Abstract—The discovery of insulin in the early 1900s ushered in the era of research related to peptides acting as hormones and neuromodulators, among other regulatory roles. These essential gene products are found in all organisms, from the most primitive to the common ancestor of protostomes and deuterostomes. With the use of relevant animal models and modern technologies, we can gain mechanistic insight into orthologous and paralogous endogenous peptides and translate that knowledge into medically relevant insights and new treatments. Groundbreaking advances in medicine and basic science influence how signaling peptides are defined today. The precise

Mass Spectrometry Approaches Empowering Neuropeptide Discovery and Therapeutics

Krishna D. B. Anapindi, Elena V. Romanova, James W. Checco, and Jonathan V. Sweedler

Department of Chemistry and the Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign, Urbana, Illinois (K.D.B.A., E.V.R., J.V.S.) and Department of Chemistry, University of Nebraska-Lincoln, Lincoln, Nebraska (J.W.C.)

Abstract—The discovery of insulin in the early 1900s ushered in the era of research related to peptides acting as hormones and neuromodulators, among other regulatory roles. These essential gene products are found in all organisms, from the most primitive to the most evolved, and carry important biologic information that coordinates complex physiology and behavior; their misregulation has been implicated in a variety of diseases. The evolutionary origins of at least 30 neuropeptide signaling systems have been traced to the common ancestor of protostomes and deuterostomes. With the use of relevant animal models and modern technologies, we can gain mechanistic insight into orthologous and paralogous endogenous peptides and translate that knowledge into medically relevant insights and new treatments. Groundbreaking advances in medicine and basic science influence how signaling peptides are defined today. The precise

ASSOCIATE EDITOR: VIVIAN HOOK

I. Introduction: A Century of Neuropeptide/Hormone Discovery: Where Are We and Where Are We Going? 663

II. What Are Neuropeptides and Peptide Hormones? 664
A. Neuropeptide Biosynthesis and Processing 664
B. From Structure to Function: Post-Translational Modifications 664
C. Mechanisms of Action of Neuropeptides via Activation of G Protein-Coupled Receptors 665

III. Methods of Neuropeptide Discovery and Characterization 666
A. Probe-Based Methods and Multifaceted ‘Omics 666
B. Analytic Framework and Bioinformatic Tools 666
C. Characterization of PTMs 668

IV. From Discovery to Function 669
A. Decoding Chemical Signaling 669
B. Revealing Prohormone Processing and Differential Neuropeptide Expression 670
C. Neuropeptides in Defined Neuronal Circuits 670
D. Peptide Hormones and Cytokines 670

V. Unusual Routes to Peptide Formation and Function 671
A. Nonclassic Neuropeptides and Cytosolic Peptides 671
B. Mysterious Endomorphins: Neuropeptides Without a Known Gene 672
C. Microproteins and Small ORF-Coded Peptides 672

VI. The Rise of Biologics: Therapeutic Peptides 673
A. Chemical Basis of Peptide Therapeutics 673
B. Improving Stability and Function: Amino Acid Isomerization, Peptide Conjugates 673
C. The Future of Peptide Therapeutics 674

VII. Conclusions 674
References 675

Address correspondence to: Jonathan V. Sweedler, University of Illinois Urbana-Champaign, 600 S. Mathews Avenue, Urbana, IL, 61801. E-mail: jsweedle@illinois.edu

This work was supported by the National Institutes of Health National Institute on Drug Abuse [Grant P30-DA018310] to J.V.S. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

No author has an actual or perceived conflict of interest with the contents of this article.

dx.doi.org/10.1124/pharmrev.121.000423.
mechanistic pathways for over 100 endogenous peptides in mammals are now known and have laid the foundation for multiple drug development pipelines. Peptide biologies have become valuable drugs due to their unique specificity and biologic activity, lack of toxic metabolites, and minimal undesirable interactions. This review outlines modern technologies that enable neuropeptide discovery and characterization, and highlights lessons from nature made possible by neuropeptide research in relevant animal models that is being adopted by the pharmaceutical industry. We conclude with a brief overview of approaches/strategies for effective development of peptides as drugs.

I. Introduction: A Century of Neuropeptide/Hormone Discovery: Where Are We and Where Are We Going?

Recognition of endogenous peptides as a novel class of bioactive molecules has led to ground-breaking discoveries in physiology, endocrinology, neuroscience, and ultimately, medicine in a historically short period of time – about a century. Acting as neurotransmitters, neuropeptides are involved in the regulation of complex systemic functions and numerous behaviors in animals ranging from hydra to humans.

One of the first potent peptides discovered, β-alanyl-l-histidine, was described by Russian scientists in 1900, who measured the nitrogen content in samples of minced meat (Gulewitsch and Amiradzibi, 1900). Accordingly, the new dipeptide was named—“carnosine”—from the Latin term caro, carnis that means meat. Later, this beneficial substance was found in numerous mammalian organs, including the brain (Crush, 1970). A multitude of biologic effects of carnosine were quickly revealed by Russian physiologists, and physicians began using the first medicinal form of this dipeptide before World War II to treat infections and rheumatoid arthritis, peptic ulcers, and essential hypertension (https://karnopedia.com/brief-history/). In 1953, the Russian scientist Severin showed that carnosine buffers lactic acid in working muscles, increasing muscle contractility and resistance to fatigue (Severin et al., 1953). Rapid muscle recovery facilitated by carnosine became known as “Severin’s phenomenon” (Rubtsov, 2001). With the benefits of this antioxidant further investigated by the global research community (Boldyrev, 1992; Banerjee and Poddar, 2020), the dipeptide has been advertised as an ‘elixir of youth’ due to its claimed antiaging effects (Kim and Kim, 2020; Caruso et al., 2021; Prakash et al., 2021) and broadly used (de Courten et al., 2016; Menon et al., 2021; Volkmar, 2021).

