Peptidomics of the zebrafish **Danio rerio** : in search for neuropeptides

Peptidomics of the zebrafish *Danio rerio*: in search for neuropeptides

Kristien A. Van Camp[a], Geert Baggerman[a,b,c], Ronny Blust[a], Steven J. Husson[a,c]

[a] Systemic Physiological & Ecotoxicological Research, University of Antwerp, Department of Biology, Groenenborgerlaan 171, B-2020 Antwerp, Belgium

[b] Flemish Institute for Technological Research (VITO), Boeretang 200, B-2400 Mol, Belgium

[c] Centre for Proteomics (CFP), University of Antwerp - VITO, Groenenborgerlaan 171, B-2020 Antwerp, Belgium

Corresponding author: Steven J. Husson, Steven.Husson@uantwerpen.be

Keywords

Zebrafish, *Danio rerio*; neuropeptide; peptide, peptidomics; LC-MS, mass spectrometry

Abstract

(Neuro)peptides are small messenger molecules that are derived from larger, inactive precursor proteins by the highly controlled action of processing enzymes. These biologically active peptides can be found in all metazoan species where they orchestrate a wide variety of physiological processes. Obviously, detailed knowledge on the actual peptide sequences, including the potential existence of truncated versions or presence of post-translation modifications, is of high importance when studying their function. A peptidomics approach therefore aims to identify and characterize the endogenously present peptide complement of a defined tissue or organism using liquid chromatography and mass spectrometry. While the zebrafish *Danio rerio* is considered as an important aquatic model for medical research, neuroscience, development and ecotoxicology, very little is known about their peptidergic signaling cascades. We therefore set out to biochemically characterize endogenously present (neuro)peptides from the zebrafish brain. This peptidomics setup yielded > 60 different peptides in addition to various truncated versions.

Introduction

In virtually all Metazoan species, a broad range of diverse signaling molecules exist including small molecule neurotransmitters like acetylcholine (ACh), γ-aminobutyric acid (GABA), nitric oxide, excitatory amino acids like glutamate and biogenic amines such as octopamine, tyramine, serotonin (5-HT) and dopamine. In contrast to the small-molecule neurotransmitters, peptidergic signaling molecules are *in vivo* mostly derived from inactive preproproteins or peptide precursors in which one or multiple peptide sequences are contained. The bioactive peptides interact with cell surface receptors (mostly G-protein-coupled receptors (GPCRs)) to trigger an intracellular signaling pathway as to govern a diverse array of physiological processes and behaviors in fish (as in all other Metazoans) such as feeding, locomotion and reproduction. As they are structurally diverse, their signaling

cascades are highly variable, hereby harboring a tremendous potential of different effects on living cells. Because of their critical signaling role, peptides, their processing enzymes or cognate receptors can be considered as attractive targets for pharmaceuticals (1-4). In order to obtain the biologically active entities, inactive preproproteins or peptide precursors have to undergo extensive posttranslational processing in the *trans*-Golgi network and dense core vesicles to produce the bioactive (neuro)peptides. After cleavage of the aminoterminal signal peptide, proprotein convertases (PCs) cleave the remaining part of the precursor at defined cleavage motifs containing basic amino acids (mainly KR and RR, while RK and KK are found in lower frequency; sometimes, the two basic residues are separated from each other by 2, 4, 6 or 8 other residues and were earlier described as “monobasic” cleavage sites (1;4;5). In mammals, these cleavage motifs are specifically recognized by PC2 and PC1/3, reflecting their role in the processing of neuropeptide precursors (1). The neuroendocrine protein 7B2 regulates the activity of PC2 (6;7) whereas proSAAS inhibits PC1/3 activity (8). After processing by the PCs, the resulting intermediate peptides still contain basic residues at the carboxyterminus which are cleft off specific carboxypeptidases (mainly carboxypeptidase E (CPE)) (9). Finally, if a carboxyterminal glycine is present, this amino acid will be transformed into an amide functional group by the action of a bifunctional enzyme peptidylglycine α-amidating monooxygenase (PAM) (10;11). For some species (mostly the invertebrates) the two enzymatic activities of the PAM enzyme is contained in two separate enzymes: peptidylglycine α-hydroxylating monooxygenase (PHM) and peptidyl hydroxyglycine α-amidating lyase (PAL) (12).

Annotation of (neuro)peptide precursors can be quite challenging from genomic sequence information even from well-annotated protein databases. This is due largely to the absence of general discriminating features. Usually the conserved bioactive sequence is short and the only other “specific features” are the presence of a signal peptide and the existence of specific PC cleavage motifs. Even if a peptide precursor can be annotated, predicting the peptides that originate from the precursors can be difficult. Different mechanisms exist to

generate peptide diversity, which is dependent on the actual need at a specific time and place. First of all, cell-specific expression of the respective peptide precursor genes and their processing enzymes can contribute to neuropeptide diversity. Next, resulting mRNAs can be alternatively spliced and alternative proteolytic processing of the resulting precursor proteins in addition to spatiotemporal regulation of post-translational modifications also contribute to peptide diversity. As a consequence, the endogenous peptide content of a cell, tissue or organism, is spatially and temporally dynamic, which has to be taken into account when monitoring peptide profiles. After processing, bioactive peptides can be stored in dense core vesicles prior to their release within the nervous system or peripheral organ systems where most of them will act through G-protein coupled receptors (GPCR) to govern physiological processes in response to internal and external stimuli. This emphasizes the importance of detailed knowledge of the full complement of the wide diversity of actually present neuropeptides. To this end, liquid chromatography and mass spectrometry (LC-MS)-based approaches have been used to identify these peptidergic signaling entities in a plethora of different tissues from different (model) species (see (13-27) for some examples). While important as a model organism, however, no such high-throughput peptidomics analysis has been performed on the freshwater teleost zebrafish (*Danio rerio*) to date, though defined peptidergic signaling systems have been well studied in the zebrafish (28-34). As we still lack a comprehensive overview of all (bioactive) peptides present, we set out to biochemically monitor and identify endogenous peptides from the brain of the zebrafish using a peptidomics workflow.

