
# Identification of short peptide sequences in the nanofiltration permeate of a bioactive whey protein hydrolysate

Solène Le Maux, Alice B. Nongonierma, Brian Murray, Phil M. Kelly, Richard J. FitzGerald

PII: S0963-9969(15)30186-1
DOI: doi: 10.1016/j.foodres.2015.09.012
Reference: FRIN 6015

To appear in: *Food Research International*

Received date: 12 June 2015
Revised date: 10 September 2015
Accepted date: 13 September 2015

Please cite this article as: Le Maux, S., Nongonierma, A.B., Murray, B., Kelly, P.M. & FitzGerald, R.J., Identification of short peptide sequences in the nanofiltration permeate of a bioactive whey protein hydrolysate, *Food Research International* (2015), doi: 10.1016/j.foodres.2015.09.012

# Identification of short peptide sequences in the nanofiltration permeate of a bioactive whey protein hydrolysate

Solène Le Maux[a,‡], Alice B. Nongonierma[a,‡], Brian Murray[b], Phil M. Kelly[b] & Richard J. FitzGerald[a,*]

[a] *Department of Life Sciences and Food for Health Ireland (FHI), University of Limerick, Limerick, Ireland.*

[b] *Teagasc Food Research Centre, Moorepark and Food for Health Ireland (FHI), Fermoy, Co. Cork, Ireland*

‡ These two authors contributed equally

* Corresponding author: Dick.FitzGerald@ul.ie

Tel: +353 (0) 61 202598

Fax: + 353 (0) 61 331490

## Abstract

Short peptides in food protein hydrolysates are of significant interest as they may be highly bioactive whilst also being bioavailable. A dipeptidyl peptidase IV (DPP-IV) inhibitory whey protein hydrolysate (WPH) was fractionated using nanofiltration (NF) with a 200 Da MWCO membrane. The DPP-IV half maximal inhibitory concentration of the NF permeate ($IC_{50}$ = 0.66 ± 0.08 mg protein equivalent $mL^{-1}$) was significantly more potent ($P$ > 0.05) than that of the starting WPH ($IC_{50}$ = 0.94 ± 0.24 mg protein equivalent $mL^{-1}$) and associated retentate ($IC_{50}$ = 0.82 ± 0.13 mg protein equivalent $mL^{-1}$). This confirmed the contribution of short peptides within the NF permeate to the overall DPP-IV inhibitory activity. An hydrophilic interaction liquid chromatography (HILIC-) and reverse-phase (RP-) liquid chromatography tandem mass spectrometry (LC-MS/MS) strategy, based on two retention time models, allowed detection of eight free amino acids and eight di- to tetrapeptides in the NF permeate. The potential sequences of the peptides within the NF permeate were then ranked on the basis of their highest probability of occurrence. A confirmatory study with synthetic peptides showed that valine-alanine (VA), valine-leucine (VL), tryptophan-leucine (WL) and tryptophan-isoleucine (WI), displayed DPP-IV $IC_{50}$ values < 170 µM. The NF and LC-MS strategy employed herein represents a new approach for the targeted identification of short peptides within bioactive food protein hydrolysates.


**Key words:** bioactive peptides, mass spectrometry, short peptides, retention time, dipeptidyl peptidase IV inhibition

# 1. Introduction

There is a growing interest in the utilisation of naturally-derived food products with health benefits such as antioxidant, antidiabetic and antihypertensive properties (Korhonen & Pilhanto, 2006; Li-Chan, 2015). Short peptide sequences are of particular interest as they have been associated, in certain instances, with both high bioactive potency and bioavailability (Segura-Campos, Chel-Guerrero, Betancur-Ancona, & Hernandez-Escalante, 2011). This may be due to the fact that specific short peptides may be stable to gastrointestinal digestion and be absorbed intact in the intestinal epithelium (Foltz et al., 2007).

Enzymatic hydrolysis of milk proteins is a strategy frequently used for the generation of bioactive peptides. However, milk protein hydrolysates can contain a complex mixture of peptides. Therefore, fractionation of milk protein hydrolysates is employed to: (i) generate fractions enriched in bioactive peptides and (ii) reduce their compositional complexity, which may help in subsequent peptide identification.

To date, the most frequently used techniques to separate peptides have been based on membrane processing and/or chromatographic separation (Bazinet & Firdaous, 2013; Poliwoda & Wieczorek, 2009). Nanofiltration (NF) has been utilised to enrich milk protein hydrolysates in bioactive peptides. It has been shown that both the molecular mass and peptide charge affect peptide transmission through NF membranes (Pouliot, 2008; Pouliot, Gauthier, & L'Heureux, 2000).

Food protein-derived bioactive peptides can differ in length (generally < 6 amino acids) and physicochemical properties (Panchaud, Affolter, & Kussmann, 2012). For example, in the case of dietary antidiabetic peptides, several di- and tripeptides have been shown to be relatively potent *in vitro* (Lacroix & Li-Chan, 2012; Nongonierma & FitzGerald, 2014). Liquid chromatography-tandem mass spectrometry (LC-MS/MS) is a strategy

commonly used to identify specific peptide sequences within food protein hydrolysates. However, LC-MS/MS detection of short peptides within food protein hydrolysates still presents several technical challenges. This is linked to the fact that a number of different peptide sequences can correspond to the same molecular mass. Therefore, the application of different separation strategies prior to MS detection can provide additional information to allow for short peptide identification. This is the case for chromatographic separation as peptide retention on a chromatographic matrix depends both on the sequence and the chromatographic conditions (Zou, Zhang, Hong, & Lu, 1992). The retention time of short peptides can be predicted using algorithm models based on properties such as peptide size and amino acid position within the peptide sequence (i.e., at the N-, the C-terminus or within the peptide sequence) (Krokhin, 2006; Le Maux, Nongonierma, & FitzGerald, 2015; Meek, 1980; Tripet et al., 2007). Furthermore, the combination of different separation modes such as hydrophilic interaction liquid chromatography (HILIC) and reverse-phase (RP) chromatography has been shown to further enhance the accuracy of short peptide identification (Harscoat-Schiavo et al., 2012).