Perhaps the next seminal discovery—the hormone insulin—was marked by a controversial Nobel Prize in 1921. Based on the pioneering work of the Canadians Banting, Best, and Macleod, insulin became an effective treatment to lower blood glucose to normal levels in animals and persons suffering from diabetes (https://www.thecanadianencyclopedia.ca/en/article/the-discovery-of-insulin). Then in 1931, substance P was found as an unidentified factor in horse brain and gut tissue extracts that caused intestinal contraction (v Euler and Gaddum, 1931). Isolation of the peptide was a formidable task that was finally accomplished in 1971 (Chang et al., 1971). One difference between substance P and other higher molecular weight peptide regulators from the dipeptide carnosine is that they are formed from high molecular weight peptide precursors by proteolytic post-translational processing. In addition, neuropeptide effects are carried out at low concentrations via selective receptor binding. On the other hand, carnosine is formed by the synthesis of low molecular weight precursors and accumulates in tissues at levels of hundreds to thousands of mg/kg.

Despite the original active peptides being discovered a century ago, the term “neuropeptide” is relatively new, perhaps first coined in 1971 by D. de Wied, who used the term to describe a truncated peptide hormone that lacks the activity of the intact hormone but produces a different effect (Klavdiva, 1995). Understanding the roles of bioactive peptides’ in the body has progressed because we now know that they may act as neurotransmitters, neuropeptides, or hormones, depending on the sites of their synthesis and release. Respectively, they can be classified as either regulatory neuropeptides or peptide hormones.

**Significance Statement**—Neuropeptides, an important class of cell-cell signaling molecules, are involved in maintaining a range of physiological functions. Since the discovery of insulin’s activity, over 100 bioactive peptides and peptide analogs have been used as therapeutics. Because these are complex molecules not easily predicted from a genome and their activity can change with subtle chemical modifications, mass spectrometry (MS) has significantly empowered peptide discovery and characterization. This review highlights contributions of MS-based research towards the development of therapeutic peptides.
but the lines often become blurred as the same peptide can have multiple roles (Kondo and Hayashi, 2021). Only 18 bioactive neuropeptides had been formally recognized by 1978 (Iversen et al., 1978), but the numbers have ‘exploded’ since then. Why did it take so long to discover, characterize, and apply neuropeptide analogs as medicines? An in-depth review (Civelli, 2005) suggests that problem might have been twofold: the lack of robust techniques for chemical characterization and an overall research focus on classic neurotransmitters rather than peptides. Technological advances brought mass spectrometry (MS) into the neuropeptide discovery effort in 1960s. Among the earliest examples was the use of MS to sequence thyrotropin-releasing hormone (TRH) by Guillemin and Schally, recipients of the Nobel Prize in Medicine in 1977 (Burgus et al., 1969). The technology came of age with the advent of soft ionization via electrospray ionization (Konermann et al., 2013) and matrix-assisted laser desorption/ionization (MALDI) (Glish and Vachet, 2003). MS quickly enabled the analysis of neuropeptides at low levels in tissues and biologic fluids without prior derivatization.

Two decades ago, the term ‘peptidomics’ was officially introduced to define a strategy for the direct measurement and structural characterization of endogenous peptides in biologic samples in a high-throughput manner, with robust and unprecedented sensitivity (Schrader et al., 2014). Currently, a given peptidomics study may uncover hundreds, if not thousands of native peptides in a biologic sample. It is safe to say that MS-facilitated discovery of neuropeptides exceeds the rate of their functional characterization and pairing with their cognate receptor(s) (Lee, 2016). As an additional layer of complexity, receptors often bind multiple related bioactive peptides. Today, the precise mechanistic pathways for over 100 endogenous peptides in mammals are reported and lay the foundation for multiple avenues in drug development. Biologies, such as peptides, have become valuable pharmaceuticals due to their unique specificity and biologic activity, lack of toxic metabolites, and unwanted interactions. In this review, we include an overview of the characteristics of neuropeptides that are central to their pharmacological value and approaches/strategies for neuropeptide discovery and development of therapeutic drugs. Specifically, the information pertaining to the peptide receptors and their pathways has resulted in several therapeutics as will be discussed later in this review.

II. What Are Neuropeptides and Peptide Hormones?

A. Neuropeptide Biosynthesis and Processing

Neuropeptides are synthesized, stored, and released by neurons of the central and peripheral nervous system, as well as by neuroendocrine cells throughout the body (see Fig. 1). The key features that distinguish these native peptides from other classes of regulatory molecules are their biosynthesis, mechanism of secretion, and bioactivity at low concentrations. Neuropeptides are enzymatically cleaved from larger precursor proteins (called prohormones) that are translated from the protein-coding mRNA on the ribosomes of the rough endoplasmic reticulum. The prohormones by themselves do not possess biologic activity. Once translated, they are then packaged into secretory granules along with the prohormone processing enzymes, the prohormone convertases, and shuttled toward their site of release (Mains et al., 1987; Fricker, 1991). Unlike classic neurotransmitters, neuropeptides modulate slower and ongoing actions and therefore have slower turnover rates.

