Published in final edited form as:

Mol Biosyst. 2010 August 12; 6(8): 1355–1365. doi:10.1039/c003317k.

Analysis of mouse brain peptides using mass spectrometry-based peptidomics: Implications for novel functions ranging from non-classical neuropeptides to microproteins

Lloyd D. Fricker

Department of Molecular Pharmacology, Albert Einstein College of Medicine, Bronx, NY

Abstract

Peptides are known to play many important physiological roles in signaling. A large number of peptides have been detected in mouse brain extracts using mass spectrometry-based peptidomics studies, and 850 peptides have been identified. Half of these peptides are derived from secretory pathway proteins and many are known bioactive neuropeptides which activate G protein-coupled receptors; these are termed “classical neuropeptides.” In addition, 427 peptides were identified that are derived from non-secretory pathway proteins; the majority are cytosolic, and the remainder are mitochondrial, nuclear, lysosomal, or membrane proteins. Many of these peptides represent the N- or C-terminus of the protein, rather than internal fragments, raising the possibility that they are formed by selective processing rather than protein degradation. In addition to consideration of the cleavage site required to generate the intracellular peptides, their potential functions are discussed. Several of the cytosolic peptides were previously found to interact with receptors and/or otherwise influence cellular activity; examples include hemophins, hemopressins, diazepam binding inhibitor, and hippocampal cholinergic neurostimulating peptide. The possibility that these peptides are secreted from cells and function in cell-cell signaling is discussed. If these intracellular peptides can be shown to be secreted in levels sufficient to produce a biological effect, they would appropriately be called “non-classical neuropeptides” by analogy with non-classical neurotransmitters such as nitric oxide and anandamide. It is also possible that intracellular peptides function as “microproteins” and modulate protein-protein interactions; evidence for this function is discussed, along with future directions that are needed to establish this and other possible functions for peptides.

Introduction

Peptides perform many important physiological roles in mammals and lower organisms.^{1,2} A large number of peptides are secreted from cells and serve as signaling molecules; examples include neuropeptides, peptide hormones, and growth factors. Some secreted peptides function in defense mechanisms; examples include antimicrobial peptides, venoms, and toxins. Within cells, peptides can play signaling roles after release from their parent proteins, although there are relatively few examples of this.³ It is also conceivable that peptides perform roles in modulating protein-protein interactions within cells;⁴ many studies have used synthetic peptides as reagents to either mimic or inhibit protein-protein interactions, demonstrating that peptides are capable of a large number of cellular roles.^{5,6,7,8}

Peptidomics is defined as the analysis of peptides within a tissue.^{9,10,11,12,13} Early studies on peptides required purification of each peptide so that sequencing by Edman degradation could

be performed. With advances in mass spectrometry and related proteomics analyses, it became possible to sequence large numbers of peptides in complex mixtures. Early applications of these techniques focused on lower organism and on mouse pituitary.^{9,14,2} Attempts to study brain using standard mass spectrometry-based peptidomic approaches and extraction conditions previously used for analysis of peptides by radioimmunoassays failed to detect more than a handful of neuropeptides.^{15,16} One of the problems was the high background of peptides that arose from breakdown of cytosolic proteins. A solution to this problem was to affinity-purify neuropeptide processing intermediates from extracts of *Cpe^{fat/fat}* mouse brains.¹⁷ These mice lack carboxypeptidase E (CPE) activity due to a naturally-occurring point mutation, and in the absence of CPE activity the levels of peptide processing intermediates with C-terminal Lys and/or Arg residues accumulate to high levels.¹⁸ These intermediates were purified on an anhydrotrypsin affinity column which bound peptides with C-terminal basic residues.¹⁷ Mass spectrometric analysis of the affinity column eluates revealed a number of known peptide-processing intermediates, demonstrating the usefulness of the technique.¹⁷ In addition, a number of novel peptides were identified and many of these may function as neuropeptides.^{17,19}

While the anhydrotrypsin affinity column approach with *Cpe^{fat/fat}* mouse brain extracts was useful to detect novel peptides, the requirement for mutant mice limited the utility of this technique. A simpler solution to the problem of protein breakdown was to inactivate brain proteases by rapidly heating the brain by microwave irradiation, either as a means of sacrificing live animals or immediately after decapitation and prior to dissection of the brain.^{13,16} Both methods greatly reduced the background level of peptides that arose from cytosolic proteins and enabled many endogenous neuropeptides to be detected in wild-type mouse brain. However, while the levels of some protein fragments were greatly reduced by the introduction of microwave irradiation to heat the brain tissue, a large number of peptides derived from cytosolic and mitochondrial proteins were still detected. Some of these appeared to arise during the extraction procedure as a result of the brief exposure to heat (70°C) and dilute acid (10 mM HCl), based on the relatively high fraction of cleavages at Asp-Pro bonds which are especially sensitive to acid-induced cleavage.²⁰ A comparison of hot and cold acid extraction procedures showed that many of the Asp-Pro cleavages were unique to the hot acid, indicating that they were post-extraction artifacts.²⁰ Still, when extractions were performed without hot acid, a large number of fragments of cytosolic and mitochondrial proteins were detected in the brain peptidome, suggesting that these peptides are normally present in brain.²⁰

In this review, the past five years of peptidomics studies in my laboratory are described and analyzed, but in a distinct angle from previously published reports. Many of our published studies used quantitative peptidomic approaches involving stable isotopic labels to measure relative changes in peptide levels in various mutant mice, or in mice treated with drugs or by food deprivation.^{21,16,20,22,23,24,25,26,27,28} Because peptides are often regulated in ways that reflect their physiological function, quantitative information of relative peptide levels can potentially reveal new functions. This has been the focus of most of the previous quantitative studies, and is an important area for further work. The present analysis is focused on the peptidome itself, rather than on the regulation of peptides. This analysis addresses several important questions: what are the major peptides present in brain?; what proteins do these peptides come from?; and for those peptides from cytosolic and mitochondrial proteins, are there any common features that allow for prediction as to the enzymes that generate them? Putative functions of the brain peptides are also briefly discussed.

Methods used for peptidomics analysis

All data included in the meta-analysis were from studies that subjected the mouse brains to microwave irradiation immediately upon decapitation, prior to removal of the brain and

dissection.^{29,16} Also, all studies included in the present meta-analysis involved extraction conditions that minimize post-extraction modifications by avoiding hot acid.²⁰ The present analysis did not consider the regulation of the levels of peptides, but focused mainly on peptides found in WT mouse brain; these included neuropeptides and other secretory pathway peptides as well as peptides derived from cytosolic and mitochondrial proteins. In addition, some studies included in the present analysis employed anhydrotrypsin affinity columns with *Cpe^{fat/fat}* mouse brain extracts to detect neuropeptide processing intermediates that contain C-terminal Lys and/or Arg residues.¹⁷

The details of the microwave irradiation, peptide extraction and purification, and mass spectrometric analysis have been previously published.^{16,20} In brief, immediately after decapitation the head was placed in an ordinary microwave oven and irradiated for 8 seconds to raise the brain temperature to 80°C. The brain was removed and either used directly or dissected into regions such as amygdala, hippocampus, hypothalamus, prefrontal cortex, cortex other than prefrontal cortex, thalamus, periaqueductal grey, cerebellum, olfactory bulb, and striatum. In some experiments, the striatum was further dissected into caudate putamen and nucleus accumbens. Brain tissues were sonicated in water and heated at 70°C for 20 minutes, then cooled in ice and acidified to 10 mM HCl final concentration.²⁰ Homogenates were centrifuged at 4°C in a microfuge to remove insoluble material. For studies involving *Cpe^{fat/fat}* mice, the extracts were purified on anhydrotrypsin columns, as described.^{17,30} For other studies, the extracts were labeled with trimethylammoniumbutyrate tags, as described.^{20,25} For both types of studies, peptides were filtered through 10 kDa microfiltration units to select for peptides and eliminate proteins. Additional purification was performed on C₁₈ resins to remove salts and other inorganic compounds. Mass spectrometry and analysis of data was performed as described.^{21,16,20,22,23,24,25,26,27,28}

