosus*  |
| LLGSPT |  |  |  |  |  | - |
| IIGSPT |  |  |  |  |  | - |
| LIGSPT |  |  |  |  |  | tRNA modification GTPase GTPBP3, mitochondrial: *Danio rerio*  |
| AAPF | 82 | 405.2118 | 1 | 13.59 | -3.4 | numerous |
| DSRPAP | 82 | 321.6737 | 2 | 5.69 | 30.5 | - |
| MTLCDT | 82 | 342.1868 | 2 | 8.87 | 135.5 | - |
| MTICDT |  |  |  |  |  | - |
| LDAHGP | 82 | 305.1502 | 2 | 6.09 | -9.9 | estrogen receptor: *Oreochromis aureus*  |
| IDAHGP |  |  |  |  |  | - |
| ATAGGAPHE | 82 | 405.6897 | 2 | 5.29 | -2.4 | - |
| LEKPM(+15.99) | 82 | 317.1666 | 2 | 5.2 | -2.5 | transmembrane protein 201: *Danio rerio*  |
| IEKPM(+15.99) |  |  |  |  |  | myosin heavy chain, fast skeletal muscle: *Cyprinus carpio*  |
| LELM | 81 | 505.2676 | 1 | 14.01 | -2.8 | numerous |
| LEIM |  |  |  |  |  | numerous |
| IELM |  |  |  |  |  | numerous |
| IEIM |  |  |  |  |  | numerous |
| VGCP | 81 | 375.2187 | 1 | 8.03 | 131.1 | acetylcholinesterase: *Electrophorus electricus*  |
| M(+15.99)DLP | 81 | 491.2467 | 1 | 15.89 | 60.5 | numerous |
| M(+15.99)DIP |  |  |  |  |  | numerous |
| LGAGQ | 81 | 445.2359 | 1 | 5.25 | -10.3 | numerous |
| IGAGQ |  |  |  |  |  |  - |
| LLM(+15.99)S | 81 | 479.2595 | 1 | 8.55 | 12.7 | numerous |
| IIM(+15.99)S |  |  |  |  |  | numerous |
| LIM(+15.99)S |  |  |  |  |  | numerous |
| ILM(+15.99)S |  |  |  |  |  | numerous |
| QF | 81 | 294.1429 | 1 | 7.91 | -6.7 | numerous: *Capros aper*  |
| R | 81 | 175.1172 | 1 | 0.86 | -9.9 | numerous: *Capros aper*  |
| TEAPLNPK | 81 | 435.2422 | 2 | 7.72 | 5.1 | actin, cytoplasmic 1 : *Salmo salar*  |
| TEAPINPK |  |  |  |  |  | - |
| NLGPQGPK | 81 | 405.6874 | 2 | 5.37 | -97.8 | - |
| NIGPQGPK |  |  |  |  |  | - |
| F | 80 | 166.0864 | 1 | 7.7 | 0.8 | numerous: *Capros aper*  |
| VAVEV | 80 | 516.2968 | 1 | 12.53 | -11.8 | numerous |
| NPADSKKPE | 80 | 493.2503 | 2 | 8.29 | -1.6 | - |
| LPGA | 80 | 357.2119 | 1 | 5.68 | -3.8 | numerous |
| IPGA |  |  |  |  |  | numerous |
| HLLSLT | 80 | 342.1854 | 2 | 9.04 | -66.0 | V(D)J recombination-activating protein 1: *Danio rerio* and *Oncorhynchus* *mykiss*  |
| HIISIT |  |  |  |  |  |  - |
| HLLSIT |  |  |  |  |  |  - |
| HILSIT |  |  |  |  |  |  - |
| HLISIT |  |  |  |  |  |  - |
| HIISLT |  |  |  |  |  |  - |
| HILSLT |  |  |  |  |  |  - |
| HLISLT |  |  |  |  |  |  - |

ALC: Average local confidence is an average of the individual local confidence score for each amino acid and is based on the PEAKS algorithm (Ma et al., 2003).