B. From Structure to Function: Post-Translational Modifications

Mature neuropeptides, formed via the action of prohormone convertases, often undergo additional enzymatic (and nonenzymatic) post-translational modifications (PTMs) leading to the addition or removal of chemical groups to select amino acid residues, such as glycosylation, phosphorylation, C-terminal amidation, acetylation, and sulfation, to reach their final bioactive forms (Hook et al., 2018). Such PTMs dramatically increase the diversity of native peptides and serve to regulate cellular activity. Often, N-terminal PTMs (e.g., acetylation, pyroglutamylation) enhance neuropeptide stability by blocking the access of aminopeptidases. Modification of the C-terminal, which often serves as the recognition site for receptors and protein complexes, may alter a peptide’s binding ability to its receptor and therefore has functional implications. For example, C-terminal amidation, found in a significant fraction of known bioactive peptides (Kumar et al., 2014), including calcitonin gene-related peptide (CGRP), substance P, and neuropeptide Y, is essential for maintaining their biologic activity (Eipper et al., 1992). Peptides may contain complex PTMs that protect both the N- and C-termini, e.g., TRH, which is modified by pyroglutamate at the N-terminal and by amidation at the C-terminal (Bulant et al., 1990). Phosphorylation of neuropeptides plays a decisive role in the inhibition of bacterial growth (Goumon et al., 1998), proteolytic processing of gastrin precursor (Baldwin et al., 1983), and the affinity of a-melanocyte-stimulating hormone (a-MSH) toward melanocortin receptors (Secher et al., 2016), to name a few. Even the active peptide components of cone snail venoms have a high-degree of disulfide bonds that are essential for maintaining their functionality (Craig et al., 1999).

The entire gamut of peptide PTMs can be broadly classified into two classes based on the nature of the reactions that lead to them: reversible and irreversible modifications. PTMs that result from the covalent addition of chemical groups to amino acid residues, such as phosphorylation, acetylation, and hydroxylation, are often catalyzed by enzymes that can act
reversibly to eliminate the PTM. This class of transient PTMs is usually involved in processes such as signal propagation. Alternatively, irreversible modifications are often a product of chemical changes to the amino acid residue itself and/or facilitated by a non-enzymatic chemical process. For example, glycation, which is an irreversible addition of sugar molecules to proteins and lipids, is an indication of high blood glucose levels, and glycated peptides are often responsible for triggering proinflammatory pathways and oxidative stress. Similarly, deamidation is a non enzymatic process that is indicative of protein degradation (Liddy et al., 2013). These PTMs enable peptides to reach their final bioactive forms and bind to the appropriate receptors to elicit a range of downstream actions. Because a majority of known bioactive neuropeptides are post-translationally modified, structural characterization with respect to PTMs is a requirement of peptide discovery efforts.

C. Mechanisms of Action of Neuropeptides via Activation of G Protein-Coupled Receptors

The mature neuropeptides are released upon cell stimulation and bind to specific cell surface receptors on the target cell to elicit a cascade of events that eventually leads to a change in the biologic state of a cell. Most known neuropeptide receptors belong to a class of transmembrane G protein-coupled receptors (GPCRs). Neuropeptides, which serve as agonists to the GPCRs, bind to specific extracytoplasmic segments of the receptor. Once bound, the neuropeptide ligands induce a structural change in the cytoplasmic domain of the GPCR, which leads to a series of downstream events that eventually alters cell state (Hoyer and Bartfai, 2009). GPCRs are one of the most sought-after drug targets in the pharmaceutical industry due to their involvement in mediating a range of physiologic processes (Rosenbaum et al., 2009).

The human genome encodes ~800 GPCRs, of which ~400 so-called transmitter GPCRs bind various signaling molecules, including (neuro)peptides, and therefore are central to signal transduction throughout the body (Takeda et al., 2002; Vassilatis et al., 2003). Originally many were known as ‘orphan’ receptors when they did not have a known cognate ligand(s); these GPCRs have become valuable for drug discovery efforts and currently are the targets of ~35% of approved drugs (Hauser et al., 2017; Sriram and Insel, 2018), according to databases such as ChEMBL, GtoPdb, DrugBank, and the Drug Repurposing Hub by the Broad Institute.
However, a significant number of transmitter GPCRs still lack information on their physiologic agonists and the marathon to discover endogenous ligands for them continues. One difficulty in deorphanizing GPCRs stems from their transmembrane localization, complicating their three-dimensional modeling and the ability to target them for structure-based drug designs (Hruby, 2002). Novel endogenous ligands for an orphan receptor often illuminate unique physiologic mechanisms and druggable targets. The similarity of numerous remaining orphan GPCRs to structural features of known peptide-activated GPCRs suggest that they may bind peptides (Fricker and Devi, 2018; Foster et al., 2019). The mass characterization of endogenous bioactive peptides accelerated at the end of last century with the advent of effective, unbiased methods of chemical tissue analysis and structure identification. Several neuropeptide receptors that have been deorphanized in the recent past proved to be crucial mediators in various diseases and physiologic functions, including pain, hypertension, depression, obesity, sleep, and learning and memory, to name a few (Hoyer and Bartfai, 2009).