The aim of this study was to validate a strategy involving NF, LC-MS/MS and retention time prediction for the selective enrichment and specific identification of short peptides in a bioactive protein hydrolysate. To support this approach, an antidiabetic *in vitro* bioassay measuring DPP-IV inhibition was employed. Dipeptidyl peptidase IV (DPP-IV) is a metabolic enzyme which has been identified in the degradation of incretins (glucose dependent insulinotropic polypeptide (GIP) and glucagon-like peptide-1 (GLP-1)), resulting in a loss in their insulinotropic properties *in vivo* (Juillerat-Jeanneret, 2014). To date, di- and tripeptide sequences have mainly been identified using *in silico* digestion followed by confirmatory studies with synthetic peptides along with peptide library approaches (Hikida, Ito, Motoyama, Kato, & Kawarasaki, 2013; Lan et al., 2015; Nongonierma & FitzGerald,

2013a; Tulipano, Sibilia, Caroli, & Cocchi, 2011). However, to our knowledge, no di- and tripeptides with DPP-IV inhibitory properties have to date been directly identified within milk protein hydrolysates using LC-MS.

This study was carried out using a whey protein hydrolysate (WPH) with DPP-IV inhibitory activity. NF of the WPH was used to enrich for short peptides in the permeate. Milk protein-derived short peptide sequences were detected in the NF permeate and their overall contribution to DPP-IV inhibitory properties was assessed using a synthetic peptide confirmatory study.

## 2. Materials and methods

### *2.1.Reagents*

Trifluoroacetic acid (TFA), tris(hydroxymethyl)aminomethane (TRIS), Gly-Pro-pNA, diprotin A (IPI), porcine DPP-IV (≥ 10 units $mg^{-1}$ protein), high pressure liquid chromatography (HPLC) and MS grade water and acetonitrile (MeCN) were obtained from Sigma Aldrich (Dublin, Ireland). Hydrochloric acid (HCl) and sodium hydroxide (NaOH) were from VWR (Dublin, Ireland). The synthetic peptides cysteine-isoleucine-valine-leucine (CIVL) and leucine-cysteine-valine-leucine (LCVL) were from Genscript (Piscataway, NJ, USA) while glycine-isoleucine (GI), leucine-glycine (LG), serine-valine (SV), tyrosine-isoleucine (YI), tyrosine-leucine (YL), isoleucine-tyrosine (IY) and isoleucine-glycine (IG) were from Bachem (Bubendorf, Germany). Standard peptides (purity ≥ 95% (*w/w)*) used for the retention time prediction models were purchased from Bachem, GenScript or Thermo Fisher Scientific (Waltham, MA, USA) as per Le Maux et al. (2015). The whey protein substrate (88.3% (*w/w*) protein) was obtained from Carbery Food Ingredients (Ballineen,

Ireland). All other chemicals were obtained from Sigma Aldrich and were of analytical grade unless otherwise stated.

*2.2.Enzymatic hydrolysis of whey proteins and NF fractionation of the hydrolysate*

The WPH was generated using a food-grade pancreatic proteinase preparation at semi-pilot scale (200 L) as described by Nongonierma and FitzGerald (2013b). After 4 h hydrolysis, heat inactivation of the enzyme preparation was achieved by continuous heating to 85°C with 25 s holding time in a Unison H17 plate heat exchanger (Unison Engineering Services Ltd., Limerick, Ireland). The hydrolysate was then nanofiltered using a membrane housing containing a 200 Da molecular weight cut off (MWCO) spiral wound Synder Filtration NF membranes, NFX-2A-3838 with 31 mil spacer, at 50°C and 22 bar inlet operating pressure. NF was conducted until 10.5% (*w/v*) solids were reached in the retentate. Different samples were collected throughout the process (hydrolysate, 200 Da permeate and 200 Da retentate). These were freeze-dried (FreeZone 18L, Labconco, Kansas City, U.S.A.) and stored at -20°C until utilization.

*2.3.Determination of total nitrogen*

Total nitrogen content of the WPH, NF permeate and retentate was determined in triplicate (n=3) by Kjeldahl analysis as previously described by Connolly, Piggott, and FitzGerald (2013), with a KjelFlex K-360 (BUCHI Labortechnik AG, Flawil, Switzerland). A conversion factor of 6.38 was used to calculate the protein equivalent in the different samples (Cerbulis, Woychik, & Wondolowski, 1972).

*2.4.DPP-IV inhibition assay*

The protein hydrolysates were dispersed in HPLC-grade water at concentrations ranging from $31.3 \times 10^{-3}$ to 5.0 mg $mL^{-1}$ (final concentration). The DPP-IV inhibition assay was carried out as described by Nongonierma and FitzGerald (2013a). Each sample was analysed in triplicate. The DPP-IV half maximal inhibitory concentration ($IC_{50}$) values were determined by plotting the percentage inhibition as a function of the concentration of test compound.

*2.5.Reverse-phase ultra-performance liquid chromatography (RP-UPLC) of the hydrolysates and associated NF fractions*

The peptide profiles of the different samples were determined by RP-UPLC (Acquity - Waters, Dublin, Ireland) equipped with a 2.1 × 50 mm, 1.7 µm Acquity UPLC BEH C18 column and an Acquity BEH C18 1.7 µm vanguard pre-column (Waters) as described by Nongonierma and FitzGerald (2012). For each sample, 50 µg (final concentration: 5 g $L^{-1}$, injection volume: 10 µL) was injected onto the column.

*2.6.Peptide detection with liquid chromatography-mass spectrometry (LC-MS/MS)*

Samples were analysed by LC-MS/MS using a UPLC system (Waters) coupled to a quadrupole time-of-flight mass spectrometer (Impact HD™, Bruker Daltonics GmbH, Bremen, Germany). The MS was equipped with an electrospray ionisation (ESI) source used in positive ion mode. Instrument control and data acquisition were performed using Hystar™ software (Bruker Daltonics).