Materials and Methods

Materials

Water and Acetonitrile (ACN) were LC-MS grade and purchased from Biosolve. Methanol (LC-MS grade) and acetic acid (HPLC grade) were obtained from Sigma-Aldrich. N-hexane,

ethyl acetate, TFA (HPLC grade) and formic acid (FA) were purchased from VWR. Formic acid (FA) was purchased from Merck-Millipore.

Animals and dissection

Wild type male adult zebrafish (*D. rerio*) were obtained from the Zebrafishlab (LA2100621; University of Antwerp) and were maintained on a photoperiod of 14 h light: 10 h dark in US-EPA medium hard water at 25°C; they were fed twice a day *ad libitum* with Vipan flakes. All experimental procedures were approved by the ethical committee of the University of Antwerp (ECD2013-43); fish were sacrificed by decapitation on ice.

Peptide extraction and sample preparation

The entire brain was dissected as quickly as possible and immediately placed in ice cold peptide extraction solvent consisting of methanol:water:acetic acid (90:9:1; v:v:v). The sample was kept on ice and sonicated using a bar sonicator (Branson Sonifier SLPe cell disruptor) prior to centrifugation. The supernatant was collected from which the methanol was evaporated by using a Speedvac concentrator (Eppendorf 5301 concentrator centrifugal evaporator). Lipids were removed from the remaining aqueous residue, containing the peptides, by re-extraction with ethyl acetate and n-hexane. The aqueous solution was subsequently desalted by solid phase extraction using a Pierce C_{18} Spin Column (Thermo scientific) according to the manufacturers procedure. The obtained peptide sample was stored at 4 °C prior to analysis. For analysis by nanoLC-ESI-LTQ-Orbitrap Velos, reconstitution of the sample was performed in 10 µL water containing 2 % ACN and 0.1 % FA.

Peptidomics analysis by liquid chromatography and mass spectrometry

LC-MS analysis was performed on a Eksigent nanoLC-Ultra system connected to a Thermo Scientific LTQ-Orbitrap Velos mass spectrometer. 5 µl of the sample was loaded on the trapping column (pepmap C18 300 µm x 20 mm, Dionex) with an isocratic flow of 2 % ACN in water with 0.1 % FA a flow rate of 5 µl/min. After 2 min, the column-switching valve was switched, placing the pre-column online with the analytical capillary column, a Pepmap C18, 3 µm 75 µm x 150 mm nano column (Dionex). Separation was conducted using a linear gradient from 2 % ACN in water, 0.1 % FA to 40 % ACN in water, 0.1 % FA in 45 minutes. The flow rate was set at 350 nl/min. The LTQ-Orbitrap Velos was set up in a data dependent MS/MS mode where a full scan spectrum (350–5000 m/z, resolution 60000) was followed by a maximum of ten collision-induced dissociation (CID) tandem mass spectra (100 to 2000 m/z). Peptide ions were selected as the 10 most intense peaks of the MS1 scan; single charged ions were rejected. CID scans were acquired in the LTQ ion trap part of the mass spectrometer. The normalized collision energy used was 35 % in CID. We applied a dynamic exclusion list of 45 s. Automated gain control (AGC) target value was set at 5×10^4 ions, maximum inject time was 100 ms, minimum signal threshold 500 counts. Proteome discoverer (1.3) software (Thermo Scientific) was used to perform database searching against the NCBI database nr_20130601 filtered for taxonomy *Dano rerio* using Mascot. Following settings were applied: precursor mass tolerance of 10 ppm, fragment mass tolerance of 0.5 Da. No enzyme was specified as digesting enzyme and 2 missed cleavages are allowed. Pyroglutamic acid, carboxyterminal amidation and methionine-oxidation were set as variable modifications. The results were filtered using following settings: Only medium and high confident peptides with a global FDR < 5 % and first ranked peptides were included in the results.

Results and Discussion

In a first attempt to biochemically identify endogenously present peptides from brain tissues of the zebrafish *Danio rerio*, a peptidomics workflow was employed. The entire brain region of 6 male adult zebrafishes were carefully dissected and 6 independent (neuro)peptide extracts were made using an extraction protocol that is extremely efficient in avoiding the presence of protein degradation products. The 6 peptide samples were analyzed using a nanoLC instrument that is directly coupled with an LTQ-Orbitrap mass spectrometer to yield biochemical identifications of 62 peptides (Table 1, Table S1) that belong to 34 different peptide precursor proteins (Table S2). In addition, a remarkable amount of shortened forms (aminoterminally or carboxyterminally truncated) could be identified, yielding a total number of 105 peptide identifications (all fragmentation spectra are shown in figure S1). These peptide variants may result from further *in vivo* processing in the vesicles or might be the result of extracellular (*in vivo*) peptide processing by specific peptidases. Alternatively, they can also occur from in vitro degradation during sample processing, as has previously been noted (35). Surprisingly, however, the majority of identified peptides don't resemble the predicted canonical "bioactive" peptides. This might be caused by the fact that bioactive peptides and non-tryptic endogenous peptides often have unfavorable ionization properties. In the case of bioactive peptides specifically, peptides are cleaved form the precursor by prohormone convertases at basic amino acid residues. These are then removed by carboxypeptidases. As a result a typical bioactive peptide lacks basic amino acids and on top often has a modified N-terminus resulting in lower ion intensities and, in addition, will produce less predictive and often less informative fragmentation patterns when compared to tryptic peptides. The often low ion intensity and less predictable fragmentation might be an explanation for the relative low numbers of canonical bioactive peptides.