### III. Methods of Neuropeptide Discovery and Characterization

Historically, sequencing via Edman degradation (Edman, 1950) was the earliest effective method for elucidating the primary sequence of peptides. Developed in 1950, the method identifies a protein/peptide sequence by sequentially removing the N-terminal amino acid. Although it is robust and easy to perform, Edman sequencing is slow, requires a high degree of peptide purification, and is best-suited to peptides without N-terminal PTMs. The limitations of Edman sequencing are easily overcome by MS, which can simultaneously characterize the sequences and modifications of low-abundance peptides in complex biological samples (<100 fmol), and now is the primary method for peptide sequencing (Mann, 2016). The fact that MS can perform this characterization without a priori knowledge of the peptide sequence or structure has led to the discovery of hundreds of novel peptides (Hummon et al., 2006; Boonen et al., 2008; Schrader et al., 2014; Hook and Bandeira, 2015).

#### A. Probe-Based Methods and Multifaceted ‘Omics

Neuroanatomical localization for peptides with known sequences can be achieved using probe-based assays, such as immunohistochemistry (IHC) and in-situ hybridization (ISH). Although IHC and ISH by themselves provide little detail regarding peptide function, obtaining information on immune-reactive signal or relevant mRNA distribution in tissues and organs can be crucial to elucidating their functional roles. Given the high sensitivity and specificity of IHC and ISH, virtually every known bioactive peptide has been mapped using these techniques, aiding their functional annotation. For example, mammalian therapeutic peptides, including vasoressin (Gruber et al., 2012) and oxytocin (Lee et al., 2009) followed this roadmap.

The discovery and characterization of peptides with opioid activity deserve special mention due to their crucial role as antipain therapeutics. For example, CGRP, a potent vasodilator involved in nociception, appetite regulation, and temperature regulation, was extensively characterized by several groups via ISH and IHC (Kresse et al., 1992; Van Rossum et al., 1997). Better understanding of the localization of CGRP mRNA in various neuroanatomical regions and CGRP-like immunoreactivity eventually led to the development of drugs for migraine that target the CGRP peptide and its receptor, such as fremanezumab (Silberstein et al., 2017), galcanezumab (Urats et al., 2020), and erenumab (Andreou et al., 2020).

While MS is often considered the go-to method for peptide characterization, its utility is strengthened through integration with probe-based methods. Specifically, for models without a well-annotated genome or an extensively characterized proteome, initial cloning of the pro-hormone of interest and/or mapping of its mRNA may lead to follow-up MS peptide characterization from cells or tissues of interest (Collins et al., 2010; Checco et al., 2018b; Zhang et al., 2018a). A homology annotation against the proteomes of related species was particularly useful in characterizing venom peptides, also termed as integrated venomics (Muttenthaler et al., 2021). The integrated venomics approach combines information obtained from venom duct transcriptomics with tandem MS (MS/MS) de novo sequencing data to deduce the sequence and function of novel venom peptides with therapeutic potential. A multitude of highly potent and neurologically active venomous peptides target a wide range of proteins, including ion channels, enzymes, and GPCRs. However, the genomes of many venomous species are still largely unknown.

#### B. Analytic Framework and Bioinformatic Tools

Here, we briefly outline the peptide discovery process (see Fig. 2). The analysis of neuropeptides from tissues can be categorized into five major steps: 1) postmortem tissue stabilization, 2) peptide extraction and desalting, 3) separation of analyte molecules into semi-pure fractions to reduce the chemical complexity, 4) MS-based analysis, and 5) peptide sequence identification via bioinformatic tools.

The first step in terms of tissue stabilization is critical. After neuropeptides are released, they either bind to a receptor or are degraded, terminating the signal. Thus, tissues tend to have extensive enzymatic pathways that degrade peptides, and postmortem degradation of peptides and proteins by proteolytic enzymes is one of the obstacles to capturing the chemical state of biologic tissues closest to their in vivo state.
To ensure that the endogenous peptide complement is preserved, it is essential to arrest or delay the action of these enzymes. Several approaches have been developed to address this issue in tissues, including flash freezing, heating in boiling water, and microwave denaturation. Enzyme denaturation by rapid conductive heating of tissue has been shown to be particularly effective compared with other methods of preserving endogenous neuropeptides. Once stabilized, the tissue of interest is subjected to peptide extraction followed by desalting to ensure that the salts that interfere with MS are removed before the analysis.

Raw tissue extracts typically present high chemical complexity and a large concentration range of compounds. Techniques that can separate the analytes prior to measurement are often coupled with MS for peptide analysis. Compound separation prior to MS enables both deeper precursor-protein coverage and more accurate peptide quantitation. Reversed-phase liquid chromatography, size-exclusion chromatography, and strong cation exchange are several well-used chromatographic platforms for peptidomics experiments. These systems are either coupled online (via electrospray ionization) or offline (via MALDI) to MS for analysis. Another separation technique for sample separation, capillary electrophoresis, can handle 1–2 orders of magnitude lower sample volumes than liquid chromatography systems, making capillary electrophoresis a useful technique for cellular and sub-cellular peptidomics.