Most peptides were identified from tandem mass spectrometry (MS/MS) fragmentation data and database searches (typically using the Mascot program and public NCBI databases). In addition to the MS/MS data, other criteria were considered such as a precursor mass within the error range of the mass spectrometer (typically 40 parts per million on the older instruments, and within several parts per million for the newer instruments). In addition, the number of isotopic tags incorporated into the peptide had to match the number of free amines present on the peptide (i.e. unmodified N-terminus and Lys side chains) and the observed charge state had to be consistent with the theoretical charge state. For some peptides, there was either no MS/MS data, or the quality of these data was insufficient for a conclusive identification of the peptide. In these cases, if all other criteria matched a known peptide or predicted peptide (based on cleavage sites known to occur), then the peptide was considered tentatively identified and is listed in the supplemental tables with parentheses surrounding the amino acid sequence; all other peptides without parentheses were conclusively identified by MS/MS data as well as the additional criteria.

Caveats and other considerations

Two approaches were used in the studies included in the meta-analysis: affinity columns with extracts of *Cpe^{fat/fat}* mice, and isotopic labeling studies with extracts from WT mice. Both approaches used microwave-irradiated mouse brains and extraction with hot water rather than hot acid in order to minimize postmortem and post-extraction modifications. Still, some post-extraction modifications occurred, such as Met-oxidation which was detected in the present analysis. Peptides identified from the mass spectrometry analysis using both approaches fell into one of two groups: one group consisted of secretory pathway peptides, including neuropeptides, neuropeptide-processing intermediates, and other peptides that may function as neuropeptides (supplemental Table S1); the other group consisted of peptides derived from intracellular proteins such as those in the cytosol or mitochondria (supplemental Table S2).

Altogether, 850 peptides were identified from MS/MS sequencing and other criteria (described above). Over one thousand additional peptides were detected that could not be conclusively identified because there was either no MS/MS information, or the quality of the MS/MS data was not sufficient for positive identification.

The number of times each peptide was found is included in Tables S1 and S2, although this number is not necessarily an accurate reflection on the abundance of the peptides. First, many experiments were done with brain regions such as hypothalamus and only a few were done with other regions (such as cerebellum); therefore a peptide that is expressed only in cerebellum would not be frequently found in the supplementary tables. Second, the counting of peptides was based on the number of times listed in the spreadsheet, which included entries for each charge state observed. While many peptides were found in only a single charge state, some peptides were found in multiple charge states and because each m/z ion was distinct, these were counted separately in the meta-analysis. Third, and most importantly, mass spectrometry favors some peptides, and is optimal for peptides in the 1-3 kDa range that are not highly negatively charged. For example, the proenkephalin precursor contains four copies of Met-enkephalin and only one copy of the heptapeptide sequence, but the heptapeptide was found 117 times while Met-enkephalin was found only 44 times (Table S1). Similarly, prothyrotropin-releasing hormone contains five copies of the bioactive 359 Da tripeptide (thyrotropin-releasing hormone) and single copies of a number of other peptides in the 900-3,700 Da range; while many of these larger peptides were detected multiple times, thyrotropin-releasing hormone was not detected in the mass spectrometry analysis. Because thyrotropin-releasing hormone is known to be produced in brain, the inability to detect this peptide by mass spectrometry clearly demonstrates that the numbers of times each peptide was found is not an absolute reflection of the levels of the peptide. The peptides found infrequently in Tables S1 and S2 are not necessarily low abundance. On the other hand, it is very likely that peptides found many times are both abundant and produce strong signals in electrospray ionization mass spectrometry.

One last point to mention in this section on caveats is that the lists of peptides presented here are only a preliminary report of the mouse brain peptidome. Over one thousand additional peptides were detected but not identified yet, and it is likely that many more peptides are present but not detected either because of the limitations of the mass spectrometry approaches or because their abundance is not sufficient to be detected above the background of other peptides present in the extracts. Advances in mass spectrometry will likely enable the sequence identification of many additional peptides in the next several years.

Neuropeptides and other secretory pathway peptides

Altogether, 423 different peptides were identified that arose from 44 distinct secretory pathway proteins (Table S1). Of these 423 peptides, 271 were detected in WT mouse brain extracts. Peptides found in many experiments include a large number of peptides derived from proSAAS, including little SAAS (found 246 times), big LEN (found 150 times), GAV and PEN (147 times each), little SAAS 1-16 (found 132 times), and many others. Other peptides found more than 50 times include a number of fragments of cerebellin 1 precursor protein, cerebellin 4, chromogranin A and B, procholecystokinin, prodynorphin, proenkephalin, prohormone convertase 2, promelanin concentrating hormone, proneuropeptide Y, proneurotensin, pronociceptin/orphanin FQ, proopiomelanocortin, propeptidyl-amidating-monooxygenase, protachykinin A and B, prothyrotropin releasing hormone, provasopressin, secretogranin II and III, and VGF (Table S1).

Nearly all of the major secretory pathway peptides detected a large number of times arise from cleavage at sites containing basic amino acids that are consensus sites for prohormone

convertases 1/3 and 2. Of the major peptides detected 50 times or more, there are only a few exceptions that do not include a basic residue at a processing site required to generate the peptides; these are cerebellin 1 and cerebellin 2 peptides, a few peptides from procholecystokinin, and a vasopressin C-terminal fragment. Similar analysis of the other peptides found in Table S1 reveals that most represent cleavages at sites containing basic amino acids and only a small percentage reflect cleavages by enzymes other than the secretory pathway prohormone convertases and carboxypeptidases.

In addition to the 271 peptides detected in WT mouse brain extracts, 160 secretory pathway peptides were found after affinity purification from *Cpe^{fat/fat}* mouse brain extracts, and all contained C-terminal basic residues (Table S1). The small numbers of times many of these peptides were found reflects the relatively small number of experiments done with affinity purified extracts of *Cpe^{fat/fat}* mouse brain. The vast majority of these peptides (153 out of 160) were unique to the *Cpe^{fat/fat}* mouse extracts. However, 98 of these were detected in WT mice without the C-terminal basic residues (Table S1), consistent with the efficient removal of C-terminal basic residues by CPE present in the secretory pathway of WT mice.^{31,32}

Intracellular peptides

A total of 427 different peptides were identified that arose from 112 distinct intracellular proteins (Table S2). Approximately 2/3 of these peptides arise from cytosolic proteins (Figure 1A). Peptides from mitochondrial proteins constitute approximately 1/4 of the identified intracellular peptides, and the remaining 10% arise from proteins located in the nucleus, lysosomes, or membrane compartments (Figure 1A). The 21 major peptides detected 50 times or more in the various experiments are derived from 13 different proteins, most of which are very abundant within cells: actin, ATP synthase F0 complex, clathrin light chain A, cytochrome c oxidase subunit 7b, hemoglobin alpha, myelin basic protein, peptidylprolyl isomerase A, thioredoxin 1, thymosin beta-4, thymosin beta-10, tubulin beta, ubiquinol-cytochrome c reductase complex 11 kDa protein, and vacuolar proton pump subunit G2 (Table S2). All of these proteins, including hemoglobin, are known to be expressed in brain cells.^{33,34}