Since in data dependent analysis peptides are selected in a semi-stochastic way, repeated analysis will increase the number of peptides identified. Six independent peptidomics extracts and LC-MS experiments were performed from which accumulative amounts of peptide identifications are plotted in figure 1. From this figure, it shows that further increasing

the number of additional LC-MS runs would only marginally increase the total amount of identified peptides. In fact, this graph suggests that, within the technical limits in terms of sensitivity, most if not all peptides present were sampled. In the next section, we describe the identified peptides together with relevant literature data. In addition to the peptides derived from the classical neuropeptide precursors (having an N-terminal signal sequence for entry into the excretory pathway and displaying the canonical basic cleavage sites), several peptides originating from cytosolic proteins that lack an N-terminal signal sequence could be found (Table S3).

proSAAS is a peptide precursor of which the embedded carboxyterminal peptide binds to PC1/3, hereby inhibiting its prohormone processing function (8). Recently, proSAAS homologs were identified in the lower vertebrates *Xenopus* and the zebrafish (36). Expression was shown to occur in several neural tissues including telencephalon, diencephalon, midbrain, hindbrain, spinal cord, cranial nerves and the pituitary of the zebrafish (36). Though these novel proSAAS homologs are poorly conserved compared to their mammalian counterparts, *Xenopus* and zebrafish proSAAS seem to be functionally conserved as they efficiently inhibit PC1/3 activity (36). However, no information about the actual processing products of the precursors (i.e. the peptides) was available until now. Our peptidomics survey identified 7 endogenous biological peptides in addition to several truncated variants: KPLSAMRGGV, ELRDSVPYEAQMISYPSADF (and the shorter form DSVPYEAQMISYPSADFKS), SNDYYPSEVL, DLSAVSSIERPIKPAL (and the truncated version DLSAVSSIERPIKPA), SLDSAPGPQAEASLLRV (together with the shortened forms SLDSAPGPQAEASLL and SLDSAPGPQAEAS), IDTDLPPPK and ALSYDPALIAQHIL. As can be judged from the position of the 7 proSAAS peptides in the precursor sequence, canonical cleavage motifs were used for the *in vivo* processing that yield the peptides, except for the latter peptide. At times, however, internal basic residues are used to form truncated variants. Similarly, the neuroendocrine protein 7B2, which is also called secretogranin-5, functions as a molecular chaperone for PC2. It is required for the activation of PC2, but the

carboxyterminal peptide inhibits the enzymatic activity of the mature PC2 enzyme (6;37). The peptide VPNPYLMGQRLDNVVA from the zebrafish 7B2 precursor could be sequenced; note that this sequence is flanked by KRRKRS and KK in the 7B2 precursor, suggesting that the aminoterminal S residue has been cleft off.

The secretogranin-2 (Scg2) precursor (also called chromogranin C) has a major role in the formation of secretory granules (38). Processing of the precursor results in secretoneurin (SN), manserin and the EM66 peptide. From the Scg2a precursor, we were able to identify the peptide sequence LSGTEPVSYGPPSQLRPPPSAEMLRAL, which is aminoterminally and carboxyterminally flanked by an arginine. Five additional carboxyterminally truncated versions could also be identified (Figure 2). Two additional peptides that are flanked by typical basic cleavage motifs, SVLKMASPT and YREYPIMFEDDQPL, also appear to be present. From the latter peptide, the two aminoterminally amino acids were also found to be chopped off, suggesting that the R at position two serves as a monobasic cleavage site. The resulting aminoterminal glutamic acid (E) residue was found in both the not-modified form and was found to be transformed to a pyroglutamate (pE). The cyclization of an aminoterminal glutamate (E) (and also glutamine (Q)) to pE occurs in a lot of the known biologically active peptides. The methionine (M) occurred in both oxidized and non-oxidized forms. The previously reported secretoneurin A with the sequence TNENAEEQYTPQKLATLQSVFEELSGIASSKTNT (39;40), however, could not be detected. ATEDLDEQYTPQSLANMRSIFEELGKLSAAQ is the sequence of secretoneurin B (40), from which only the two degradation products and their carboxyterminally truncated versions could be identified: ATEDLDEQYTPQSLANm (note: the small-case m indicates an oxidized methionine (M) residue) and SIFEELGKLSAAQ. The entire secretoneurin B sequence was also detected, although the score for identification was just below the 5 % FDR and hence not included in the list. The presence of the two degradation products most likely result from a targeted cleavage at the middle R residue. Other identified peptides from the Scg2b precursor include RPVADGESPAGDYAGFVKPH (including a variant that is missing the aminoterminal arginine), YPLMFEDEENGRDN, ALGDISEQGMENME (including 10

carboxyterminally truncated variants in which the 2 methionines occur in both oxidized and non-oxidized forms), AAKLLTEYPDTSSSN (including 2 amino- and carboxyterminally truncated variants) and AIDSAANGQLPYEL.

From the chromogranin A (ChgA) precursor, we were able to identify VEmSPKVDE, which is aminoterminally flanked by KR and carboxyterminally flanked by R.

Somatostatins (SST) are involved in the regulation of growth processes in vertebrates and are therefore also referred to as growth hormone–inhibiting hormones (GHIH). A comparative evolutionary study revealed 6 SST genes in the zebrafish (41), while humans only possess one SST gene. Somatostatin has two active forms produced by alternative cleavage of the carboxyterminally part of the peptide precursor at the RK motif that is contained in the middle of the long SST form. In our peptidomics survey, we could identify the aminoterminal parts of the long SST peptides from 3 different precursors: AAGPMLAPRE (including a carboxyterminally truncated version indicating that the R acts as a monobasic cleavage site, and a version with an oxidized M), SAESSNHIPARE (from SST2), and AVYNRLSQLPQ (from SST3, also named cortistatin). We also identified a peptide, APSDAKLRQLLQ (and a carbyxoterminally truncated version), that is flanked by the signal peptide and an R residue in the SST precursor.