Strategically, MS-based neuropeptide characterization can be categorized into two approaches, targeted (where the experimenter looks for a pre-defined set of peptides) and untargeted (where measurable peptides are profiled without pre-determined sequences). A distinct set of bioinformatics tools are often used in conjunction with MS techniques for neuropeptide characterization. For example, statistical tools such as NeuroPred, NeuroPID, NeuroPIpred, and NeuroPP are routinely implemented to generate a list of peptides predicted to result from prohormone proteins, which then can be validated by MS. Once MS/MS data are available, bioinformatics search engines are used to deduce neuropeptide sequences from the raw MS spectra. These search engines deduce the peptide sequence either based on the peptide fragmentation pattern (known as de novo search) or by comparing the acquired MS spectrum.
to the predicted fragmentation spectrum derived from an existing proteome database (known as a database search). Some of the newer search engines, such as PEAKS Studio (Zhang et al., 2012), perform a hybrid search that combines the strengths of de novo and database methods. In addition to deducing the peptide sequence, a search engine can use the information available in an MS/MS experiment to identify the PTMs on specific amino acid residues based on their corresponding mass difference in the MS/MS spectrum of the peptide. This approach works for most bioactive peptides that have only a few PTMs. However, for peptides with multiple complex modifications, researchers have implemented open searching approaches (PTM-Shepherd) (Geiszler et al., 2021), Bayesian models (PTMProphet) (Shteynberg et al., 2019) and various other specialized scoring algorithms (Na et al., 2012; Spencer et al., 2013).

C. Characterization of PTMs

A challenge of the structural characterization of endogenous peptides is the assignment of PTMs. Because the presence of most PTMs cannot be inferred from gene or transcript sequences alone, they must be measured from the isolated peptides. Probe-based approaches that rely on antibodies targeted to the predicted peptide sequence often cannot distinguish subtle PTMs that would be present in a peptide (unless the antibody was raised specifically against that modification and has high specificity). MS is particularly well-suited for characterizing PTMs (Fig. 3) in neuropeptides because most PTMs lead to changes in mass that are distinguished by most mass spectrometers (Secher et al., 2016; Yu et al., 2017; Lietz et al., 2018; Madsen et al., 2020).

Some PTMs are more challenging to characterize and may require custom-designed approaches specifically targeting the PTM of interest. For example, sulfation is a labile modification that results in a similar mass change as phosphorylation (+79.9568 Da for sulfation and +79.9663 Da for phosphorylation). With a mass difference of 0.01Da, these PTMs are essentially isobaric and can be accurately measured only with high-resolution instruments in combination with isotopic pattern modeling. Practical approaches for high-throughput differentiation between sulfated and phosphorylated peptides in complex sample matrices are often based on the recognition of characteristic product ions by MS/MS (Chen et al., 2018; Yang et al., 2018). Specifically, a combination of higher-energy collisional dissociation and electron-transfer/higher-energy collision dissociation techniques were used to acquire the fragmentation data which was manually deconvoluted to identify a neutral loss signal characteristic of sulfation PTM. Another example of an unconventional PTM is γ-carboxylation of glutamate, particularly in invertebrates. Although previously known to be present in mammals, the role of γ-carboxylglutamate as a modification on an intercellular signaling molecule was reported by our research group for the first time (Jakubowski et al., 2006). We were able to successfully localize the modification on glutamate using the collective information from isotopic labeling, tailored collision energies, and enrichment via prefractionation.

L- to D-residue isomerization is a subtle PTM that results in a D-amino acid-containing peptide (DAACP), and is endogenously found in different phyla (Richter et al., 1987; Kamatani et al., 1989; Heck et al., 1994; Soyez et al., 1994; Torres et al., 2002; Bai et al., 2013; Livnat et al., 2016). DAACPs often have higher bioactivity than their all-L-residue counterparts, and can be more stable to proteolytic degradation (Heck et al., 1994; Jilek et al., 2005; Checco et al., 2018a; Checco et al., 2018b). Despite their high functional importance, characterizing DAACPs via MS is a challenging task because this PTM does not lead to a change in the peptide’s mass. A variety of approaches have been developed to aid in the identification of DAACPs, including enzyme-assisted MS screening (Bai et al.,

![Fig. 3. PTM characterization via MS and their functional implications. MS is effective at discovering both common and unexpected chemical modifications to peptides; these modifications impact receptor binding and lifetime and, hence, need to be characterized to test a peptide's function.](https://pharmacy.aspetjournals.org/2023/09/17/ptm-characterization-via-ms-and-their-functional-implications.ms-is-effective-at-discovering-both-common-and-unexpected-chemical-modifications-to-peptides-these-modifications-impact-receptor-binding-and-lifetime-and-hence-need-to-be-characterized-to-test-a-peptides-function/)
IV. From Discovery to Function

Historically, neuropeptides and peptide hormones were discovered using biologic activity-based screening of extracts of brain or endocrine tissue, followed by rounds of purification where the active fractions were tracked until a pure substance could be isolated. Insulin, substance P, oxytocin, and many others were purified from complex extracts via heroic efforts in processing mountains of organs until a sole fraction retaining the activity was isolated (Du Vigneaud et al., 1953; Burgus et al., 1969; Schally et al., 1971; Brazeau et al., 1973; Miyata et al., 1989). As examples, the purification of pituitary adenylate cyclase activating enzyme required 5000 ovine hypothalami, and neuropeptide Y was isolated from 400 kg of pig brain (Tatemoto et al., 1982; Miyata et al., 1989). With modern MS-based multiomics and shotgun peptidomics, hundreds and even thousands of novel peptides can be discovered in one measurement from samples as small as single hypothalamus or brain nucleus (Anapindi et al., 2019). By narrowing down the output results to mature peptides with PTMs we can target putative bioactive peptides for follow up functional studies. Here, we discuss MS-based approaches aimed at providing insights into the functional significance of peptides.