Although the observation that the most frequently detected peptides represent abundant cellular proteins implies a simple correlation between protein levels and protein degradation fragments, further analysis suggests a more complex picture. First, peptides from many other abundant cellular proteins were not found in the analyses. Second, the identified peptides are often the N- or C-terminal fragment of the protein, and not random fragments expected from protein turnover. For example, of the 21 peptides detected 50 times or more, only 5 are internal fragments while 14 are either the N- or C-terminal peptide and 2 represent intact small cellular proteins. When all intracellular peptides are considered, 21% are C-terminal peptides and 21% are N-terminal peptides (Figure 1B). Because the proteolytic machinery that gives rise to the peptides is likely to vary for the different cellular compartments, the data were analyzed separately. Of the cytosolic peptides, approximately 20% represent N- and 20% represent C-terminal peptides. Mitochondrial peptides show a higher contribution from C-terminal peptides (Figure 1B). In contrast, similar analysis of secretory pathway peptides shows the vast majority represent internal fragments of the precursor rather than N- or C-terminal peptides (Figure 1B).

The over-representation of N- and C-terminal fragments in the mouse brain peptidome of intracellular peptides suggests either that these peptides are more stable than other peptides or that these peptides are selectively generated, possibly by limited proteolysis of the protein rather than complete degradation of the protein into peptides. Of the N-terminal fragments, approximately 60% have an N-terminal acetyl group; this modification is expected to stabilize the peptides within the cell by preventing aminopeptidase-mediated digestion. However, the C-terminal and internal fragments are not generally modified on either the N- or C-terminus.

In considering all 427 intracellular peptides, N-terminal Ala is the most abundant residue, either unmodified or acetylated (Figure 2A). However, Ala is also the most abundant amino acid present in the peptide sequences (Figure 3). Adjustment of the N-terminal residue of the peptide based on the frequency of that amino acid in all 427 peptides shows that a broad range of amino acids are present on the N-termini of the peptides (Figure 2B). Because Cys was so rarely found within the peptide sequences, this amino acid was eliminated from this analysis. With the exception of Gln and several acetylated amino acids (AcThr and AcGlu), all other amino acids show only a 2-3-fold variation in relative abundance as N-terminal residues. Thus, selective stability to aminopeptidase activity cannot explain the frequency by which these peptides are found.

To investigate the cleavages responsible for generation of the intracellular peptides, the residues in the P1 position of the cleavage sites were compared. For the N-terminal and C-terminal fragments, the formation of the peptides requires only a single cleavage while the internal fragments require two cleavages. For this analysis, all cleavages were considered together, without consideration for N-terminal, C-terminal, internal, or intracellular location (cytosol, mitochondria, etc); this was done to increase the number of peptides so that individual amino acids can be compared. The most common residue in the P1 position of the cleavage site is Leu, with other common residues including Phe, Ala, Arg, and Asp (Figure 4A). Rare cleavage sites are Ile, Pro, and Gln (Figure 4A). To factor in the relative frequency of occurrence of each amino acid, the amino acid composition of all intracellular peptides was determined; Ala and Leu are the most abundant while Cys and Trp are rarely found (Figure 4A). When the frequency of appearance of residues in the P1 position within the cleavage site was adjusted to the amino acid abundance within the peptides, the hydrophobic residues Phe, Leu, Trp, Tyr, and Met and the basic residue Arg were found to be ~5-10 fold more common in the cleavage site than polar residues such as Glu, Lys, Gln, and Ser (Figure 4B). Interestingly, the hydrophobic residue Ile was rarely found in the P1 position (Figure 4B). Other residues found less frequently in the P1 position relative to other positions within the peptides include Gly and Pro. Because Cys was rarely found within peptides but occasionally found in the cleavage site (Figure 4A), the relative abundance of this amino acid could not be accurately determined.

Discussion: implications, speculation and further directions

The major purpose of this meta-analysis study was to describe the mouse brain peptidome that has been determined over the past five years, involving dozens of experiments and hundreds of LC/MS runs. Altogether, 850 peptides were identified, most by MS/MS sequencing, and over 1000 additional peptides were detected but could not be identified. Of the identified peptides, half are secretory pathway peptides, including many known neuropeptides and others that are likely to have biological functions in cell-cell signaling. Because many known neuropeptides function through G-protein coupled receptors, a reasonable strategy to explore potential functions would be to screen cell lines expressing receptors (both known and “orphan”) with the secretory pathway peptides listed in Table S1. However, it is also possible that peptides bind to other extracellular targets and exert a biological effect independent of G-protein coupled receptors; there are examples where the biological activity of peptide is thought to be through inhibition of enzymes rather than through interactions with receptors.^{35,36,37,38,39}

The finding that half of the peptides identified in mouse brain extracts arise from intracellular proteins, and not secretory pathway proteins, was unexpected. While proteins are constantly undergoing degradation within the cell as part of their normal life cycle and many proteins are specifically regulated by intracellular degradation, the protein degradation fragments are thought to be very unstable and rapidly broken down by intracellular peptidases. Some studies quote a half-life of several seconds for unmodified (i.e. non-acetylated) intracellular peptides,

although only a small number of peptides seem to have been tested in these half-life experiments.^{40,41} Intracellular peptide degradation is thought to involve endopeptidases and/or aminopeptidases.^{42,43,44,45,41,46,47,48,49,50,51} Recent evidence suggests that cytosolic carboxypeptidases are also involved in peptide turnover.^{21,52,53} Many of these intracellular peptidases are thought to be limited to digestion of peptides by structural features that prevent proteins from binding to the active site, thus preventing “nibbling” of the N- or C-termini of proteins in the cytosol.

In considering the large number of peptidases within cells (Table S2) and previous studies showing rapid degradation of unmodified peptides,^{40,41} it is remarkable that so many intracellular peptides are present in the brain. Furthermore, while some of the detected peptides represent abundant proteins, there are many other proteins of equal or greater abundance for which no peptides have been detected. In addition, approximately half of the intracellular peptides represent the N- or C-terminus of the protein. If these peptides were generated from degradation of the proteins during normal protein turnover, more peptides should result from internal cleavages, as found for secretory pathway peptides (Figure 1B). One possible explanation is that the observed intracellular peptides are more stable within the cell than the other degradation fragments from these and other proteins. But why would these peptides be more stable? Some are acetylated, which would protect them from aminopeptidases but not from cytosolic endo- and carboxypeptidases. An intriguing idea is that these peptides are prevented from degradation because they are sequestered, possibly by binding to cellular proteins, RNA, or DNA. Another possibility is that these peptides do not reflect protein degradation but are specifically produced by selective proteolysis. Although the proteasome is considered to be primarily degradative and to release only small peptides once proteins enter the internal chamber, there are other enzymes that have more limited cleavage specificities. For example, the calpains are a class of calcium-activated cytosolic enzymes that perform limited cleavages of proteins, often at hydrophobic amino acids;⁵⁴ this fits the observation that many of the intracellular peptides have hydrophobic residues in the P1 position of the cleavage site (Figure 4). The caspases are another class of intracellular proteases that perform limited cleavages of proteins;^{55,56,57} however, these enzymes cleave at P1 Asp residues, and most of the observed peptides do not fit this pattern (Figure 4). There are many other intracellular proteases although most have limited specificity and are unlikely to contribute to the formation of the observed peptides; examples include secretases, autophagin, and separase.^{58,59,60,61,62} Finally, there are mitochondrial and lysosomal proteases (cathepsins), and some of these are known to be released into the cytosol where they could possibly cleave proteins and/or peptides.^{63,64}

Most proteomics studies do not examine the native forms of peptides, and so little information is available regarding whether the forms of proteins present in cells have undergone processing at the N- and/or C-termini. A recent study using a novel approach to specifically test for the forms of a protein present in cells found many examples of selective trimming of N- and/or C-termini but only under conditions where apoptosis was triggered; under basal conditions there was little evidence that proteins were proteolytically modified.^{65,66} However, this technique may not have detected small changes in the size of the protein due to cleavage near the N- or C-termini. Many of the peptides observed in the peptidomics analyses represent cleavages <20 amino acids from the N- or C-termini (Table S2), which are hard to detect in proteomic studies. Further studies are needed to determine the enzymatic pathways involved in the generation of the peptides that come from intracellular proteins.