Two complementary chromatographic modes were used for peptide separation prior to MS/MS analysis. RP and HILIC separations were used as compounds elute differently on these two matrices. RP-UPLC was performed using an Acquity BEH C18 column (2.1 x 150

mm, 1.7 µm) equipped with an Acquity BEH C18 1.7 µm vanguard pre-column (Waters). Mobile phase A was 0.1% (*v*/*v*) TFA in water whereas mobile phase B was 0.1% (*v*/*v*) TFA in MeCN. A linear gradient from 100 to 60% solvent A was applied for 120 min at a flow rate of 0.2 mL $min^{-1}$. The column temperature was maintained at 40°C. Samples were diluted in solvent A (final concentration of 5 g $L^{-1}$) and filtered through 0.2 µM cellulose acetate filters before injection (2 µL). The UPLC system was equipped with a tunable UV detector set at 214 and 280 nm. The operating conditions were initially investigated to obtain the optimal separation of a complex milk protein hydrolysate sample. MS measurements were performed over a 70-700 *m/z* acquisition range. MS data were processed on Compass DataAnalysis 4.0 SP5 (Bruker Daltonics).

HILIC chromatography was performed on an Acquity BEH amide column as previously described by Le Maux et al. (2015). The MS operating conditions were similar to those described above for RP-LC-MS/MS.

### *2.7.Retention time prediction models and identification of short peptides by LC-MS*

Retention time prediction models were coupled to MS/MS data in order to improve peptide identification. The HILIC retention time model used was as described by Le Maux et al. (2015). Another model based on RP separation was developed herein using a similar strategy. Briefly, a training set of 153 standard peptides was used to determine the algorithm of the RP prediction model. These peptides were selected based on their range of hydrophilicity/hydrophobicity, peptide size (di- to tetrapeptides) and sequence. Matlab (version 2014b, The MathWorks Inc., Natick, MA, USA) was used to generate coefficients representing the impact of each amino acid on peptide apparent hydrophobicity. These hydrophobicity coefficients were determined depending on amino acid location (C-, N-terminus or within the peptide sequence). The apparent hydrophobicity of a peptide (*H*) was

determined as the sum of each amino acid coefficient in the peptide chain (Meek, 1980). Subsequently, a linear regression was performed to predict retention time using both peptide sequence and length. The RP retention time model was statistically validated as previously described (Le Maux et al., 2015).

The approach to identify short peptides within the nanofiltrate was also as described by Le Maux et al. (2015). A list of potential peptides, which corresponded to the properties of the detected peptides, was generated using a mass tolerance set at 0.1 Da. An in-house bovine milk protein database was built in order to discard non relevant peptides using all the available genetic variants, given in PubMed, of the major bovine milk proteins (β-lactoglobulin, α-lactalbumin, bovine serum albumin, lactoferrin, $\alpha_{s1}$-, $\alpha_{s2}$-, β- and κ-casein). As several peptide sequences could correspond to the retention time and molecular mass of one peptide signal, these potential peptides were ranked as a function of their retention time differences (observed minus predicted retention times).

### *2.8.Statistical analysis*

Means comparison was carried out using a one way ANOVA. Post-hoc tests were conducted following a Student Newman-Keuls test using SPSS (version 22, SPSS Inc., Chicago, IL, USA) at a significance level $P < 0.05$.

## 3. Results

### *3.1.Peptide profile of the WPH and associated NF fractions*

The RP-UPLC profiles for the WPH and its associated NF retentate and permeate are depicted in Fig. 1. The WPH and its NF retentate had similar peptide profiles showing a large

number of peptide peaks eluting during the first 25 min of the MeCN gradient. In contrast, the NF permeate shows a less complex profile, with 3 main peaks eluting at 2.9, 4.8 and 7.7 min. Similar peaks were also present in the WPH and the NF retentate, but at a lower intensity in the WPH than in the NF permeate.

### *3.2.DPP-IV inhibitory activity of the WPH and associated NF fractions*

The DPP-IV $IC_{50}$ values were 0.94 ± 0.24, 0.82 ± 0.13 and 0.66 ± 0.08 mg protein equivalent $mL^{-1}$ for the WPH, NF retentate and permeate, respectively (Table 1). The DPP-IV $IC_{50}$ for the NF permeate was significantly lower ($P < 0.05$) than that of the WPH and NF retentate.

The NF permeate peptide profile showed a significantly reduced complexity (i.e., three main peptide peaks, Fig. 1). In addition, it displayed the highest DPP-IV inhibitory potency (Table 1). Therefore, LC-MS/MS analysis was applied to the NF permeate with a view to specifically identify short peptides contributing to the overall DPP-IV inhibitory properties of the WPH.

### *3.3.Development of the reverse phase (RP) retention time prediction ($RT_{pred}$) model*

The retention time prediction ($RT_{pred}$) model was best described by the equation:

$$RT_{pred} = \frac{(H-a)}{b} \qquad \text{(Equation 1)}$$

The constants *a* and *b* were optimised through iterations of the amino acid hydrophobic coefficients and were defined in the operating conditions as - 1.384 ± 0.194 and 0.9383 ± 0.006, respectively. The established amino acid hydrophobic coefficients that allowed equation 1 to have the highest $R^2$ are described in Supplementary Table S1. The peptide apparent hydrophobicity (*H*) was calculated using these coefficients. Tryptophan (W)

had the highest hydrophobic coefficients, followed by phenylalanine (F) > leucine (L) and isoleucine (I) > tyrosine (Y), methionine (M) and valine (V). The basic and polar with uncharged side group amino acids exerted a very low influence on the retention time. However, depending on their position in the peptide sequence, some of these amino acids displayed a large difference between their hydrophobic coefficients. For instance, the hydrophobic coefficients for threonine (T) were low but displayed large variations (from -5.387 to 5.563 in C- and N-terminus, respectively) compared to W coefficients which were high but varied within a narrow range (from 25.659 to 28.989 in N- and C-terminus, respectively). The position of some amino acids in the peptide sequence had an impact on the retention time. Indeed, twenty-five pairs of homologous peptides (peptides with the same amino acid composition but in a different sequence) were analysed, nineteen of these peptides had significantly ($P < 0.05$) different retention time (data not shown). For example, the dipeptides WD and DW had retention times of 22.076 ± 0.077 and 33.477 ± 0.019 min, respectively. There were no peptides eluting before 2.1 min, which represented the void volume of the column.