A. Decoding Chemical Signaling

Neuropeptides are the largest and most diverse set of chemical signals that govern coherent biologic response. The release of neuropeptides can be affected by environmental, pharmacological, and genetic factors and a multitude of disorders. Subtle alterations in secreted molecular signals underpinning a pathology or behavior can be revealed by MS analysis (Song et al., 2019). The sensitivity, specificity, and wide dynamic range for detection of low-abundance secreted peptides makes MS exceptionally well suited for profiling as well as the structural and quantitative characterization of the peptidome in various models, and enables further discovery of pharmacological targets. Secretomics studies, most often based on liquid chromatography-MS/MS technology, have been applied to the analysis of neuropeptides in releasates from cell and organ cultures, brain microdialysates, and cerebrospinal fluid (Bernay et al., 2009; Yin et al., 2012; Tillmaand et al., 2015; Gahoi and Gautam, 2016; Al-Hasani et al., 2018; Beumer et al., 2020; Royds et al., 2021).

Another MS-based approach to probing secreted peptides is the analysis of extracellular vesicles in relation to neurodegenerative and autoimmune diseases and tumors (Kreimer et al., 2015; Bandu et al., 2019; Osti et al., 2019; Xu et al., 2020). As one example, MS analysis of extracellular vesicles in cerebrospinal fluid revealed the presence of peptides from priogenic β-amyloid protein, providing evidence for the postulated mechanisms of the spread of neurodegeneration throughout the nervous system via extracellular vesicles (Chiasserini et al., 2014). As reported by a National Institute on Aging and the Alzheimer’s Association workgroup (McKhan GM, 2011), the presence of β-amyloid peptide is used as a diagnostic tool for dementia. Additionally, MS identified thousands of peptides and proteins in this study (Chiasserini et al., 2014), including peptides from proSAAS, proenkephalin-B, neuroendocrine protein 7B2, cerebellin, and secretogranin, presenting a valuable resource for biomarker discovery in neurologic diseases.

The repertoire of neuropeptides and their localization within an organ or tissue may offer insights into potential pathways and the molecular mechanism of normal physiologic changes and various pathologies. Mapping of endogenous neuropeptide locations can be achieved with mass spectrometry imaging (MSI), which uses a laser or ion beam that is scanned across a tissue to provide spatially resolved localization of a multitude of molecules (Vu et al., 2021). Since MSI is independent of molecular labeling, and thousands of molecules can be measured simultaneously in a single experiment with high molecular specificity directly in tissues and at times, from individual cells (Neumann et al., 2019). An added advantage of MSI is the ability to correlate a tissue’s morphology with its spatial chemistry. Recently, this approach was used to demonstrate alterations in enkephalins, dynorphins, tachykinins, and neurotensin under experimental parkinsonism and L-DOPA therapy in multiple regions of both rat and primate brains (Bourdenx et al., 2014; Hulme et al., 2020). Further, by using MSI to track the opioid peptide levels in brain slices, it has been shown that intracerebral administration of a µ-opioid receptor agonist has a therapeutic effect on severe locomotor deficits in a primate model of Parkinson’s disease. This recent finding suggests a need to reassess Parkinson’s disease pharmacology and points to the importance of the development brain-delivered drug treatments (Bezard et al., 2020).

An alternative to MSI is high-throughput chemical phenotyping of neurons or neuroendocrine cells by single-cell MS-profiling to offer insights into the complexity and heterogeneity of chemical specificity among individual cells, large cell populations, and even subcellular structures (Ong et al., 2015; Comi et al., 2017a; Do et al., 2018; Zhang et al., 2018b). Phenotypic heterogeneity is vital; for example, pancreatic islets of Langerhans are comprised of thousands of cells secreting a well-established repertoire...
of signaling peptides and peptide hormones expressed by distinct cell types (alpha, beta, delta, and gamma). The heterogeneity of islets of Langerhans, which facilitates the endocrine regulation of glucose homeostasis, has been extensively studied by MS at the level of the secretome, peptidome, and proteome of whole islets and their individual cells (Boonen et al., 2007; Waanders et al., 2009; Stewart et al., 2011; Jansson et al., 2016; Schmulbach et al., 2016; Yin et al., 2018).

B. Revealing Prohormone Processing and Differential Neuropeptide Expression

Cells can regulate neuropeptide signaling by changing the gene expression and processing of neuropeptide prohormones and their receptors. Alternative gene splicing and tissue-specific post-translational proteolytic processing of precursor proteins are additional mechanisms of neuropeptide regulation. Both splicing and differential prohormone processing result in subsets of peptides with different sequences and different molecular masses, which can be reliably assessed with MS. As one example, corticotropin hormone, which is processed from the proopiomelanocortin (POMC) prohormone, has been shown to have potential in treating multiple sclerosis, arthritis, lupus, several eye conditions, skin/kidney/lung diseases, and immune system disorders (Bohm and Graessel, 2012). POMC is processed differently in the pituitary gland, secreting adrenocorticotropic hormone for peripheral action, and hypothalamic neurons, where α-MSH and β-endorphin are the main products (Burbach and Wiegant, 1990). In a study combining optical imaging, molecular techniques, and MS-based peptidomics, the POMC peptides were characterized in cultured hypothalamic neurons generated from human pluripotent stem cells and human brain, demonstrating that physiologically-relevant concentrations of leptin hormone modulate levels of secreted POMC peptides, which may have downstream effects on energy homeostasis (Kirwan et al., 2018). Peptidomic analysis of the proprotein convertase activation of neuropeptides, substance P and neuropeptide K in spinal cord demonstrated that the C-terminal truncated peptides retained neuropeptide K receptor affinities and may have a role in nociception (Saidi et al., 2016; Salem et al., 2018).