The body can produce different families of opioid peptides that are involved in several situations including stress responses, pain, emotion, motivation and control of food intake. We were able to identify the peptide SYFVDDTNPQVLQ from the dynorphin precursor that was originally characterized in the zebrafish in 2007 (42). The dynorphin precursor shows remarkable sequence similarities with three different enkephalin precursors (43) from which we could identify 6 peptides: GGLYALESGVRELQ, VGRPDWWQE, GGLYDLESGVRELQ, AAEIGTGAPAESDGTGAIS, VGRPDWLDNQKSGGLL and TWEEGGETALPDMQ, in addition to several truncated variants.

The neuropeptide FF precursor is processed to yield different peptides that display the RFamide motif at the carboxyterminus; at least one such neuropeptide was found in the

zebrafish: NPSVLHQPQRFamide, in addition to the ITQDEALEQN sequence that is found directly after the predicted signal peptide. From the related neuropeptide VF precursor, contained RFamide peptides could only found at FDR > 5 %; peptides LRLPLSGE (directly after the signal sequence) and TPESPSFPKE were successfully identified. Neuropeptide Y (NPY) is involved in several functions including food intake, fat storage, stress responses, pain modulation, etc. Though the typical 36-amino acid long NPY was not detected in our proteomics study, we were able to identify SSADTLISDLLIGETESRPQT which at least indicates that the respective precursor protein is actually expressed and processed.

Galanin is considered to function as an inhibitory, hyperpolarizing neuropeptide that is involved in diverse biological functions like pain, sleep regulation, feeding, regulation of mood etc. The zebrafish galanin peptide (GWTLNSAGYLLGPYAH) could be identified. In addition, a peptide with the sequence MVIAAKE located directly after the signal peptide of the galanin X1 isoform was also found to be present.

From the corticotropin releasing factor precursor, the peptides SPADSFPETSQYPK and ALDSMERE could be identified. The corticotropin-releasing hormone (CRH) originates from the hypothalamus and regulates release of corticotropin from the pituitary gland.

Thyrotropin-releasing hormone (TRH, also named thyrotropin-releasing factor (TRF) or thyroliberin) is produced in the hypothalamus and stimulates the release of the thyroid-stimulating hormone (TSH, also called thyrotropin) and prolactin from the anterior pituitary. The carboxyterminal peptide contained in the precursor, AELEDELPGLE, and a truncated variant could be identified in the zebrafish.

The very large group of tachykinin peptides with members like the neurokinins (NK) and the well-studied substance P, were initially named due to their ability to induce gut contractions (44). Though typical neurokinin peptides (NK-F and NK-B) displaying the conserved carboxyterminal FxGLMamide (where x is a hydrophobic residue) could only be observed at FDR > 5 %, we were able to rigorously identify the SVSSESPSFRMSTHNLL peptide (and a

carboxyterminally truncated variant from the zebrafish tachykinin 3 precursor. Cloning of the respective Tac3 gene from any fish species took until 2012 and the encoded zebrafish NK-B peptide has been suggested to play an important role in fish reproduction (45).

The glucagon/secretin family of bioactive peptides includes the vasoactive intestinal peptide (VIP) that is functionally involved in vasodilatation, regulation of blood pressure, muscle contractions etc. From the zebrafish VIP type I precursor, the sequence VSDDLMEDQAPM could be identified. Pituitary adenylate cyclase-activating polypeptide (PACAP) is similar to VIP and functions to stimulate adenylate cyclase hereby increasing cAMP levels in target cells. SPPSLTDDLYTLYYPEKRTE and TPIGMTFPKm, in addition to three other variants (TPIGmTFPKM, TPIGmTFPKm, and TPIGMTFPK) were robustly identified from the PACAP precursor. A signal peptide for this precursor could not unambiguously be identified using the Signal P 4.1 server; though we hereby provide a strong biochemical suggestion that the signal peptide contains amino acids 1-24 as the identified peptide starts at position 25, which is in line with the obtained scores from the Signal P algorithm. The identified HADGLLDRAL sequence is the aminoterminal part from the so-called “hormone 2” contained in the PACAP precursor, indicating that the internal R in this putative hormone sequence actually acts as an internal cleavage place.

Mammalian VIP and PACAP homologues from nonmammalian vertebrates were previously named growth hormone-releasing hormones (GH-RH), though they were unable to demonstrate GH-releasing activities. In 2007, GHRH peptides were *in silico* identified from translated zebrafish cDNA sequences (46). From the newly identified precursor, the SPVYPALKFamide peptide, which is located after the signal peptide and in front of the predicted GH-RH peptide, could be biochemically identified in our peptidomics survey. GH-RH peptides function to stimulate the secretion of growth hormone from the anterior pituitary.

Peptides encoded by the cocaine- and amphetamine-regulated transcript precursor (CART) function as endogenous psychostimulants that play a role in feeding behavior, stress and

reward mechanisms, energy homeostasis, learning and memory; similar behaviors are elicited as when exogenous cocaine or amphetamine are administered. Our peptidomics survey yielded the peptide NESEPEIEVELDT which is aminoterminally flanked by the signal peptide and aminoterminally by the RxyR cleavage recognition site in the zebrafish CART precursor. Spatiotemporal distributions of CART peptides in the zebrafish brain revealed immunoreactivities in diverse olfactory, photic, and acoustic-mechanosensory systems, suggesting a function of the CARP peptide in sensory perception, which is in line with conserved evolutionary trends among vertebrates (47).