A potential role as “non-classical neuropeptides”

A major question is whether the peptides derived from intracellular proteins are functional. A number of peptides derived from cytosolic or mitochondrial proteins have been reported to

interact with proteins on other cells, or otherwise function in cell-cell signaling. Some of the well-studied examples include diazepam binding inhibitor,⁶⁷ hippocampal cholinergic stimulating neuropeptide,^{68,69} and the hemorphins.^{70,71} All of these peptides are derived from intracellular proteins and proposed to function on distant cells after secretion. Because there is no known mechanism for the secretion of intracellular peptides, many neuroscientists are skeptical that these peptides play a role in cell-cell communication. For the hemorphins, which are derived from beta hemoglobin, another problem was that hemoglobin is a well-known component of red blood cells and was assumed to be absent from brain. However, recent studies have convincingly demonstrated that hemoglobin is produced in neurons and glia.^{33,34} In addition to the hemorphins, other hemoglobin-derived peptides have been detected in brain and many of these are distinct from the hemoglobin-derived peptides detected in blood or heart.⁷² This observation suggests that hemoglobin-derived peptides are produced in brain tissue and not just present in brain extracts as a result of blood that is present in brain. Recently, a family of peptides named the hemopressins were found to interact with CB1 cannabinoid receptors: the shorter form is an antagonist/inverse agonist while the longer forms are agonists.^{73,23} Hippocampal cholinergic neurostimulating peptide, the longer forms of hemopressin and hemorphin, and peptides very similar to diazepam binding inhibitor have been found in our peptidomics analysis (Table S2).

For intracellular-derived peptides to be accepted as cell-cell signaling molecules, it will be important to show that these peptides are secreted in levels that are biologically active and also regulated either at the level of secretion or synthesis. Typical neuropeptides (such as the enkephalins) are secreted via the regulated secretory pathway (Figure 5), much like classical neurotransmitters (acetylcholine, dopamine, etc.). In contrast, the non-classical neurotransmitters (anandamide, nitric oxide, etc., which are also called neuromodulators) are not synthesized in advance, but are instead synthesized upon stimulation and then diffuse out of the cell and stimulate other cells through receptor-mediated (anandamide) or non-receptor-mediated (nitric oxide) mechanisms. It is possible that intracellular proteins are processed into peptides upon stimulation and then diffuse out of the cells to exert their biological effect (Figure 5). To differentiate these peptides from the secretory pathway peptides, we propose the term “non-classical neuropeptides” by analogy with non-classical neurotransmitters. (Until they can be shown to be secreted in biologically relevant levels, they should be called “putative” or “potential” non-classical neuropeptides.) Secretory pathway peptides (such as enkephalin) could be called “classical neuropeptides” to differentiate them from the non-classical neuropeptides (Figure 5).

Although the mechanism of secretion is not known, a number of proteins and peptides have been shown to be secreted from the cytosol. Well known examples include cytokines such as interleukins, which are produced in the cytosol but which function in cell-cell signaling.⁷⁴ In addition, many other cytosolic proteins or peptides have been found to be secreted, including large proteins such as thimet oligopeptidase and insulin degrading enzyme, small proteins such as thymosin, and peptides such as hemorphin.^{75,76,77,78,79} Several mechanisms of secretion have been proposed, including direct translocation across the plasma membrane, lysosomal-dependent pathways, microvesicle shedding, and exosome release.⁸⁰ Alternative mechanisms are also conceivable, such as transport into the ER by the TAP transporter followed by routing through the Golgi and secretory pathway; this is the route by which antigenic peptides are transported from the cytosol to the cell surface.⁸¹ Elucidation of the mechanism will help establish the concept that peptides generated in the cytosol of a cell can be secreted and function in cell-cell signaling as non-classical neurotransmitters.

A potential role as “microproteins”

It is possible that some of the intracellular peptides are functional within the cells in which they are synthesized. There are several known examples where the intracellular portion of a membrane protein is cleaved and travels to the nucleus to turn gene transcription on or off.^{3, 82,83} It is also possible that intracellular peptides function by binding to proteins that normally interact with the peptide precursor. If the intracellular peptides are biologically active by mimicking their parent proteins and interacting with target proteins, this would be analogous to microRNAs. Therefore, an appropriate term to refer to peptides that mimic the effects of the parent protein is “microproteins” (Figure 5). These interactions could be either positive (promoting the effect caused by protein binding) or negative (not producing an effect and preventing protein binding). Furthermore, these effects could be in *cis* (only the parent protein is affected) or in *trans* (other protein targets are affected). Much like microRNA, there are several intracellular enzymes that modify the forms of intracellular peptides without degrading them. These cytosolic endopeptidases include thimet oligopeptidase (also known as endopeptidase 24.15), neurolysin (endopeptidase 24.16), insulin-degrading enzyme, prolyl oligoendopeptidase, and prolyl-like oligoendopeptidase.^{84,85,86,87,88} The function of these enzymes is not clear, and although they may contribute to peptide degradation, they have limited specificities and seem to play more of a modulatory role than a degradative one. Further studies are needed to address the exciting possibility that intracellular peptides play a modulatory role as microproteins.

Although there is not yet direct proof that endogenous intracellular peptides play a role in regulating protein-protein interactions, it is clear that the introduction of synthetic peptides into cells can influence protein-protein interactions.^{5,6,7,8} In some cases, low nM concentrations of short peptides 8-10 amino acids in length have biological effects. In most of these studies, the synthetic peptide was not considered to be endogenous because of the dogma that peptides are unstable within cells and rapidly degraded.^{40,41} However, based on the peptidomics results described here and in studies from other laboratories,^{15,79,89} it is clear that many peptides can be detected in brain extracts. Furthermore, some of the identified intracellular peptides are likely to have biological activities, based on studies using synthetic peptides. For example, a peptide corresponding to the N-terminal region of stathmin (residues 6-24) has been found to possess many of the same properties as the full-length protein and to affect tubulin polymerization.⁹⁰ Although stathmin 6-24 has not yet been found in brain, a slightly longer peptide (residues 2-27) has been detected in our analysis (Table S2), and other peptidomic studies have detected other N-terminal stathmin fragments.¹⁵ It remains to be demonstrated if the observed stathmin peptides have the same biological activity as stathmin 6-24 peptide, and if levels of the “endogenous” peptides are sufficient to affect the dynamics of tubulin assembly.