ppm:mass error, calculated as 106 × (experimental mass - theoretical mass) / theoretical mass. ‘-‘ : peptides did not show homology with the protein sequences of in the ‘in-house’ database which consisted of 5,300 reviewed proteins from the superclass Actinopterygii and 132 unreviewed protein sequences from the family Caproidae (boarfishes)**.** **\***Peptides displayed using one letter code for the amino acids.

**Table S2:** Sequences of the peptides identified by liquid chromatography-mass spectrometry/mass spectrometry in boarfish (*Capros aper*) protein hydrolysate fraction 29 (F29) using the *de novo* sequencing function of PEAKS software.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Peptide sequence\*** | **ALC (%)** | **Mass (m/z)** | **z** | **Retention time (min)** | **Error (ppm)** | **Homologous Protein: Species** |
| LDLL | 96 | 473.3015 | 1 | 16.15 | 9.6 | numerous |
| IDII |  |  |  |  |  | numerous |
| LDIL |  |  |  |  |  | numerous |
| LDII |  |  |  |  |  | numerous |
| LDLI |  |  |  |  |  | numerous |
| IDLI |  |  |  |  |  | numerous |
| IDIL |  |  |  |  |  | numerous |
| IDLL |  |  |  |  |  | numerous |
| VPDPR | 96 | 292.1678 | 2 | 5.97 | 14.5 | sarcoplasmic/endoplasmic reticulum calcium ATPase 1: *Makaira nigricans*/ Serine/threonine-protein kinase SIK3 homolog: *Danio rerio*  |
| VVVT | 96 | 417.2753 | 1 | 7.04 | 10.9 | numerous |
| VPTP | 94 | 413.2453 | 1 | 8.11 | 14.2 | numerous |
| VGVT | 94 | 375.2275 | 1 | 6.07 | 10.0 | numerous |
| VLVS | 94 | 417.2761 | 1 | 7.66 | 12.9 | numerous |
| VIVS |  |  |  |  |  | numerous |
| TGLP | 93 | 387.2267 | 1 | 9.06 | 7.6 | numerous |
| TGIP |  |  |  |  |  | numerous |
| ALDLL | 92 | 544.3345 | 1 | 16.83 | 0.8 | numerous |
| AIDII |  |  |  |  |  | triple functional domain protein: *Danio rerio* |
| ALDLI |  |  |  |  |  | numerous |
| ALDIL |  |  |  |  |  | numerous |
| ALDII |  |  |  |  |  | nipped-B-like protein A: *Danio rerio*/TATA box-binding protein-like protein 1: *Danio rerio* |
| AIDLL |  |  |  |  |  | ubiquitin carboxyl-terminal hydrolase 22: *Danio rerio* |
| AIDIL |  |  |  |  |  | numerous |
| AIDLI |  |  |  |  |  | numerous |
| LEAPPH | 92 | 332.1832 | 2 | 7.51 | 19.9 | - |
| IEAPPH |  |  |  |  |  | myosin heavy chain, fast skeletal muscle: *Cyprinus carpio*  |
| GHREP | 92 | 298.1542 | 2 | 2.02 | 10.8 |  - |
| VELK | 92 | 244.6606 | 2 | 5.9 | 12.4 | numerous |
| VEIK |  |  |  |  |  | numerous |
| PAVP | 92 | 383.2328 | 1 | 6.6 | 10.3 | numerous |
| TVPPHL | 92 | 332.1823 | 2 | 9.48 | -37.8 | - |
| TVPPHI |  |  |  |  |  | - |
| APLDK | 91 | 272.1653 | 2 | 5.09 | 17.6 | alcohol dehydrogenase class-3 chain L: *Gadus morhua*  |