The RP retention time model was statistically validated, having an $R^2$ of 0.978, a F-statistic of $2.23 \times 10^4$ with an extremely low *P*-value ($P < 5 \times 10^{-324}$) and a Durbin–Watson statistic value of 1.762. The Cook's distance and the residual plots of this prediction model were considered acceptable as no outliers could be determined (data not shown). The root mean squared error of the model, as well as the root mean squared errors of two cross-validations, leave-one-out and tenfold cross-validations, revealed the robustness of the model as the values were of 2.724, 2.992 and 2.995, respectively. The predicted retention time had intervals of 6.8 and 8.4 min for confidence levels of 95 and 99%, respectively (Fig. 2). Therefore, a 10 min difference interval between the observed and predicted retention times was used as a cut off to select the possible peptide candidates.

*3.4.Identification of peptides within the NF permeate of the WPH*

A relatively low number (16) of compounds could be detected by LC-MS/MS within the NF permeate (Table 2). For the first time, the peptides of this NF permeate sample were subsequently identified. These compounds were analysed by correlating their molecular mass, retention time and MS/MS spectrum. As several sequences could match the properties of one compound, these sequences were defined as potential peptides within the NF permeate. The presence of potential peptide sequences was searched against an in-house bovine milk protein database in order to reject non-relevant peptides. Eight amino acids, seven dipeptides and one tetrapeptide were detected within the NF permeate. The potential peptide sequences were ranked based on their highest probability of occurrence (Table 2). This was achieved by using the two retention time models (RP and HILIC) to allow for a more accurate ranking of homologous peptides. For instance, the molecular mass of peptide No. 10 in Table 2 corresponded to 6 (GI, GL, IG, LG, AV and VA) potential peptide sequences, but only 5 (GI, GL, IG, LG and AV) of these had predicted retention times that were within 10 min of the observed peptide retention time. Moreover, these potential peptides were ranked based on their retention time differences (predicted versus observed retention times), with GI and GL having the highest probability, followed by IG and LG, and then AV. VA was excluded as a potential peptide candidate for compound No. 10 as the difference between its predicted and observed retention time in RP conditions was > 10 min. Y, F and W showed the highest intensities by UV detection at 214 nm in HILIC-MS/MS and RP-MS/MS, which is in accordance with the peptide profile in Fig. 1 where these aromatic amino acids eluted at 2.9, 4.8 and 7.7 min, respectively.

*3.5.Confirmatory study of DPP-IV inhibitory activity with synthetic peptides*

Several of the peptide sequences identified within the NF permeate have previously been identified in the literature (Table 2). Therefore, only the peptide sequences which had not previously been reported for their DPP-IV inhibitory properties were synthesised for further confirmatory study. It was conducted to determine which compounds within the NF permeate were bioactive. Previously published DPP-IV $IC_{50}$ values for the amino acids which were found within the NF permeate are shown in Table 3 (Nongonierma, Mooney, Shields, & FitzGerald, 2013). In addition, IW, LW, WL, WI VA, GL, FW, WF and VL have also been previously evaluated for their DPP-IV inhibitory potential (Lan, Ito, Ito, & Kawarasaki, 2014; Lan et al., 2015; Nongonierma & FitzGerald, 2013a, 2013d; Tulipano et al., 2011). Three peptides reported for the first time (CIVL, LCVL and GI) were evaluated for their *in vitro* DPP-IV inhibitory properties in this study (Table 2). Their DPP-IV $IC_{50}$ values are reported in Table 3. WL was the most potent DPP-IV inhibitory peptide identified within the NF permeate, with an $IC_{50}$ value of 43.6 µM. The NF permeate contained four potential DPP-IV inhibitory peptides, VL, VA, WL and WI with $IC_{50}$ values < 170 µM.

# 4. Discussion

The fractionation of complex peptide mixtures followed by the identification of bioactive peptides is of major interest as it allows further optimisation of the protein hydrolysis processes, with the view of increasing the potency of bioactive ingredients. The NF and LC-MS/MS approach describe herein allowed for the development of an integrated strategy aimed at the identification of potent short DPP-IV inhibitory peptides within a WPH. Previous studies have identified short milk protein-derived peptides within milk protein

hydrolysate fractions (Lacroix & Li-Chan, 2014; Silveira, Martínez-Maqueda, Recio, & Hernández-Ledesma, 2013). However, the peptides reported were ≥ 4 amino acid residues in length. To our knowledge, it is the first time that di and tripeptides are identified using LC-MS/MS within a milk protein hydrolysate fraction.

Identification of short peptide sequences in complex food protein hydrolysates is challenging. Therefore, a new RP retention time prediction model focused on short peptide (up to four amino acids in length) separation was developed. It was based on the amino acid location within a peptide and its impact on peptide apparent hydrophobicity. Previous studies have also highlighted the impact of N- and C-terminal amino acids on the retention time of peptides in RP conditions (Krokhin, 2006; Tripet et al., 2007). These studies showed a similar trend to the model herein concerning the influence of each residue on peptide hydrophobicity with W, F, L and I having the highest impact. Interestingly, the elution order of peptides under RP and HILIC conditions were not the mirror image of each other as the contribution of each amino acid to the RP retention time was not the exact opposite of their contribution in HILIC conditions (Le Maux et al., 2015). Therefore, the differences between the two separation modes showed that HILIC- and RP-LC were complementary peptide separation methods. The use of complementary separation methods has previously been shown to significantly enhance peptide identification (Harscoat-Schiavo et al., 2012). Peptide identification was improved herein by focusing on di- to tetrapeptides and discriminating on the basis of the amino acid position within the peptide chain. This peptide identification strategy permitted the compilation of a short ranked list of potential sequences for the unknown peptides detected in the NF permeate (Table 2).