Indirect evidence of the microprotein hypothesis has been provided by studies on *Purkinje cell degeneration (pcd)* mice. Purkinje cells degenerate in these mice between 3-5 weeks of age, and several other cell types also degenerate such as olfactory bulb mitral cells, retinal photoreceptor cells, some thalamic cells, and spermatocytes.^{91,92,93,94,95,96,97} Prior to degeneration, many cellular changes have been observed: reduction and alteration in the shape of the endoplasmic reticulum,^{98,94} including a partial replacement of rough endoplasmic reticulum cisternae by free ribosomes and other changes in ribosome distribution;^{99,100} signs of ER stress;¹⁰¹ an increase in DNA damage/repair foci;⁹⁹ transcriptional inhibition, nucleolar segregation, and the reorganization of nuclear speckles of splicing factors and Cajal bodies;⁹⁹ altered transcription of many genes;^{102,103,104} elevated nitrosylation of Tyr residues presumably due to production of nitric oxide;¹⁰¹ morphological changes in Purkinje cell spines;⁹⁴ and elevated markers of autophagy^{105,21} and apoptosis.^{102,101} The large number of cellular and molecular changes suggested that the mutation was in a gene with a very broad function within the cell. In 2002, Fernandez-Gonzalez and colleagues¹⁰⁶ mapped the *pcd* mutation to

a gene encoding a protein with homology to metallocarboxypeptidases, initially named Nna1¹⁰⁷ and later renamed cytosolic carboxypeptidase 1/Nna1 (CCP1/Nna1).^{52,53} This enzyme is thought to function in peptide turnover by degrading cytosolic peptides formed by the ubiquitin-proteasome system or other intracellular proteases; this hypothesis is based on the recent finding that *pcd* mice show a large increase in the levels of most intracellular peptides.²¹ The broad array of cellular and molecular changes that occur in the *pcd* mice may therefore be due to the elevated levels of intracellular peptides. Further studies are needed to test this possibility, and to obtain proof that in WT mice the levels of endogenous intracellular peptides are sufficient to influence cellular functions.

Peptide identification and post-translational modifications

An important area for further direction is the identification of more peptides. While on one level it would seem like the focus should be to determine the function of those peptides already identified, many additional peptides have been detected by mass spectrometry but not yet identified and it is possible that one or more of these will correspond to peptides already known to be functional in studies using synthetic peptides. Recent improvements in mass spectrometers and in search programs will likely allow for the identification of many additional peptides. Some of the unidentified peptides either had no MS/MS data or weak MS/MS data that could not be interpreted; instruments with greater sensitivity and accuracy will help address this. Additional fractionation procedures should also allow for more peptides to be sequenced. For example, adding an ion exchange step prior to the reverse phase LC/MS analysis typically results in a large increase in the number of identified peptides.¹⁰⁸

Improvements in databases and search strategies may also lead to increases in the numbers of peptides identified from the MS/MS data. Some of the peptides identified from manual analysis of the MS/MS data are polymorphisms or mouse strain-specific variants of known peptides. In some cases, these polymorphisms and variants are found in databases, while others appear to be previously unreported. Current search strategies focus on databases composed of known proteins (such as Swiss-Prot) or theoretical proteins based on translation of mRNAs. While the vast majority of known proteins are contained in these databases, it is possible that peptides are produced from translation of other RNAs within a cell. For example, the 24-residue peptide named humanin is encoded by a short open reading frame of mitochondrial 16s ribosomal RNA.¹⁰⁹ Further studies are needed to conclusively prove the existence of humanin; only immunoreactive peptide has been found and this may represent cross reactivity with another peptide or protein. Still, it is possible that RNA other than mRNA is translated to directly produce cellular peptides; current search programs do not usually take these RNAs into account.

Post-translational modifications present another problem with the current database search programs; unless a modification is specified, the Mascot search program (and other popular programs) cannot identify the peptide. Currently manual interpretation has been necessary, but new search strategies that do not depend on the mass of the parent peptide (and thus don't require knowledge of post-translational modifications) may help identify more peptides.¹¹⁰ Some of the post-translational modifications found in the mouse brain peptidome are modifications known to occur *in vivo*, such as N-terminal pyroglutamylation, Ser and Thr phosphorylation, and C-terminal amidation of secretory pathway proteins, and N-terminal acetylation of cytosolic proteins (Tables S1 and S2). Others are modifications known to occur during extraction and analysis, such as oxidation of Met, and although these could also form *in vivo*, this is difficult to determine. Some modifications could be post-extraction artifacts, such as the loss of water (-18 Da) from peptides. However, unlike Met oxidation that occurred frequently, the loss of water was limited to a small number of peptides and was found for long- and short-forms of the same peptides (Table S1). Thus, it is possible that this modification

represents a specific enzymatic step. Other modifications are not clear, except for the mass added by the modification and, in cases where there was clear MS/MS data, the residue that was modified. For example, +25 Da was found for peptides with an N-terminal Cys residue, and this may represent the addition of a cyano group on the sulfhydryl. But, in both cases where this modification was found, the N-terminal amine of the peptide was not labeled with the isotopic reagent that reacts with primary amines. Because the labeling of amines by the isotopic reagent is virtually complete, the failure of the peptide to incorporate an isotopic tag on the N-terminus suggests that this N-terminal amine is blocked, possibly forming a heterocyclic ring with the carbon of the cyano group on the Cys. This modification has not been previously described. Another strange modification found in a small number of peptides is one that adds 54 Da; this was found in 3 different secretory pathway peptides (Table S1) and one cytosolic peptide (Table S2). In one case, the addition was mapped to a Lys, and in all other cases the peptides contain a Lys residue. Malondialdehyde is a known modification of Lys that adds a net mass of 54 Da, although it is not clear if this is the modification responsible for the observed peptides. Although only a small number of peptides were found with this modification, as well as the others discussed above, these were identified from manual analysis of the MS/MS data and only a small fraction of the data was scanned using this approach. The computer searches used for most of the data analysis did not consider peptides with an additional 25 Da on Cys, 54 Da on Lys, or any other unknown modification, and so these (and other) modifications may be much more common than represented in the data in Tables S1 and S2. Further studies using alternative methods of sequencing peptides that allow for unknown modifications are needed to address this issue.

Conclusion

Many peptides have been detected, and 850 of these identified in the past few years of peptidomics analysis. Many more peptides remain to be identified, and it is likely that advances in mass spectrometers and in search programs will address this. A key question is the function of the numerous peptides present in mouse brain. Many of the peptides secreted via the regulated secretory pathway are likely to function as classical neuropeptides. Only a fraction of the secretory pathway peptides found in mouse brain are known bioactive peptides, and it is likely that additional bioactivities will be found either from those peptides already identified (Table S1), or from novel peptides not yet sequenced. Intracellular peptides may also be secreted; this has already been demonstrated for some peptides in a number of systems but the mechanism of secretion is currently unknown. If secreted in sufficient levels, these peptides may function in cell-cell signaling, and would appropriately be called non-classical neuropeptides. There are many examples of peptides derived from intracellular proteins that have biological activities, but further studies are needed to establish that these are physiologically relevant and meet the criteria to be considered non-classical neuropeptides. Finally, it is possible that intracellular peptides function in modulating protein-protein interactions, or have other roles inside the cell, much like small RNAs can alter the function of mRNAs. If this can be demonstrated to occur with endogenous peptides at physiological levels, an appropriate term to distinguish them from other biologically active peptides would be microproteins.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Many scientists contributed to the work presented in the supplementary tables, with special thanks to present or former members of the laboratory: Fa-Yun Che, Xin Zhang, Hui Pan, Jihyeon Lim, Myrasol Callaway, Reeta Biswas,

Sebastian Tanco, Dries Cardoen, Larkin Elderon, Cain Morano, Chava Ruderman, Iryna Berezniuk, Julia Gelman, Jonathan Wardman, and Juan Sironi. Mass spectrometry was performed in the Laboratory for Macromolecular Analysis and Proteomics of the Albert Einstein College of Medicine, directed by Dr. Ruth Angeletti, and in Brazil through the Rede de Proteoma do Estado de São Paulo in the Laboratório Nacional de Luz Sincrotron, Campinas, SP, Brazil, and in the laboratory of Prof. Fabio Gozzo, Universidade de Campinas. Thanks to Prof. Emer Ferro, Leandro Castro, and Denise Berti for assistance with the mass spectrometry. Major funding was provided by National Institutes of Health grant DA-04494 (to L.D.F.).