| APIDK |  |  |  |  |  |  - |
| WYPR | 91 | 311.1566 | 2 | 5.73 | -13.7 |  - |
| LGVDK | 91 | 266.1653 | 2 | 6 | 18.4 |  - |
| IGVDK |  |  |  |  |  |  - |
| TVGALPT | 91 | 329.6837 | 2 | 6.92 | -25.7 | - |
| TVGAIPT |  |  |  |  |  | - |
| APLER | 90 | 293.1751 | 2 | 5.25 | 12.7 | filamin A-interacting protein 1-like: *Danio rerio*/tail-anchored protein insertion receptor WRB: *Danio rerio*  |
| APIER |  |  |  |  |  |  - |
| TEAPLNPK | 90 | 435.2436 | 2 | 7.72 | 8.3 | actin, gamma-enteric smooth muscle: Gallus gallus  |
| YSGL | 90 | 439.2249 | 1 | 10 | 14.2 | numerous |
| YSGI |  |  |  |  |  | numerous |
| FDLP | 90 | 491.254 | 1 | 15.68 | 8.2 | numerous |
| FDIP |  |  |  |  |  | numerous |
| AALP | 90 | 371.2334 | 1 | 9.43 | 12.2 | numerous |
| AAIP |  |  |  |  |  | numerous |
| VPDAPK | 90 | 313.6839 | 2 | 6.07 | 15.6 | - |
| LDLK | 90 | 244.6597 | 2 | 6.09 | 8.9 | numerous |
| IDLK |  |  |  |  |  | numerous |
| LDIK |  |  |  |  |  | numerous |
| IDIK |  |  |  |  |  | numerous |
| VSYPR | 90 | 311.158 | 2 | 5.77 | -43.2 |  - |
| AGLP | 90 | 357.2167 | 1 | 9.14 | 9.7 | numerous |
| AGIP |  |  |  |  |  | numerous |
| GPGL | 89 | 343.2014 | 1 | 8.67 | 11.3 | numerous |
| GPGI |  |  |  |  |  | numerous |
| TPGL | 89 | 387.2273 | 1 | 9.6 | 9.1 | numerous |
| TPGI |  |  |  |  |  | numerous |
| TTGLV | 89 | 490.2935 | 1 | 9.53 | 12.9 | - |
| TTGIV |  |  |  |  |  | numerous |
| GLAGQ | 89 | 445.2464 | 1 | 5.36 | 13.3 | numerous |
| GIAGQ |  |  |  |  |  | polycystin-2: *Oryzias latipes* |
| LWAAP | 88 | 279.1724 | 2 | 5.7 | 52.6 | - |
| IWAAP |  |  |  |  |  | - |
| SGLP | 88 | 373.2132 | 1 | 8.47 | 13.6 | numerous |
| SGIP |  |  |  |  |  | numerous |
| TPTV | 88 | 417.2749 | 1 | 7.2 | 97.3 | numerous |
| ASLP | 88 | 387.2277 | 1 | 9.15 | 10.2 | numerous |
| ASIP |  |  |  |  |  | numerous |
| GSLP | 88 | 373.21 | 1 | 9.04 | 4.9 | numerous |
| GSIP |  |  |  |  |  | numerous |
| LGGV | 88 | 345.2169 | 1 | 8.14 | 10.8 | numerous |
| IGGV |  |  |  |  |  | numerous |
| SGELGLQLK | 88 | 472.7416 | 2 | 7.22 | -69.0 | - |
| SGEIGIQIK |  |  |  |  |  | - |
| SGELGLQIK |  |  |  |  |  | - |
| SGELGIQLK |  |  |  |  |  | - |
| SGEIGLQIK |  |  |  |  |  | - |
| SGEIGIQLK |  |  |  |  |  | - |
| SGELGIQIK |  |  |  |  |  | - |
| SGELGLQLK |  |  |  |  |  | - |
| ANLP | 88 | 414.2394 | 1 | 8.7 | 11.5 | numerous |
| ANIP |  |  |  |  |  | numerous |
| GHRPE | 88 | 298.1543 | 2 | 2.1 | 11.0 | - |
| APLSTP | 88 | 293.1756 | 2 | 5.44 | 33.7 | - |
| APISTP |  |  |  |  |  | - |
| LGGP | 87 | 343.2006 | 1 | 8.3 | 8.9 | numerous |
| IGGP |  |  |  |  |  | numerous |
| MNPEPQ | 87 | 358.1965 | 2 | 7.97 | 109.0 | - |