Previous fractionation studies were performed on the WPH fractionation using a 2 kDa ultrafiltration membrane and an hydrophobic solid-phase extraction (SPE) resin (Nongonierma & FitzGerald, 2013b). When the SPE fraction was analysed by LC-MS/MS in

the previous study, 45 potential peptides were identified. However, this number of potential peptide candidates was too high to distinguish which peptide sequence(s) may be responsible for the DPP-IV inhibitory properties (Le Maux et al., 2015). The NF fractionation approach described herein allowed the generation of a fraction with a reduced complexity (16 compounds) and a significant higher DPP-IV inhibitory activity than the WPH. While the $IC_{50}$'s of the different samples were of the same order, this approach demonstrated that short peptides were responsible in part for the bioactivity of the WPH. This NF permeate fraction was characterised in the present study for the first time. The masses of the compounds detected in the NF permeate were compatible with the nominal MWCO of the membrane as it contained free amino acids and short peptides (2-4 amino acid in length). As the permeation of compounds through the membrane is dependant of the NF conditions (i.e., pressure, temperature, etc.) as well as the peptide conformation, it may explain the detection of a 447.295 Da peptide in the permeate. As this study focused on short peptides, the NF permeate was the only sample which was further characterised by LC-MS/MS. Moreover, the NF permeate was the most potent fraction and displayed a reduced peptide profile complexity.

The DPP-IV inhibitory properties of three amino acids found within the NF permeate (L, M and W) had previously been demonstrated. However, their DPP-IV inhibitory potency was quite low ($IC_{50}$ > 2300 µM (Nongonierma et al., 2013)). The most potent peptides ($IC_{50}$ < 170 µM) found within the NF permeate of WPH were WI, WL, VA and VL. This result was in agreement with a previous *in silico* study using a peptide alignment strategy, which showed that peptides possessing a W at the N-terminus and/or a P at position 2 generally had DPP-IV $IC_{50}$'s < 200 µM (Nongonierma & FitzGerald, 2014). Interestingly, WL ($IC_{50}$ = 43.6 µM), the most potent peptide within the NF permeate sample was of the same order of DPP-IV inhibitory potency as LKPTPEGDL ($IC_{50}$ = 45 µM) and IPAVF ($IC_{50}$ = 44.7 µM), which had previously been identified within whey protein hydrolysates (Lacroix & Li-Chan, 2014;

Silveira et al., 2013). It was shown in previous studies that specific milk protein-derived peptides with a DPP-IV $IC_{50}$ value of the same order as the most potent peptides identified within the NF permeate displayed an antidiabetic activity *in vivo*. This was the case with the DPP-IV inhibitory peptides VAGTWY (β-lactoglobulin f15-20; $IC_{50}$ = 174 μM) and LPQDIPPL (β-casein f70–77; $IC_{50}$ = 46 μM) which mediated insulinotropic and serum glucose lowering activities in small animals (Uchida, Ohshiba, & Mogami, 2011; Uenishi, Kabuki, Seto, Serizawa, & Nakajima, 2012). The results presented herein need to be assessed *in vivo* in order to understand how they may affect serum glucose regulation.

The interest in focusing on short (di- tri and tetra-) peptides was related to the fact that these sequences may be bioavailable as they may be able to survive gastrointestinal digestion and also display an increased intestinal permeability compared to larger peptides or free amino acids. Permeation of short peptides (di- and tripeptides) through Caco-2 cell monolayers has been reported in *in vitro* studies (Shimizu, Tsunogai, & Arai, 1997). Human studies have also shown that short peptides have the ability to cross the gut barrier as they were identified in the serum following the ingestion of a yoghurt enriched in the lactotripeptides, LPP and VPP (Foltz et al., 2007), or milk protein hydrolysates (Morifuji et al., 2010). *In silico* digestion of the most potent peptides (VL, VA, WL and WI) indicated that VA and VL may be stable to digestion with gastrointestinal enzymes as described in Nongonierma and FitzGerald (2013c) (data not shown). However, confirmatory *in vitro* and *in vivo* studies are required to support these results.

# 5. Conclusion

A new strategy to identify short peptides was developed, being innovative in the utilisation of a RP and HILIC approach designed for di- to tetrapeptides. This new approach allowed the differentiation between homologous peptides. This method was optimised to identify short peptides which cannot usually be identified by MS/MS and *de novo* searches. The use of further separation methods prior to MS, such as ion-exchange or capillary electrophoresis, may lead to the determination of a unique peptide sequence for each compound detected.

To our knowledge, NF has not been previously employed during the fractionation and further identification of DPP-IV inhibitory peptides from complex food protein hydrolysate samples. Specific short peptide sequences found within the NF permeate of the WPH were shown to be relatively potent DPP-IV inhibitory peptides ($IC_{50}$ < 170 µM). This demonstrated that short peptides were also responsible for the overall DPP-IV inhibitory potential of the WPH. Furthermore, these peptides were predicted *in silico* to be relatively stable to gastrointestinal digestion. However, these results need to be confirmed following *in vitro* and *in vivo* digestion of the fraction.

The results reported in this study demonstrated that, with the strategy employed, it was possible to improve the fractionation and subsequently the detection of short DPP-IV inhibitory peptides in a milk protein hydrolysate. The strategy described herein is highly relevant to the discovery of short bioactive peptide sequences.

## Acknowledgements

The work described herein was supported by Enterprise Ireland under Grant Number TC2013-0001. The research was part funded by the Science Foundation Ireland Research Infrastructure Fund. The authors express their gratitude to Paul Tobin for his help with the Kjeldahl nitrogen analysis.

## References

Bazinet, L., & Firdaous, L. (2013). Separation of bioactive peptides by membrane processes: Technologies and devices. *Recent Patents on Biotechnology, 7*, 9-27.

Cerbulis, J., Woychik, J. H., & Wondolowski, M. V. (1972). Composition of commercial wheys. *Journal of Agricultural and Food Chemistry, 20*, 1057-1059.

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 & Technology, 48*, 1670-1681.

Foltz, M., Meynen, E. E., Bianco, V., van Platerink, C., Koning, T. M. M. G., & Kloek, J. (2007). Angiotensin converting enzyme inhibitory peptides from a lactotripeptide-enriched milk beverage are absorbed intact into the circulation. *The Journal of Nutrition, 137*, 953-958.