References

1. Strand FL. *Prog.Drug Res* 2003;61:1–37. [PubMed: 14674607]
2. Hummon AB, Amare A, Sweedler JV. *Mass Spectrom.Rev* 2006;25:77–98. [PubMed: 15937922]
3. Schweisguth F. *Curr.Biol* 2004;14:R129–R138. [PubMed: 14986688]
4. Ferro ES, Hyslop S, Camargo AC. *J.Neurochem* 2004;91:769–777. [PubMed: 15525330]
5. Churchill EN, Qvit N, Mochly-Rosen D. *Trends Endocrinol.Metab* 2009;20:25–33. [PubMed: 19056296]
6. Rubinstein M, Niv MY. *Biopolymers* 2009;91:505–513. [PubMed: 19226619]
7. Arkin MR, Whitty A. *Curr.Opin.Chem.Biol* 2009;13:284–290. [PubMed: 19553156]
8. Cunha FM, Berti DA, Ferreira ZS, Klitzke CF, Markus RP, Ferro ES. *J.Biol.Chem* 2008;283:24448–24459. [PubMed: 18617518]
9. Baggerman G, Verleyen P, Clynen E, Huybrechts J, De Loof A, Schoofs L. *J.Chromatogr.B Analyt.Technol.Biomed.Life Sci* 2004;803:3–16.
10. Fricker LD, Lim J, Pan H, Che F-Y. *Mass Spectrom.Rev* 2006;25:327–344. [PubMed: 16404746]
11. Schulz-Knappe P, Zucht HD, Heine G, Jurgens M, Hess R, Schrader M. *Comb.Chem.High Throughput.Screen* 2001;4:207–217. [PubMed: 11281836]
12. Clynen E, De Loof A, Schoofs L. *Gen.Comp Endocrinol* 2003;132:1–9. [PubMed: 12765638]
13. Svensson M, Skold K, Svenningsson P, Andren PE. *J.Proteome.Res* 2003;2:213–219. [PubMed: 12716136]
14. Che F-Y, Fricker LD. *J.Mass Spectrom* 2005;40:238–249. [PubMed: 15706629]
15. Skold K, Svensson M, Norrman M, Sjogren B, Svenningsson P, Andren PE. *Proteomics* 2007;7:4445–4456. [PubMed: 18072205]
16. Che F-Y, Lim J, Biswas R, Pan H, Fricker LD. *Mol.Cell.Proteomics* 2005;4:1391–1405. [PubMed: 15970582]
17. Che F-Y, Yan L, Li H, Mzhavia N, Devi L, Fricker LD. *Proc.Natl.Acad.Sci.USA* 2001;98:9971–9976. [PubMed: 11481435]
18. Naggert JK, Fricker LD, Varlamov O, Nishina PM, Rouille Y, Steiner DF, Carroll RJ, Paigen BJ, Leiter EH. *Nature Genetics* 1995;10:135–142. [PubMed: 7663508]
19. Fricker LD, McKinzie AA, Sun J, Curran E, Qian Y, Yan L, Patterson SD, Courchesne PL, Richards B, Levin N, Mzhavia N, Devi LA, Douglass J. *J.Neurosci* 2000;20:639–648. [PubMed: 10632593]
20. Che FY, Zhang X, Berezniuk I, Callaway M, Lim J, Fricker LD. *J.Proteome.Res* 2007;6:4667–4676. [PubMed: 17979226]
21. Berezniuk I, Sironi J, Callaway MB, Castro LM, Hirata IY, Ferro ES, Fricker LD. *FASEB J*. 2010 in press.
22. Fricker LD. *Endocrinology* 2007;148:4185–4190. [PubMed: 17584973]
23. Gomes I, Grushko JS, Golebiewska U, Hoogendoorn S, Gupta A, Heimann AS, Ferro ES, Scarlata S, Fricker LD, Devi LA. *FASEB J* 2009;23:3020–3029. [PubMed: 19380512]
24. Lim J, Berezniuk I, Che F-Y, Parikh R, Biswas R, Pan H, Fricker LD. *J.Neurochem* 2006;96:1169–1181. [PubMed: 16417576]
25. Morano C, Zhang X, Fricker LD. *Anal.Chem* 2008;80:9298–9309. [PubMed: 19551992]
26. Pan H, Che FY, Peng B, Steiner DF, Pintar JE, Fricker LD. *J.Neurochem* 2006;98:1763–1777. [PubMed: 16903874]
27. Zhang X, Che FY, Berezniuk I, Sonmez K, Toll L, Fricker LD. *J.Neurochem* 2008;107:1596–1613. [PubMed: 19014391]