| ELLEKDL | 87 | 430.2334 | 2 | 11.55 | -20.4 | - |
| EIIEKDI |  |  |  |  |  | - |
| ELLEKDI |  |  |  |  |  | - |
| ELIEKDL |  |  |  |  |  | - |
| EILEKDI |  |  |  |  |  | - |
| EIIEKDL |  |  |  |  |  | - |
| ELIEKDI |  |  |  |  |  | - |
| EILEKDL |  |  |  |  |  | - |
| LDFP | 86 | 491.2557 | 1 | 16.13 | 11.6 | numerous |
| IDFP |  |  |  |  |  | numerous |
| SAGIH | 86 | 242.6318 | 2 | 4.01 | 10.2 | actin, alpha skeletal muscle: numerous |
| SAGLH |  |  |  |  |  | homeobox protein Hox-A3a: *Danio rerio*  |
| EAPPLH | 86 | 332.1812 | 2 | 9.68 | 13.7 | - |
| EAPPIH |  |  |  |  |  | - |
| GGLP | 86 | 343.2028 | 1 | 8.56 | 15.1 | numerous |
| GGIP |  |  |  |  |  | numerous |
| VSVGQVP | 86 | 343.2037 | 2 | 6.89 | 17.9 | - |
| GPLN | 86 | 400.2238 | 1 | 5.4 | 12.0 | numerous |
| GPIN |  |  |  |  |  | numerous |
| LPVDM | 86 | 574.2983 | 1 | 13.34 | 13.7 | numerous |
| IPVDM |  |  |  |  |  | numerous |
| AVLQ | 85 | 430.2684 | 1 | 7.21 | 5.6 | numerous |
| AVIQ |  |  |  |  |  | numerous |
| LLKDGDK | 85 | 263.4919 | 3 | 5.13 | 12.8 | - |
| IIKDGDK |  |  |  |  |  | - |
| LIKDGDK |  |  |  |  |  | - |
| ILKDGDK |  |  |  |  |  | - |
| MPTAAP | 84 | 294.1776 | 2 | 7.73 | 105.9 | - |
| VDIK | 84 | 237.6528 | 2 | 6 | 13.1 | numerous |
| VDLK |  |  |  |  |  | numerous |
| EAPFSL | 84 | 332.1817 | 2 | 9.78 | 32.1 | - |
| EAPFSI |  |  |  |  |  | - |
| YPLS(+79.97)L | 84 | 336.6763 | 2 | 8.18 | 66.9 | guanine nucleotide-binding protein subunit beta-5a & b: *Danio rerio* |
| YPLS(+79.97)I |  |  |  |  |  | - |
| YPIS(+79.97)I |  |  |  |  |  | - |
| YPIS(+79.97)L |  |  |  |  |  | - |
| MPAVP | 84 | 257.6586 | 2 | 6.47 | 79.0 | neurexin-1a: *Danio rerio*  |
| VVVP | 84 | 413.2444 | 1 | 8.31 | -76.3 | numerous |
| LELL | 83 | 487.3111 | 1 | 16.43 | -3.2 | numerous |
| IEII |  |  |  |  |  | numerous |
| LEIL |  |  |  |  |  | numerous |
| LELI |  |  |  |  |  | numerous |
| IELI |  |  |  |  |  | numerous |
| IEIL |  |  |  |  |  | numerous |
| LEII |  |  |  |  |  | numerous |
| IELL |  |  |  |  |  | numerous |
| EAYEK | 83 | 320.175 | 2 | 6.59 | 69.5 | pre-mRNA-splicing factor SLU7: *Danio rerio*/TBC1 domain family member 8B: *Danio rerio* |
| IGGM | 83 | 377.1893 | 1 | 9.12 | 10.6 | numerous |
| LGGM |  |  |  |  |  | numerous |
| YPIIS(+79.97) | 82 | 336.6764 | 2 | 7.92 | 67.3 | numerous |
| YPLLS(+79.97) |  |  |  |  |  | - |
| YPILS(+79.97) |  |  |  |  |  | DANRE Uracil phosphoribosyltransferase homolog: *Danio rerio*  |
| YPLIS(+79.97) |  |  |  |  |  | - |
| LRDK | 82 | 266.1638 | 2 | 6.1 | -8.6 | numerous |
| IRDK |  |  |  |  |  | numerous |
| ELLMVPQ | 82 | 415.2512 | 2 | 10.86 | 55.8 | - |