Calcitonin (CT) is highly conserved among vertebrates where it is involved in calcium homeostasis. CT is mainly produced by the thyroid in mammals and by the corpus ultimobranchiale in other vertebrates. About 5 years ago the CT - calcitonin-gene-related peptide (CGRP) family was investigated in the zebrafish with special emphasis on their spatial distribution and potential involvement in calcium homeostasis (48). While transcripts for the CT precursor were indeed found in the ultimobranchial body only, CGRP transcripts appear to be present in several tissues, but it was predominantly present in the central nervous system. In line with these results, our peptidomics survey of the zebrafish brain tissue revealed the carboxyterminal peptide DLLQSPVYL that originated from the CGRP2 precursor.

Neuromedin B (NMB) is a bombesin-related peptide in mammals where it controls various functions. Surprisingly, very little information about these peptides is available from fish. Recently, a neuromedin-B precursor entry was submitted to NCBI, originating as a (non-validated) part of the DNA sequence of a zebrafish BAC clone. From this putative precursor, We were able to sequence the SISNSQLQDSPFPVKPD peptide, which is flanked by the canonical cleavage motifs in the putative neuromedin B precursor.

The nociceptin precursor belongs to the opioid family and derived peptides like nocistatin; nociceptin / orphanin FQ and orphanin FQ2 are involved in a wide range of biological

functions, such as pain control, locomotor activity, anxiety, stress, feeding, learning and memory, attention and emotions (49-51). We were able to identify the peptide ALGADDQDRQL from the zebrafish nociceptin precursor.

The melanin-concentrating hormone (MCH) is a peptidergic neuromodulator that is involved in a number of behaviors including sleep regulation (52;53). Recently, two MCH genes could be characterized in the zebrafish, *Pmch1* and *Pmch2* (54). We were able to identify the peptide VGSDLSPNFAII and a variant EVGSDLSPNFAIIR from the PMCH1 precursor. The peptide is located just in front of the MCH sequence.

Not only (neuro)peptides from peptide precursors with a clear biological signaling function can be identified in our peptidomics survey, but also endogenously occurring short peptide fragments that are the result of (auto)calatylic processing of enzymes. The proprotein convertase 1 (PC1/3), for example, undergoes several autocatalytic processing events to acquire its catalytic activity. Three peptides that are likely the result of such autocatalytic cleavage reactions could be successfully sequenced: SIADEFGYQLV, pQIGALENHYLF (and the shorter form pQIGALENHYL) and LSEDDRVSWAEQQYE, as both peptides are flanked by the conventional basic residues that are recognized as PC1/3 cleavage motifs. From the peptidyl-glycine α-amidating monooxygenase (PAM) A, the two peptides HSIQNHMYRF and AVELQEPK (and the truncated variant AVELQEP) could be identified. Furthermore, additional peptides from proteins containing a signal peptide could be found: DDPVVALPEVQK from ephrin A1a, EEDTVVNPR from fibrinogen alpha chain (FGA), IAGGLASELEPVVDKA from apolipoprotein and finally peptides SNLPPLGQGEQDGSVFSRPA and SLISEEVLE originating from relaxin 3a.

Conclusions

The zebrafish is a well-established model organism to study vertebrate biology and gene functions. However, very little knowledge about the biochemical peptide entities was

available. In the present study we analysed the endogenous peptides from the zebrafish *Dano rerio* LC-MS to identify 62 peptides. This archive of identified endogenous peptides will aid future research in (neuro)endocrinology in this important model organism. Furthermore, the endogenous peptide content of a cell, tissue or organism, is spatially and temporally dynamic. The methodology that was developed in this paper will allow us to study the changes in peptide expression in response to changes in the organism or the environment using differential peptidomics or molecular imaging techniques such as immunocytochemistry or MALDI imaging. As such, our peptidomics data is likely to aid further functional elucidation of defined neuropeptidergic signalling systems.

Funding

The authors highly appreciate funding from the University Research Fund (Bijzonder Onderzoeksfonds, BOF).

Reference List

(1) Seidah NG, Prat A. The biology and therapeutic targeting of the proprotein convertases. Nat Rev Drug Discov 2012 May;11(5):367-83.

(2) Fosgerau K, Hoffmann T. Peptide therapeutics: current status and future directions. Drug Discov Today 2015 January;20(1):122-8.

(3) Bellmann-Sickert K, Beck-Sickinger AG. Peptide drugs to target G protein-coupled receptors. Trends Pharmacol Sci 2010 September;31(9):434-41.

(4) Fricker LD. Neuropeptide-processing enzymes: applications for drug discovery. AAPS J 2005;7(2):E449-E455.

(5) Veenstra JA. Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. Arch Insect Biochem Physiol 2000 February;43(2):49-63.

(6) Mbikay M, Seidah NG, Chretien M. Neuroendocrine secretory protein 7B2: structure, expression and functions. Biochem J 2001 July 15;357(Pt 2):329-42.

(7) Muller L, Zhu P, Juliano MA, Juliano L, Lindberg I. A 36-residue peptide contains all of the information required for 7B2-mediated activation of prohormone convertase 2. J Biol Chem 1999 July 23;274(30):21471-7.

(8) Qian Y, Devi LA, Mzhavia N, Munzer S, Seidah NG, Fricker LD. The C-terminal region of proSAAS is a potent inhibitor of prohormone convertase 1. J Biol Chem 2000 August 4;275(31):23596-601.

(9) Sapio MR, Fricker LD. Carboxypeptidases in disease: insights from peptidomic studies. Proteomics Clin Appl 2014 June;8(5-6):327-37.

(10) Eipper BA, Milgram SL, Husten EJ, Yun HY, Mains RE. Peptidylglycine alpha-amidating monooxygenase: a multifunctional protein with catalytic, processing, and routing domains. Protein Sci 1993 April;2(4):489-97.

(11) Prigge ST, Mains RE, Eipper BA, Amzel LM. New insights into copper monooxygenases and peptide amidation: structure, mechanism and function. Cell Mol Life Sci 2000 August;57(8-9):1236-59.