Harscoat-Schiavo, C., Nioi, C., Ronat-Heit, E., Paris, C., Vanderesse, R., Fournier, F., & Marc, I. (2012). Hydrophilic properties as a new contribution for computer-aided identification of short peptides in complex mixtures. *Analytical and Bioanalytical Chemistry, 403*, 1939-1949.

Hikida, A., Ito, K., Motoyama, T., Kato, R., & Kawarasaki, Y. (2013). Systematic analysis of a dipeptide library for inhibitor development using human dipeptidyl peptidase IV produced by a *Saccharomyces cerevisiae* expression system. *Biochemical and biophysical research communications, 430*, 1217-1222.

Juillerat-Jeanneret, L. (2014). Dipeptidyl peptidase IV and its inhibitors: Therapeutics for type 2 diabetes and what else? *Journal of Medicinal Chemistry, 57*, 2197–2212.

Korhonen, H., & Pilhanto, A. (2006). Bioactive peptides: Production and functionality. *International Dairy Journal, 16*, 945-960.

Krokhin, O. V. (2006). Sequence-specific retention calculator. Algorithm for peptide retention prediction in ion-pair RP-HPLC: Application to 300-and 100-Å pore size C18 sorbents. *Analytical Chemistry, 78*, 7785-7795.

Lacroix, I. M. E., & Li-Chan, E. C. Y. (2012). Evaluation of the potential of dietary proteins as precursors of dipeptidyl peptidase (DPP)-IV inhibitors by an *in silico* approach. *Journal of Functional Foods, 4*, 403-422.

Lacroix, I. M. E., & Li-Chan, E. C. Y. (2014). Isolation and characterization of peptides with dipeptidyl peptidase-IV Inhibitory activity from pepsin-treated bovine whey proteins. *Peptides, 54*, 39-48.

Lan, V. T. T., Ito, K., Ito, S., & Kawarasaki, Y. (2014). Trp-Arg-Xaa tripeptides act as uncompetitive-type inhibitors of human dipeptidyl peptidase IV. *Peptides, 54*, 166-170.

Lan, V. T. T., Ito, K., Ohno, M., Motoyama, T., Ito, S., & Kawarasaki, Y. (2015). Analyzing a dipeptide library to identify human dipeptidyl peptidase IV inhibitor. *Food Chemistry, 175*, 66-73.

Le Maux, S., Nongonierma, A. B., & FitzGerald, R. J. (2015). Improved short peptide identification using HILIC-MS/MS: Retention time prediction model based on the impact of amino acid position in the peptide sequence. *Food Chemistry, 173*, 847-854.

Li-Chan, E. C. (2015). Bioactive peptides and protein hydrolysates: Research trends and challenges for application as nutraceuticals and functional food ingredients. *Current Opinion in Food Science, 1*, 28-37.

Meek, J. L. (1980). Prediction of peptide retention times in high-pressure liquid chromatography on the basis of amino acid composition. *Proceedings of the National Academy of Sciences, 77*, 1632-1636.

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*, 8788-8797.

Nongonierma, A. B., & FitzGerald, R. J. (2012). Tryptophan-containing milk protein-derived dipeptides inhibit xanthine oxidase. *Peptides, 37*, 263-272.

Nongonierma, A. B., & FitzGerald, R. J. (2013a). Dipeptidyl peptidase IV inhibitory and antioxidative properties of milk-derived dipeptides and hydrolysates. *Peptides, 39*, 157-163.

Nongonierma, A. B., & FitzGerald, R. J. (2013b). Dipeptidyl peptidase IV inhibitory properties of a whey protein hydrolysate: Influence of fractionation, stability to simulated gastrointestinal digestion and food-drug interaction. *International Dairy Journal, 32*, 33-39.

Nongonierma, A. B., & FitzGerald, R. J. (2013c). Inhibition of dipeptidyl peptidase IV (DPP-IV) by proline containing peptides. *Journal of Functional Foods, 5*, 1909-1917.

Nongonierma, A. B., & FitzGerald, R. J. (2013d). Inhibition of dipeptidyl peptidase IV (DPP-IV) by tryptophan containing dipeptides. *Food & Function, 4*, 1843-1849.

Nongonierma, A. B., & FitzGerald, R. J. (2014). An *in silico* model to predict the potential of dietary proteins as sources of dipeptidyl peptidase IV (DPP-IV) inhibitory peptides. *Food Chemistry, 165*, 489-498.

Nongonierma, A. B., Mooney, C., Shields, D. C., & FitzGerald, R. J. (2013). Inhibition of dipeptidyl peptidase IV and xanthine oxidase by amino acids and dipeptides. *Food Chemistry, 141*, 644-653.

Panchaud, A., Affolter, M., & Kussmann, M. (2012). Mass spectrometry for nutritional peptidomics: How to analyze food bioactives and their health effects. *Journal of Proteomics, 75*, 3546-3559.

Poliwoda, A., & Wieczorek, P. (2009). Sample pretreatment techniques for oligopeptide analysis from natural sources. *Analytical and Bioanalytical Chemistry, 393*, 885-897.

Pouliot, Y. (2008). Membrane processes in dairy technology-From a simple idea to worldwide panacea. *International Dairy Journal, 18*, 735-740.

Pouliot, Y., Gauthier, S. F., & L'Heureux, J. (2000). Effect of peptide distribution on the fractionation of whey protein hydrolysates by nanofiltration membranes. *Le Lait, 80*, 113-120.

Segura-Campos, M., Chel-Guerrero, L., Betancur-Ancona, D., & Hernandez-Escalante, V. M. (2011). Bioavailability of bioactive peptides. *Food Reviews International, 27*, 213-226.

Shimizu, M., Tsunogai, M., & Arai, S. (1997). Transepithelial transport of oligopeptides in the human intestinal cell, Caco-2. *Peptides, 18*, 681-687.

Silveira, S. T., Martínez-Maqueda, D., Recio, I., & Hernández-Ledesma, B. (2013). Dipeptidyl peptidase-IV inhibitory peptides generated by tryptic hydrolysis of a whey protein concentrate rich in β-lactoglobulin. *Food Chemistry, 141*, 1072-1077.