28. Zhang X, Pan H, Peng B, Steiner DF, Pintar JE, Fricker LD. *J.Neurochem* 2010;112:1168–1179. [PubMed: 19968759]
29. Aloy P, Companys V, Vendrell J, Aviles FX, Fricker LD, Coll M, Gomis-Ruth FX. *J.Biol.Chem* 2001;276:16177–16184. [PubMed: 11278909]
30. Che F-Y, Yuan Q, Kalinina E, Fricker LD. *J.Biol.Chem* 2005;280:4451–4461. [PubMed: 15572367]
31. Fricker LD. *Ann.Rev.Physiol* 1988;50:309–321. [PubMed: 2897826]
32. Fricker, LD. *Handbook of Proteolytic Enzymes*. 2 edn.. Barrett, AJ.; Rawlings, ND.; Woessner, JF., editors. Academic Press; San Diego: 2004. p. 840-844.ch. 248
33. Richter F, Meurers BH, Zhu C, Medvedeva VP, Chesselet M-F. *J.Comp.Neurol.* 2009 in press.
34. Biagioli M, Pinto M, Cesselli D, Zaninello M, Lazarevic D, Roncaglia P, Simone R, Vlachouli C, Plessy C, Bertin N, Beltrami A, Kobayashi K, Gallo V, Santoro C, Ferrer I, Rivella S, Beltrami CA, Carninci P, Raviola E, Gustincich S. *Proc.Natl.Acad.Sci.U.S.A* 2009;106:15454–15459. [PubMed: 19717439]
35. Vanderheyden PM. *Mol.Cell Endocrinol* 2009;302:159–166. [PubMed: 19071192]
36. Demaegdt H, Lukaszuk A, De Buyser E, De Backer JP, Szemenyei E, Toth G, Chakravarthy S, Panicker M, Michotte Y, Tourwe D, Vauquelin G. *Mol.Cell Endocrinol* 2009;311:77–86. [PubMed: 19643163]
37. Albiston AL, McDowall SG, Matsacos D, Sim P, Clune E, Mustafa T, Lee J, Mendelsohn FA, Simpson RJ, Connolly LM, Chai SY. *J.Biol.Chem* 2001;276:48623–48626. [PubMed: 11707427]
38. Davies KP. *J.Sex Med* 2009;6(Suppl 3):286–291. [PubMed: 19267851]
39. Wisner A, Dufour E, Messaoudi M, Nejd A, Marcel A, Ungeheuer MN, Rougeot C. *Proc.Natl.Acad.Sci.U.S.A* 2006;103:17979–17984. [PubMed: 17101991]
40. Reits E, Griekspoor A, Neijssen J, Groothuis T, Jalink K, van Veelen P, Janssen H, Calafat J, Drijfhout JW, Neefjes J. *Immunity* 2003;18:97–108. [PubMed: 12530979]
41. Reits E, Neijssen J, Herberths C, Benckhuijsen W, Janssen L, Drijfhout JW, Neefjes J. *Immunity* 2004;20:495–506. [PubMed: 15084277]
42. Towne CF, York IA, Neijssen J, Karow ML, Murphy AJ, Valenzuela DM, Yancopoulos GD, Neefjes JJ, Rock KL. *J.Immunol* 2005;175:6605–6614. [PubMed: 16272315]
43. Dando, PM.; Barrett, AJ. *Handbook of Proteolytic Enzymes*. 1 edn.. Barrett, J.; Rawlings, ND.; Woessner, JF., editors. Academic Press; San Diego: 1998. p. 1013-1016.ch. 343
44. Goldberg AL, Cascio P, Saric T, Rock KL. *Mol.Immunol* 2002;39:147–164. [PubMed: 12200047]
45. Matsui M, Fowler JH, Walling LL. *Biol.Chem* 2006;387:1535–1544. [PubMed: 17132098]
46. Rock KL, York IA, Saric T, Goldberg AL. *Adv.Immunol* 2002;80:1–70. [PubMed: 12078479]
47. Rock KL, York IA, Goldberg AL. *Nat.Immunol* 2004;5:670–677. [PubMed: 15224092]
48. Saric T, Graef CI, Goldberg AL. *J.Biol.Chem* 2004;279:46723–46732. [PubMed: 15328361]
49. Berti DA, Morano C, Russo LC, Castro LM, Cunha FM, Zhang X, Sironi J, Klitzke CF, Ferro ES, Fricker LD. *J.Biol.Chem* 2009;284:14105–14116. [PubMed: 19282285]
50. Silva CL, Portaro FC, Bonato VL, de Camargo AC, Ferro ES. *Biochem.Biophys.Res.Commun* 1999;255:591–595. [PubMed: 10049755]
51. Barrett AJ, Brown MA, Dando PM, Knight CG, McKie N, Rawlings ND, Serizawa A. *Methods Enzymol* 1995;248:529–556. [PubMed: 7674943]
52. Kalinina E, Biswas R, Berezniuk I, Hermoso A, Aviles FX, Fricker LD. *FASEB J* 2007;21:836–850. [PubMed: 17244818]
53. Rodriguez de la Vega M, Sevilla RG, Hermoso A, Lorenzo J, Tanco S, Diez A, Fricker LD, Bautista JM, Aviles FX. *FASEB J* 2007;21:851–865. [PubMed: 17244817]
54. Vosler PS, Brennan CS, Chen J. *Mol.Neurobiol* 2008;38:78–100. [PubMed: 18686046]
55. Denault JB, Salvesen GS. *Curr.Protoc.Protein Sci.* 2002 Chapter 21, Unit-
56. Pop C, Salvesen GS. *J.Biol.Chem* 2009;284:21777–21781. [PubMed: 19473994]
57. Timmer JC, Salvesen GS. *Cell Death.Differ* 2007;14:66–72. [PubMed: 17082814]
58. Vassar R, Kovacs DM, Yan R, Wong PC. *J.Neurosci* 2009;29:12787–12794. [PubMed: 19828790]
59. Golde TE, Wolfe MS, Greenbaum DC. *Semin.Cell Dev.Biol* 2009;20:225–230. [PubMed: 19429495]
60. Wolfe MS. *Chem.Rev* 2009;109:1599–1612. [PubMed: 19226105]

61. Marino G, Uria JA, Puente XS, Quesada V, Bordallo J, Lopez-Otin C. *J.Biol.Chem* 2003;278:3671–3678. [PubMed: 12446702]
62. Pellman D, Christman MF. *Nat.Cell Biol* 2001;3:E207–E209. [PubMed: 11533673]
63. Fruitier I, Garreau I, Piot JM. *Biochem.Biophys.Res.Commun* 1998;246:719–724. [PubMed: 9618278]
64. Turk B, Turk V. *J.Biol.Chem* 2009;284:21783–21787. [PubMed: 19473965]
65. Dix MM, Simon GM, Cravatt BF. *Cell* 2008;134:679–691. [PubMed: 18724940]
66. Simon GM, Dix MM, Cravatt BF. *ACS Chem.Biol* 2009;4:401–408. [PubMed: 19415908]
67. Costa E, Guidotti A. *Life Sci* 1991;49:325–344. [PubMed: 1649940]
68. Ojika K, Mitake S, Tohdoh N, Appel SH, Otsuka Y, Katada E, Matsukawa N. *Prog.Neurobiol* 2000;60:37–83. [PubMed: 10622376]
69. Kim HG, Kim KL. *J.Neurosci.Res* 2007;85:2898–2908. [PubMed: 17628502]
70. Brantl V, Gramsch C, Lottspeich F, Mertz R, Jaeger KH, Herz A. *Eur.J.Pharmacol* 1986;125:309–310. [PubMed: 3743640]
71. Nyberg F, Sanderson K, Glamsta EL. *Biopolymers* 1997;43:147–156. [PubMed: 9216251]
72. Gelman JS, Sironi J, Castro LM, Ferro ES, Fricker LD. *J.Neurochem.* 2010 in press.
73. Heimann AS, Gomes I, Dale CS, Pagano RL, Gupta A, de Souza LL, Luchessi AD, Castro LM, Giorgi R, Rioli V, Ferro ES, Devi LA. *Proc.Natl.Acad.Sci.U.S.A* 2007;104:20588–20593. [PubMed: 18077343]
74. Rubartelli A, Cozzolino F, Talio M, Sitia R. *EMBO J* 1990;9:1503–1510. [PubMed: 2328723]
75. Ferro ES, Tullai JW, Glucksman MJ, Roberts JL. *DNA Cell Biol* 1999;18:781–789. [PubMed: 10541437]
76. Zhao J, Li L, Leissring MA. *Mol.Neurodegener* 2009;4:4. [PubMed: 19144176]
77. Bulloj A, Leal MC, Xu H, Castano EM, Morelli L. *J.Alzheimers.Dis* 2010;19:79–95. [PubMed: 20061628]
78. Goldstein AL, Hannappel E, Kleinman HK. *Trends Mol.Med.* 2005
79. Annangudi SP, Luszpak AE, Kim SH, Ren S, Hatcher NG, Weiler IJ, Thornley KT, Kile BM, Wightman RM, Greenough WT, Sweedler JV. *ACS Chem.Neurosci.* 2010 in press.
80. Nickel W, Rabouille C. *Nat.Rev.Mol.Cell Biol* 2009;10:148–155. [PubMed: 19122676]
81. Herget M, Tampe R. *Pflugers Arch* 2007;453:591–600. [PubMed: 16710701]
82. Raghow R, Yellaturu C, Deng X, Park EA, Elam MB. *Trends Endocrinol.Metab* 2008;19:65–73. [PubMed: 18291668]
83. Rajagopal C, Stone KL, Francone VP, Mains RE, Eipper BA. *J.Biol.Chem* 2009;284:25723–25734. [PubMed: 19635792]
84. Barrett, AJ.; Chen, JM. *Handbook of Proteolytic Enzymes.* 1 edn.. Barrett, AJ.; Rawlings, ND.; Woessner, JF., editors. Academic Press; San Diego: 1998. p. 1108–1112.ch. 371
85. Rioli V, Gozzo FC, Heimann AS, Linardi A, Krieger JE, Shida CS, Almeida PC, Hyslop S, Eberlin MN, Ferro ES. *J.Biol.Chem* 2003;278:8547–8555. [PubMed: 12500972]
86. Fernandez-Gamba A, Leal MC, Morelli L, Castano EM. *Curr.Pharm.Des* 2009;15:3644–3655. [PubMed: 19925417]
87. Gass J, Khosla C. *Cell Mol.Life Sci* 2007;64:345–355. [PubMed: 17160352]
88. Martens K, Derua R, Meulemans S, Waelkens E, Jaeken J, Matthijs G, Creemers JW. *Biol.Chem* 2006;387:879–883. [PubMed: 16913837]
89. Parkin MC, Wei H, O'Callaghan JP, Kennedy RT. *Anal.Chem* 2005;77:6331–6338. [PubMed: 16194096]
90. Clement MJ, Jourdain I, Lachkar S, Savarin P, Gigant B, Knossow M, Toma F, Sobel A, Curmi PA. *Biochemistry* 2005;44:14616–14625. [PubMed: 16262261]
91. Blanks JC, Mullen RJ, LaVail MM. *J.Comp Neurol* 1982;212:231–246. [PubMed: 7153375]
92. Blanks JC, Spee C. *Exp.Eye Res* 1992;54:637–644. [PubMed: 1623950]
93. Greer CA, Shepherd GM. *Brain Res* 1982;235:156–161. [PubMed: 7188319]
94. Landis SC, Mullen RJ. *J.Comp Neurol* 1978;177:125–143. [PubMed: 200636]