(12) Kolhekar AS, Roberts MS, Jiang N, Johnson RC, Mains RE, Eipper BA, Taghert PH. Neuropeptide amidation in *Drosophila*: separate genes encode the two enzymes catalyzing amidation. J Neurosci 1997 February 15;17(4):1363-76.

(13) Che FY, Biswas R, Fricker LD. Relative quantitation of peptides in wild-type and Cpe(fat/fat) mouse pituitary using stable isotopic tags and mass spectrometry. J Mass Spectrom 2005 February;40(2):227-37.

(14) Jurgens M, Appel A, Heine G, Neitz S, Menzel C, Tammen H, Zucht HD. Towards characterization of the human urinary peptidome. Comb Chem High Throughput Screen 2005 December;8(8):757-65.

(15) Li B, Predel R, Neupert S, Hauser F, Tanaka Y, Cazzamali G, Williamson M, Arakane Y, Verleyen P, Schoofs L, Schachtner J, Grimmelikhuijzen CJ, Park Y. Genomics,

transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. Genome Res 2008 January;18(1):113-22.

(16) Li J, Xu X, Xu C, Zhou W, Zhang K, Yu H, Zhang Y, Zheng Y, Rees HH, Lai R, Yang D, Wu J. Anti-infection peptidomics of amphibian skin. Mol Cell Proteomics 2007 May;6(5):882-94.

(17) Verleyen P, Huybrechts J, Sas F, Clynen E, Baggerman G, De Loof A, Schoofs L. Neuropeptidomics of the grey flesh fly, *Neobellieria bullata*. Biochem Biophys Res Commun 2004 April 9;316(3):763-70.

(18) Yuan X, Desiderio DM. Human cerebrospinal fluid peptidomics. J Mass Spectrom 2005 February;40(2):176-81.

(19) Baggerman G, Cerstiaens A, De Loof A, Schoofs L. Peptidomics of the larval *Drosophila melanogaster* central nervous system. J Biol Chem 2002 October 25;277(43):40368-74.

(20) Baggerman G, Boonen K, Verleyen P, De Loof A, Schoofs L. Peptidomic analysis of the larval *Drosophila melanogaster* central nervous system by two-dimensional capillary liquid chromatography quadrupole time-of-flight mass spectrometry. J Mass Spectrom 2005 February 10;40(2):250-60.

(21) Clynen E, Baggerman G, Veelaert D, Cerstiaens A, Van Der Horst D, Harthoorn L, Derua R, Waelkens E, De Loof A, Schoofs L. Peptidomics of the pars intercerebralis-corpus cardiacum complex of the migratory locust, *Locusta migratoria*. Eur J Biochem 2001 April;268(7):1929-39.

(22) Hummon AB, Richmond TA, Verleyen P, Baggerman G, Huybrechts J, Ewing MA, Vierstraete E, Rodriguez-Zas SL, Schoofs L, Robinson GE, Sweedler JV. From the Genome to the Proteome: Uncovering peptides in the *Apis* brain. Science 2006;314:647-9.

(23) Husson SJ, Clynen E, Baggerman G, De Loof A, Schoofs L. Discovering neuropeptides in *Caenorhabditis elegans* by two dimensional liquid chromatography and mass spectrometry. Biochem Biophys Res Commun 2005;335(1):76-86.

(24) Husson SJ, Landuyt B, Nys T, Baggerman G, Boonen K, Clynen E, Lindemans M, Janssen T, Schoofs L. Comparative peptidomics of *Caenorhabditis elegans* versus *C. briggsae* by LC - MALDI-TOF MS. Peptides 2009;30:449-57.

(25) Huybrechts J, Nusbaum MP, Bosch LV, Baggerman G, De Loof A, Schoofs L. Neuropeptidomic analysis of the brain and thoracic ganglion from the Jonah crab, *Cancer borealis*. Biochem Biophys Res Commun 2003 August 29;308(3):535-44.

(26) Huybrechts J, Verleyen P, Schoofs L. Mass spectrometric analysis of head ganglia and neuroendocrine tissue of larval *Galleria mellonella* (Arthropoda, Insecta). J Mass Spectrom 2005 February;40(2):271-6.

(27) Jimenez CR, Spijker S, de Schipper S, Lodder JC, Janse CK, Geraerts WP, van Minnen J, Syed NI, Burlingame AL, Smit AB, Li K. Peptidomics of a single identified neuron reveals diversity of multiple neuropeptides with convergent actions on cellular excitability. J Neurosci 2006 January 11;26(2):518-29.

(28) Abraham E, Palevitch O, Gothilf Y, Zohar Y. The zebrafish as a model system for forebrain GnRH neuronal development. Gen Comp Endocrinol 2009 November;164(2-3):151-60.

(29) Cerda-Reverter JM, Agulleiro MJ, RG R, Sanchez E, Ceinos R, Rotllant J. Fish melanocortin system. Eur J Pharmacol 2011 June 11;660(1):53-60.

(30) Lowry CA, Moore FL. Regulation of behavioral responses by corticotropin-releasing factor. Gen Comp Endocrinol 2006 March;146(1):19-27.

(31) Matsuda K, Kang KS, Sakashita A, Yahashi S, Vaudry H. Behavioral effect of neuropeptides related to feeding regulation in fish. Ann N Y Acad Sci 2011 March;1220:117-26.

(32) Panula P. Hypocretin/orexin in fish physiology with emphasis on zebrafish. Acta Physiol (Oxf) 2010 March;198(3):381-6.

(33) Sherwood NM, Adams BA, Isaac ER, Wu S, Fradinger EA. Knocked down and out: PACAP in development, reproduction and feeding. Peptides 2007 September;28(9):1680-7.

(34) Tsutsui K, Bentley GE, Kriegsfeld LJ, Osugi T, Seong JY, Vaudry H. Discovery and evolutionary history of gonadotrophin-inhibitory hormone and kisspeptin: new key neuropeptides controlling reproduction. J Neuroendocrinol 2010 July;22(7):716-27.