Tripet, B., Cepeniene, D., Kovacs, J. M., Mant, C. T., Krokhin, O. V., & Hodges, R. S. (2007). Requirements for prediction of peptide retention time in reversed-phase high-performance liquid chromatography: Hydrophilicity/hydrophobicity of side-chains at the N- and C-termini of peptides are dramatically affected by the end-groups and location. *Journal of Chromatography A, 1141*, 212-225.

Tulipano, G., Sibilia, V., Caroli, A. M., & Cocchi, D. (2011). Whey proteins as source of dipeptidyl dipeptidase IV (dipeptidyl peptidase-4) inhibitors. *Peptides, 32*, 835-838.

Uchida, M., Ohshiba, Y., & Mogami, O. (2011). Novel dipeptidyl peptidase-4-inhibiting peptide derived from β-lactoglobulin. *Journal of Pharmacological Sciences, 117*, 63-66

Uenishi, H., Kabuki, T., Seto, Y., Serizawa, A., & Nakajima, H. (2012). Isolation and identification of casein-derived dipeptidyl-peptidase 4 (DPP-4)-inhibitory peptide LPQNIPPL from gouda-type cheese and its effect on plasma glucose in rats. *International Dairy Journal, 22*, 24-30.

Zou, H., Zhang, Y., Hong, M., & Lu, P. (1992). Effect of column temperature on retention of dipeptide isomers in reversed-phase high performance liquid chromatography. *Journal of Liquid Chromatography, 15*, 2289-2296.

## Figure captions

**Fig. 1.** Reverse-phase ultra-performance liquid chromatographic (RP-UPLC) profile of (A) the whey protein hydrolysate (WPH), (B) its associated 200 Da retentate and (C) permeate. MeCN: acetonitrile. The amino acids Y, F and W elute at 2.9, 4.8 and 7.7 min, respectively.

**Fig. 2.** Plot of predicted versus observed retention times of the set of 153 training peptides. The observed retention times (RTobs) corresponds to the mean of three replicates (n = 3). Peptide trendline, black solid line; confidence limits of the prediction at 95%, black dashed line; confidence limits of the prediction at 99%, grey solid line.

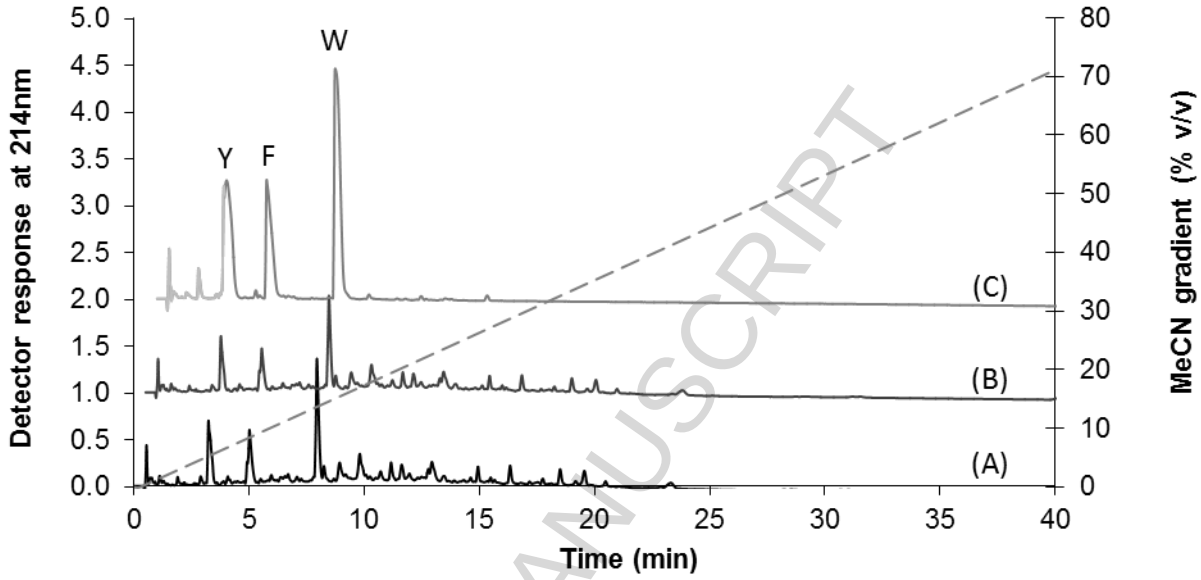


**Fig. 1.**

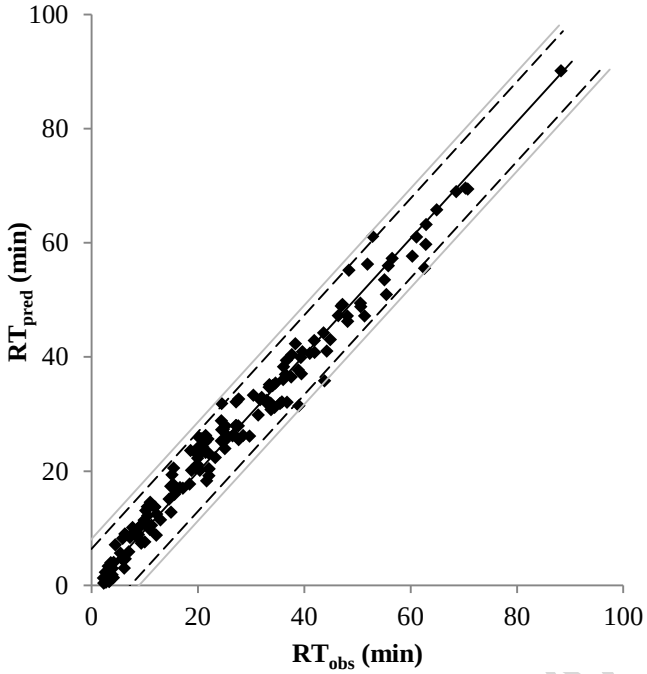


**Fig. 2.**

# Table captions

**Table 1** Dipeptidyl peptidase IV (DPP-IV) inhibitory potency (concentration of sample inducing 50% DPP-IV inhibition - $IC_{50}$) of the whey protein hydrolysate (WPH) and its associated nanofiltration (NF) retentate and permeate.