C. Neuropeptides in Defined Neuronal Circuits

Peptidomics has been actively employed to explore the neurochemistry in defined neuronal circuits in animal models of human diseases. Because a wide range of degenerative diseases and psychiatric disorders arise from the dysfunction of neural circuits, understanding how they influence physiology and behavior aids biomedical research. Exploring the neurochemistry of central and peripheral circuits clarifies their organization, explains normal function, and illuminates changes that lead to neurologic disorders. One interesting group of neurons are located in the dorsal root ganglia, which are structurally associated with the spinal nerve and its dorsal and ventral roots. The dorsal root ganglia circuitry innervates a range of peripheral organs, including skin, muscles, and joints, and is involved in mechanoreception, thermoreception, nociception, and limb proprioception (Haberberger et al., 2019). Many common pain medications, including prescription opioids, target this circuit, and localization of several therapeutically important cell-to-cell signaling peptides have been prominently identified within the dorsal root ganglia and spinal cord by various MS approaches (Sui et al., 2017; Do et al., 2018; Sandor et al., 2018; Anapindi et al., 2019).

Another neuronal circuit that has been explored by MS is the brain reward system, activated by reinforcing stimuli from all pleasurable and often habit-forming experiences like food, sex, and, unfortunately, illicit drugs (Natividad et al., 2018). Although relatively rarely measured, the habenal nuclei functionally link the forebrain and the midbrain and are involved in sleep regulation, reward-based decision-making, mood-related behavior, and drug addiction. A multiplatform MS analysis of the two habenula substructures, the medial and lateral subnuclei, revealed highly overlapping but distinct neuropeptide profiles and characterized novel proteolytic products of pituitary adenylate cyclase activating enzyme, protachikinin-1, and secretogranin prohormones. In addition, the presence of sulfotyrosine on several mature secretogranin peptides pointed to the possibility of ligand-receptor interactions by association with studies in other animal models, which makes novel secretogranin peptides potential candidates for further evaluation of biologic activity (Yang et al., 2018). By applying quantitative peptidomics, chronic nicotine consumption was shown to upregulate proenkephalin opioid peptides, but not prodynorphin peptides, in the dorsal striatum, a region implicated in habitual learning and in action initiation (Petruzziello et al., 2013).

D. Peptide Hormones and Cytokines

Many of the neuropeptides in the nervous and endocrine systems were first discovered in the periphery, where they function as hormones, e.g., hormones of the hypothalamic-pituitary-gonadal axis, such as corticotropin, melanocortin, neurtensin, oxytocin, vasoactive intestinal polypeptide, angiotensin II, vasopressin, and atrial natriuretic peptide. Peptide hormones are of great interest to the pharmaceutical industry as they lack the cytotoxicity of xenobiotics while offering higher specificity (Flatt and Conlon, 2018). The most recognized therapeutic peptide hormones, insulin, glucagon, and glucagon-like peptide-1 (GLP-1), work together to regulate glucose blood levels. Secretion of GLP-1 from endocrine gut cells
...primarily work on a different aspect of homeostasis...of magnitude lower than peptide hormones. Cytokines...circulate at picomolar concentrations, or three orders...mechanisms, with one notable distinction in that they...is facilitated by autocrine, paracrine, and endocrine...types or discrete glands. Like hormones, their action...synthesis of cytokines is not limited to specialized cell...response to inflammatory or immune stimuli. The...enous modulators are synthesized specifically in...neuropeptides and peptide hormones that are pre-...tides but as growth factors and cytokines. Unlike...pathogenesis of type 1 diabetes (Wan et al., 2018;...Wiles and Delong, 2019; Wiles et al., 2019). Hybrid...insulin peptides form via a cross-linking reaction between a C-terminal carboxylic acid group of one...peptide and the N-terminal amine group of another...peptide, a unique PTM that cannot happen spontaneously...under physiologic conditions. The chemical activation of a carboxylic acid group that is required for...reaction is thought to occur through reversed proteolysis...and molecular crowding in secretory granules of...pancreatic beta cells (Mishto et al., 2012). Ongoing...efforts in islet peptidome and hybrid insulin peptide...characterization are facilitated by MS analysis of pancreatic...tissue, individual islets, and their secretome...Jansson et al., 2016; Comi et al., 2017b; Yin et al.,...Prentice et al., 2019; Donohue et al., 2021).

Another prominent example is human growth hormone and its mediator, insulin growth factor-1, used...therapeutically to treat longitudinal growth and somatic development issues in children, pituitary pathologies, infertility, as well as for doping/performance enhancement...in sports for improved oxygen utilization and maximal power output (Bidlingmaier and Strasburger, 2010). Traditionally, peptide hormones have been detected and...quantified by immunoassay in both research and clinical settings. The wide spectrum of molecular human growth hormone isoforms, however, illuminates the limitation of such approaches. Recently, MS has been shown to be effective for the analysis of human growth hormone proteoforms, PTMs, and related peptides (Ketha and Singh, 2016; Tanna et al., 2020; Pratt et al., 2021).