(35) Lee JE, Atkins N, Jr., Hatcher NG, Zamdborg L, Gillette MU, Sweedler JV, Kelleher NL. Endogenous peptide discovery of the rat circadian clock: a focused study of the suprachiasmatic nucleus by ultrahigh performance tandem mass spectrometry. Mol Cell Proteomics 2010 February;9(2):285-97.

(36) Kudo H, Liu J, Jansen EJ, Ozawa A, Panula P, Martens GJ, Lindberg I. Identification of proSAAS homologs in lower vertebrates: conservation of hydrophobic helices and convertase-inhibiting sequences. Endocrinology 2009 March;150(3):1393-9.

(37) Martens GJ, Braks JA, Eib DW, Zhou Y, Lindberg I. The neuroendocrine polypeptide 7B2 is an endogenous inhibitor of prohormone convertase PC2. Proc Natl Acad Sci U S A 1994 June 21;91(13):5784-7.

(38) Courel M, Soler-Jover A, Rodriguez-Flores JL, Mahata SK, Elias S, Montero-Hadjadje M, Anouar Y, Giuly RJ, O'Connor DT, Taupenot L. Pro-hormone secretogranin II regulates dense core secretory granule biogenesis in catecholaminergic cells. J Biol Chem 2010 March 26;285(13):10030-43.

(39) Montero-Hadjadje M, Vaingankar S, Elias S, Tostivint H, Mahata SK, Anouar Y. Chromogranins A and B and secretogranin II: evolutionary and functional aspects. Acta Physiol (Oxf) 2008 February;192(2):309-24.

(40) Trudeau VL, Martyniuk CJ, Zhao E, Hu H, Volkoff H, Decatur WA, Basak A. Is secretoneurin a new hormone? Gen Comp Endocrinol 2012 January 1;175(1):10-8.

(41) Liu Y, Lu D, Zhang Y, Li S, Liu X, Lin H. The evolution of somatostatin in vertebrates. Gene 2010 September 1;463(1-2):21-8.

(42) Gonzalez-Nunez V, Marron Fd, V, Arsequell G, Valencia G, Rodriguez RE. Identification of dynorphin a from zebrafish: a comparative study with mammalian dynorphin A. Neuroscience 2007 January 19;144(2):675-84.

(43) Khalap A, Bagrosky B, Lecaude S, Youson J, Danielson P, Dores RM. Trends in the evolution of the proenkephalin and prodynorphin genes in gnathostomes. Ann N Y Acad Sci 2005 April;1040:22-37.

(44) Maggio JE. Tachykinins. Annu Rev Neurosci 1988;11:13-28.

(45) Biran J, Palevitch O, Ben-Dor S, Levavi-Sivan B. Neurokinin Bs and neurokinin B receptors in zebrafish-potential role in controlling fish reproduction. Proc Natl Acad Sci U S A 2012 June 26;109(26):10269-74.

(46) Lee LT, Siu FK, Tam JK, Lau IT, Wong AO, Lin MC, Vaudry H, Chow BK. Discovery of growth hormone-releasing hormones and receptors in nonmammalian vertebrates. Proc Natl Acad Sci U S A 2007 February 13;104(7):2133-8.

(47) Mukherjee A, Subhedar NK, Ghose A. Ontogeny of the cocaine- and amphetamine-regulated transcript (CART) neuropeptide system in the brain of zebrafish, *Danio rerio*. J Comp Neurol 2012 March 1;520(4):770-97.

(48) Lafont AG, Wang YF, Chen GD, Liao BK, Tseng YC, Huang CJ, Hwang PP. Involvement of calcitonin and its receptor in the control of calcium-regulating genes and calcium homeostasis in zebrafish (*Danio rerio*). J Bone Miner Res 2011 May;26(5):1072-83.

(49) Okuda-Ashitaka E, Minami T, Tachibana S, Yoshihara Y, Nishiuchi Y, Kimura T, Ito S. Nocistatin, a peptide that blocks nociceptin action in pain transmission. Nature 1998 March 19;392(6673):286-9.

(50) Gonzalez-Nunez V, Gonzalez-Sarmiento R, Rodriguez RE. Cloning and characterization of a full-length pronociceptin in zebrafish: evidence of the existence of two different nociceptin sequences in the same precursor. Biochim Biophys Acta 2003 October 1;1629(1-3):114-8.

(51) Meunier JC. Nociceptin/orphanin FQ and the opioid receptor-like ORL1 receptor. Eur J Pharmacol 1997 December 4;340(1):1-15.

(52) Peyron C, Sapin E, Leger L, Luppi PH, Fort P. Role of the melanin-concentrating hormone neuropeptide in sleep regulation. Peptides 2009 November;30(11):2052-9.

(53) Torterolo P, Scorza C, Lagos P, Urbanavicius J, Benedetto L, Pascovich C, Lopez-Hill X, Chase MH, Monti JM. Melanin-Concentrating Hormone (MCH): Role in REM Sleep and Depression. Front Neurosci 2015;9:475.

(54) Berman JR, Skariah G, Maro GS, Mignot E, Mourrain P. Characterization of two melanin-concentrating hormone genes in zebrafish reveals evolutionary and physiological links with the mammalian MCH system. J Comp Neurol 2009;517(5):695-710.