**Table 2** Peptide and free amino acid identification in the nanofiltration (NF) whey protein hydrolysate (WPH) permeate using the in-house milk protein database and the reverse phase (RP) and hydrophilic interaction liquid chromatography (HILIC) retention time prediction models. The number of possible compounds was based on the $M_w$ determined experimentally by LC-MS/MS with an error of 0.1 Da.

**Table 3** Concentration of peptides, identified within the nanofiltration (NF) whey protein hydrolysate (WPH) permeate, inducing 50% inhibition ($IC_{50}$) of dipeptidyl peptidase IV (DPP-IV).

**Table 1**

| Test sample | DPP IV $IC_{50}$ (mg protein equivalent $mL^{-1}$)* |
|---|---|
| WPH | $0.94 \pm 0.24^{b}$ |
| WPH NF retentate | $0.82 \pm 0.13^{b}$ |
| WPH NF permeate | $0.66 \pm 0.08^{a}$ |

*Values represent mean $IC_{50}$ values ± confidence interval ($P = 0.05$), n=3. Values with different superscript letters are significantly different ($P < 0.05$). The DPP-IV $IC_{50}$ of IPI (diprotin A) was of 0.0013 ± 0.0002 mg $mL^{-1}$.

**Table 2**

| Peptide No. | $M_w$+H (Da) | No. of peptides with the targeted $M_w$ | No. of potential sequences after the models and milk database | Potential sequences (ranked in peptide order)* | Sequences previously assessed for their DPP-IV inhibition** |
|---|---|---|---|---|---|
| 1 | 118.086 | 1 | 1 | V | V |
| 2 | 132.101 | 2 | 2 | I, L | I, L |
| 3 | 132.101 | 2 | 2 | I, L | I, L |
| 4 | 150.058 | 1 | 1 | M | M |
| 5 | 166.086 | 1 | 1 | F | F |
| 6 | 175.118 | 3 | 1 | R | R |
| 7 | 182.080 | 1 | 1 | Y | Y |
| 8 | 189.122 | 6 | 4 | AV, VA > IG, **GI** | AV, VA |
| 9 | 189.122 | 6 | 4 | AV, VA > IG, **GI** | AV, VA |
| 10 | 189.122 | 6 | 5 | **GI**, GL > IG, LG > AV | GL, LG, AV |
| 11 | 205.096 | 7 | 1 | W | W |
| 12 | 231.168 | 6 | 4 | IV, VI, LV, VL | IV, VI, LV, VL, |
| 13 | 295.165 | 6 | 4 | IY, YI, LY, YL | YI, LY, YL |
| 14 | 318.182 | 105 | 4 | IW, WI, LW, WL | IW, WI, LW, WL |
| 15 | 447.295 | 1082 | 2 | **CIVL** > **LCVL** | - |

* Sequences separated by a comma have the same probability of occurrence, whereas ">" showed the sequences with higher probability of occurrence. Novel peptides selected for the DPP-IV inhibitory confirmatory study are highlighted in bold.

** Peptides previously identified for their DPP-IV inhibitory properties as described elsewhere (Lan et al., 2015; Nongonierma & FitzGerald, 2014).

**Table 3**

| Peptide sequence /amino acid | DPP IV $IC_{50}$ (μM)* | Reference |
|---|---|---|
| V | nm | (Nongonierma et al., 2013) |
| L | 3419.25 | (Nongonierma et al., 2013) |
| I | nm | (Nongonierma et al., 2013) |
| M | 2381.51 | (Nongonierma et al., 2013) |
| F | nm | (Nongonierma et al., 2013) |
| W | 4280.40 | (Nongonierma et al., 2013) |
| R | nm | (Nongonierma et al., 2013) |
| Y | nm | (Nongonierma et al., 2013) |
| IPI | 3.73 ± 0.67[a] | this study |
| CIVL | nm | this study |
| LCVL | nm | this study |
| IW | nm | (Nongonierma et al., 2013) |
| LW | 993.4 | (Nongonierma & FitzGerald, 2013c) |
| WL | 43.6 | (Nongonierma & FitzGerald, 2013c) |
| WI | 138.7 | (Nongonierma & FitzGerald, 2013c) |
| VA | 168.2 | (Nongonierma & FitzGerald, 2013a) |
| AV | nm | (Lan et al., 2015) |
| IG | nm | this study |
| GI | nm | this study, (Lan et al., 2015) |
| GL | 2615 | (Nongonierma & FitzGerald, 2013a) |
| LG | nm | this study, (Lan et al., 2015) |
| YI | 1488.50 ± 229.79[c] | this study |
| | nm | (Lan et al., 2015) |
| YL | 940.20 ± 279.00[b] | this study |
| | nm | (Lan et al., 2015) |
| IV | nm | (Lan et al., 2015) |
| VI | nm | (Lan et al., 2015) |
| LV | nm | (Lan et al., 2015; Tulipano, Sibilia, Caroli, & Cocchi, 2011) |
| VL | 74 | (Lan et al., 2014) |
| IY | nm | this study |
| LY | nm | (Lan et al., 2015) |

*Values represent mean $IC_{50}$ values ± confidence interval (*P = 0.05*), n=3. Values with a

different superscript letters are significantly different ($P < 0.05$). nd: not determined; nm: not measureable, i.e.,% DPP-IV inhibition < 50% when tested at the highest concentration evaluated during the dose response experiment.

## Highlights

- Nanofiltration (NF) was used to fractionate a whey protein hydrolysate (WPH).
- An HILIC and RP-UPLC peptide retention time model was developed.
- This strategy allowed characterisation of short peptides within the NF permeate.
- DPP-IV inhibitory activity was increased in the NF permeate compared to WPH.
- VA, VL, WL and WI displayed DPP-IV $IC_{50}$ values lower than 170 µM.