| EIIMVPQ |  |  |  |  |  | - |
| ELIMVPQ |  |  |  |  |  | - |
| EILMVPQ |  |  |  |  |  | - |
| NYSDL | 82 | 306.1496 | 2 | 6.71 | 40.6 | DANRE Hyccin: *Danio rerio*  |
| NYSDI |  |  |  |  |  | - |
| GPSL | 82 | 373.2125 | 1 | 8.86 | 11.7 | numerous |
| GPSI |  |  |  |  |  | numerous |
| LM(+15.99) | 82 | 279.1418 | 1 | 3.29 | 16.2 | numerous |
| IM(+15.99) |  |  |  |  |  | numerous |
| EAPPHL | 82 | 332.1817 | 2 | 9.73 | 15.3 | - |
| EAPPHI |  |  |  |  |  | myosin heavy chain, fast skeletal muscle: *Cyprinus carpio*  |
| QQLA | 82 | 459.2613 | 1 | 6.67 | 11.2 | numerous |
| QQIA |  |  |  |  |  | numerous |
| LPVVDGW | 81 | 393.2249 | 2 | 9.62 | 29.7 | - |
| IPVVDGW |  |  |  |  |  | - |
| WVSVQGP | 81 | 386.7188 | 2 | 7.5 | 40.7 | - |
| KNAPPH | 81 | 332.1805 | 2 | 7.56 | -5.4 | - |
| TVLMTV | 81 | 332.1814 | 2 | 9.84 | -28.7 | - |
| TVIMTV |  |  |  |  |  | - |
| YPLT(+79.97)V | 81 | 336.6764 | 2 | 7.86 | 67.2 | growth/differentiation factor 8: numerous |
| YPIT(+79.97)V |  |  |  |  |  | - |
| FVEP | 80 | 491.2587 | 1 | 16.38 | 17.7 | numerous |
| VFLD | 80 | 493.2671 | 1 | 12.93 | 3.0 | numerous |
| VFID |  |  |  |  |  | numerous |
| YPLVT(+79.97) | 80 | 336.6752 | 2 | 7.97 | 63.5 | DANRE Pleckstrin homology domain-containing family H member 1: *Danio rerio* |
| YPIVT(+79.97) |  |  |  |  |  | - |
| AVLGER | 80 | 322.6761 | 2 | 6.11 | -42.9 | - |
| AVIGER |  |  |  |  |  | - |
| NYSPM | 80 | 306.1514 | 2 | 6.67 | 75.6 | - |
| QF | 80 | 294.1482 | 1 | 7.98 | 11.5 | numerous |
| DLLL | 80 | 473.3022 | 1 | 16.04 | 11.2 | numerous |
| DIII |  |  |  |  |  | numerous |
| DLIL |  |  |  |  |  | numerous |
| DLLI |  |  |  |  |  | numerous |
| DLII |  |  |  |  |  | numerous |
| DIIL |  |  |  |  |  | numerous |
| DILI |  |  |  |  |  | numerous |
| DILL |  |  |  |  |  | numerous |
| WWQVGP | 80 | 386.7183 | 2 | 7.69 | 66.9 | - |
| GALP | 80 | 357.2178 | 1 | 8.97 | 12.9 | numerous |
| GAIP |  |  |  |  |  | numerous |
| APVP | 80 | 383.2314 | 1 | 7.57 | 6.6 | numerous |
| VSVGNLP | 80 | 343.2017 | 2 | 6.95 | 12.1 | - |
| VSVGNIP |  |  |  |  |  | - |

ALC: Average local confidence is an average of the individual local confidence score for each amino acid and is based on the PEAKS algorithm (Ma et al., 2003).

ppm:mass error, calculated as 106 × (experimental mass - theoretical mass) / theoretical mass. ‘-‘ : peptides did not show homology with the protein sequences of in the ‘in-house’ database which consisted of 5,300 reviewed proteins from the superclass Actinopterygii and 132 unreviewed protein sequences from the family Caproidae (boarfishes)**.** **\***Peptides displayed using one letter code for the amino acids.