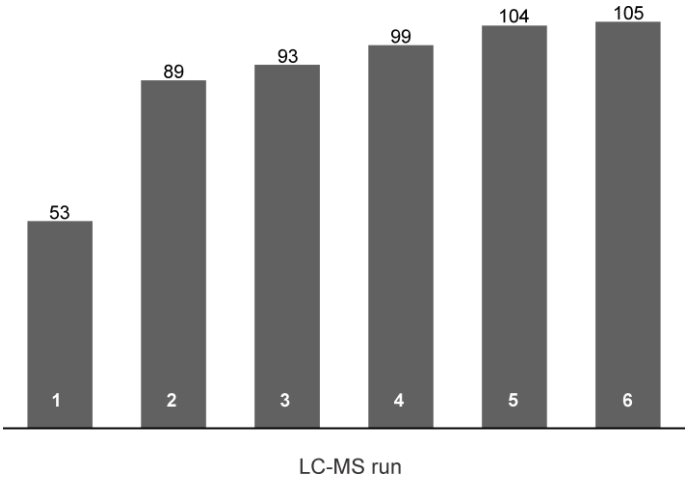
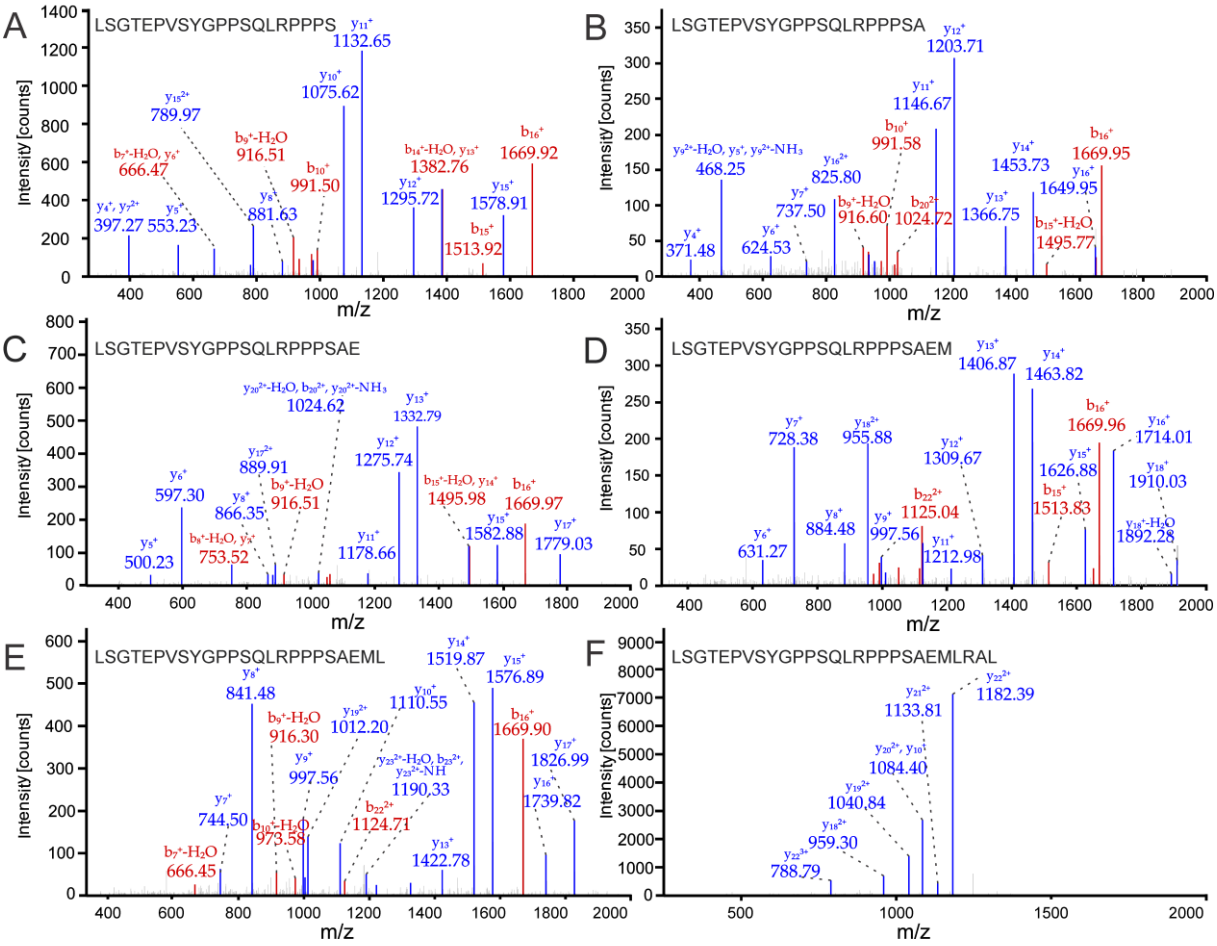


Figure 1: Accumulative count of unique peptide identifications from six LC-MS/MS analysis.

Figure 2: Fragmentation spectra of carboxyterminally truncated peptides from the secretogranin 2a precursor.



A, MS/MS spectrum of LSGTEPVSYGPPSQLRPPPS with $[MH]^+$= 2066.060 Da, ΔM= 2.05 ppm, Z=+2, , ion score= 104; B, MS/MS spectrum of LSGTEPVSYGPPSQLRPPPSA with $[MH]^+$= 2137.085 Da, ΔM= -3.62 ppm, Z=+2, ion score= 63; C, MS/MS spectrum of LSGTEPVSYGPPSQLRPPPSAE with $[MH]^+$= 2266.14141 Da, ΔM= 2.78 ppm, Z=+2, ion score= 59; D, MS/MS spectrum of LSGTEPVSYGPPSQLRPPPSAEM with $[MH]^+$= 2397.17851 Da, ΔM= 1.22 ppm, Z=+2, ion score= 70; E, MS/MS spectrum of LSGTEPVSYGPPSQLRPPPSAEML with with $[MH]^+$= 2510.24956 Da, ΔM= -4.03 ppm, Z=+2, ion score= 50; F, MS/MS spectrum of LSGTEPVSYGPPSQLRPPPSAEMLRAL with $[MH]^+$= 2850.48209 Da, ΔM= -0.04 ppm, Z=+3, ion score= 54. b-type and y-type fragment ions are indicated.