95. LaVail MM, Blanks JC, Mullen RJ. *J.Comp Neurol* 1982;212:217–230. [PubMed: 7153374]
96. Handel MA, Dawson M. *Gamete Research* 1981;4:185–192.
97. Wang T, Morgan JI. *Brain Res* 2007;1140:26–40. [PubMed: 16942761]
98. O'Gorman S. *J.Comp Neurol* 1985;234:298–316. [PubMed: 3988986]
99. Valero J, Berciano MT, Weruaga E, Lafarga M, Alonso JR. *Mol.Cell Neurosci* 2006;33:283–295. [PubMed: 16978877]
100. Okubo A, Sameshima M, Unoki K, Uehara F. *Jpn.J.Ophthalmol* 1995;39:152–161. [PubMed: 8538071]
101. Kyuhou S, Kato N, Gemba H. *Neurosci.Lett* 2006;396:91–96. [PubMed: 16356646]
102. Ford GD, Ford BD, Steele EC Jr, Gates A, Hood D, Matthews MA, Mirza S, Macleish PR. *Biochem.Biophys.Res.Commun* 2008;377:556–561. [PubMed: 18930027]
103. Rong Y, Wang T, Morgan JI. *Brain Res.Mol.Brain Res* 2004;132:128–145. [PubMed: 15582153]
104. Xiao R, Park Y, Dirisala VR, Zhang YP, Um SJ, Lee HT, Park C. *Mol.Cells* 2005;20:219–227. [PubMed: 16267396]
105. Chakrabarti L, Eng J, Ivanov N, Garden GA, La Spada AR. *Mol.Brain* 2009;2:24. [PubMed: 19640278]
106. Fernandez-Gonzalez A, La Spada AR, Treadaway J, Higdon JC, Harris BS, Sidman RL, Morgan JI, Zuo J. *Science* 2002;295:1904–1906. [PubMed: 11884758]
107. Harris A, Morgan JI, Pecot M, Soumare A, Osborne A, Soares HD. *Mol.Cell Neurosci* 2000;16:578–596. [PubMed: 11083920]
108. Baggerman G, Boonen K, Verleyen P, De Loof A, Schoofs L. *J.Mass Spectrom* 2005;40:250–260. [PubMed: 15706625]
109. Arakawa T, Kita Y, Niikura T. *Curr.Med.Chem* 2008;15:2086–2098. [PubMed: 18781936]
110. Menschaert G, Vandekerckhove TT, Baggerman G, Landuyt B, Sweedler JV, Schoofs L, Luyten W, Van Crielinge W. *J.Proteome.Res.* 2010
111. Zhou A, Webb G, Zhu X, Steiner DF. *J.Biol.Chem* 1999;274:20745–20748. [PubMed: 10409610]

Figure 1.

Analysis of the mouse brain peptidome. A: The mouse brain peptidome is composed of both secretory pathway and intracellular peptides. For this analysis, the protein precursor that gave rise to each of the 850 identified peptides was considered, and the major intracellular location of the protein was determined from literature sources. B: Analysis of the location of peptide within the protein precursor. Approximately 44% of the intracellular peptides corresponded to the N- or C-terminus of the precursor protein, whereas only a small fraction of the secretory pathway peptides were from the N- or C-terminus (taking into account removal of signal peptide). Bottom row: the intracellular peptides from cytosolic, mitochondrial, and other

locations (nuclear, lysosomal, membrane) were analyzed separately. Numbers in parentheses indicate the peptides in each category.

Figure 2.

Analysis of the N-terminal residue of the intracellular peptides. A: Total number of times each amino acid was present on the N-terminus of the 427 identified intracellular peptides. B: The frequency of each N-terminal amino acid, relative to the abundance of that amino acid within all of the intracellular peptides (i.e. the total amino acid composition of the identified intracellular peptides).

Figure 3.

Total amino acid composition of the identified intracellular peptides. The number of times each amino acid occurred in any position within the identified peptides is indicated for the 427 intracellular peptides.

Figure 4.

Analysis of the cleavage sites required to generate the observed intracellular peptides. A: The number of times each amino acid occurred in the P1 position of the cleavage site is indicated for both upstream and downstream sites. B: The data in panel A was adjusted to the amino acid composition of all intracellular peptides.

Figure 5.

Model for potential functions of intracellular peptides. Top: Classical neuropeptides are produced within the secretory pathway by the selective processing of precursor proteins in the late Golgi and secretory vesicles by endopeptidases (prohormone convertases 1/3 and 2) and carboxypeptidase E.^{111,31} Following synthesis, peptides are stored in secretory vesicles and then secreted in a regulated fashion. Secreted peptides can interact with G protein-coupled receptors (GPCRs) located on adjacent cells, and are degraded by extracellular peptidases. Middle: The term “non-classical neuropeptide” is used to refer to peptides that are formed within the cytosol or other intracellular compartments and are secreted via an unknown mechanism where they can interact with GPCRs or other targets (enzymes, etc) on a nearby cell. In this proposed mechanism, signaling does not have to proceed from presynaptic to postsynaptic but can also occur in a retrograde fashion (similar to the signaling of non-classical neurotransmitters like NO and anandamide). The enzymes that produce the bioactive peptides are not known; candidates include proteasomes, calpains, caspases, and other intracellular proteases. Bottom: Intracellular peptides may also function by binding to proteins and either activating them (i.e. mimicking a protein-protein interaction) or blocking them (i.e. inhibiting a protein-protein interaction). This potential function is termed “microprotein” by analogy with microRNA. In addition to the intracellular proteases that cleave proteins into peptides, there are several peptidases within cells that further process peptides into smaller ones; these enzymes include thimet oligopeptidase (also known as endopeptidase 24.15), neurolysin (endopeptidase 24.16), insulin-degrading enzyme, and prolyl oligopeptidase. Degradation of cyclic peptides occurs via aminopeptidases and possibly via carboxypeptidases.²¹