A category of larger signaling peptides, or small proteins, generated by non-neuronal cells in the nervous system are not formally referred to as neuropeptides but as growth factors and cytokines. Unlike neuropeptides and peptide hormones that are preformed and stored in secretory granules, these endogenous modulators are synthesized specifically in response to inflammatory or immune stimuli. The synthesis of cytokines is not limited to specialized cell types or discrete glands. Like hormones, their action is facilitated by autocrine, paracrine, and endocrine mechanisms, with one notable distinction in that they circulate at picomolar concentrations, or three orders of magnitude lower than peptide hormones. Cytokines primarily work on a different aspect of homeostasis...restoration, including inflammation and innate and adaptive immunity, and are increasingly being used as therapeutics (Schoollink and Rose-John, 2002). Quantitative MS, as well as a shotgun approach for diagnostic cytokine profiling, are gaining momentum as robust alternatives to antibody-based assays (Kupcova Skalnikova et al., 2017).

V. Unusual Routes to Peptide Formation and Function

Generally, to be classified as a cell-cell signaling molecule, a peptide must satisfy several criteria: i) be synthesized via the secretory pathway and released upon stimulation; ii) show specificity toward receptors on target cells and elicit downstream chemical changes upon binding; and iii) be regulated in the temporal and/or spatial domain (Malandrino and Smith, 2018). However, there have been increased reports of peptides that perform a signaling role but do not follow the above-stated rules. Here, we discuss a few examples that challenge this conventional view of a signaling peptide, the therapeutic potential of these examples, and the role of MS in better understanding their biosynthetic routes.

A. Nonclassic Neuropeptides and Cytosolic Peptides

Canonically, neuropeptides are derived from larger precursor proteins, packaged into vesicles, and released upon stimulation; however, there is now substantial evidence for bioactive peptides that do not follow this route; e.g., peptides derived from angiotensinogen by the action of angiotensin-converting enzyme (Braun-Menendez and Page, 1958). Since their discovery, other nonclassic, bioactive cell-cell signaling peptides have been uncovered, such as hemorphins (Brantl et al., 1986), hemopressins (Rioli et al., 2003), hippocampal cholinergic neurostimulating peptide (Ojika et al., 1992), and diazepam binding inhibitor (Costa and Guidotti, 1991). Hemopressins (Heimann et al., 2021) and hemorphins (Ayoub and Vijayan, 2021) are of particular interest to pharmaceutics due to their antinociceptive and vasoactive properties, both of which are thought to originate from hemoglobin. The role of MS is especially important in the characterization of nonclassic peptides as they often cannot be predicted from the genome, or they share a conserved sequence motif that prevents differentiation of related peptides using affinity methods.

Hemopressin, a nonapeptide (PVNFKFLSH in rats, and PVNFKLLSH in humans and mice) derived from the a-chain of hemoglobin, has been shown to elicit dose-dependent nociceptive and antihypertensive behavior in mice (Blais et al., 2005). Importantly, hemopressin binds to the endocannabinoid receptor (Gomes et al., 2009), whose role in addiction, pain, memory, appetite, cognition, and behavioral response...
to reward and stress is well-established (Pagotto et al., 2006). In contrast to most other peptides, hemopressin maintains its bioactivity when administered orally and crosses the blood-brain barrier (Heimann et al., 2007). However, little is known about the mechanisms, localization, and timing of hemopressin (and its several extended forms) synthesis, transport, and regulation. The hypothesis is that hemopressin and hemopressin-related peptides result from proteasomal degradation of hemoglobin. Intriguingly, hemoglobin α-chain mRNA and protein have been found in non-blood cells, including neurons (Gomes et al., 2010), but it remains uncertain as to whether hemopressin is an endogenous mature peptide or a degradation product of a related extended form known as RVD-hemopressin (Bomar and Galande, 2013). MS could be an effective tool for systematic elucidation of the biosynthesis and action of hemopressin peptides to facilitate their therapeutic potential.

Hemorphins are a group of short peptides with a common YPWT motif, ranging in size from 4 to 10 residues, derived from the N-terminal of the β chain of the hemoglobin. They are known ligands for opioid receptors, and their serum and plasma levels have been shown to correlate with muscle fatigue (Glämsa et al., 1993), obesity (Maraninchi et al., 2013), and diabetes (Feron et al., 2009). Given the involvement of hemorphins in various pathophysiological processes, the inquiry into their biosynthesis and regulation has garnered greater interest in recent years. Though the exact mechanism of their regulation is not yet understood, work published by Feron and colleagues (Feron et al., 2009) has shown that plasma levels of LVV-hemorphin, a quintessential hemorphin, are correlated with levels of the enzymes cathepsin D and dipeptidyl peptidase IV. However, as is the case with hemorphins, further research is needed to understand hemorphin regulation, and MS becomes an important technique in this effort (Mielczarek et al., 2021).

B. Mysterious Endomorphins: Neuropeptides Without a Known Gene

The opioid-receptor system in mammals controls the pain and reward pathways. Several peptides in the mammalian central and peripheral nervous systems, such as β-endorphins, enkephalins, and dynorphin, have been shown to act as endogenous ligands toward the opioid receptors (Cesselin, 1991; Aldrich and McLaughlin, 2009). Additionally, nonpeptide alkaloid opiates, such as morphine are known to bind to the μ-opioid receptor with the highest affinity (Keith et al., 1996). Although morphine is a potent analgesic, repetitive usage leads to several undesirable side effects, such as tolerance, hyperalgesia, respiratory depression, and constipation. The quest to find alternate μ-opioid receptor ligands that can elicit the same analgesic effect as morphine without the side effects led to the discovery of endomorphins (EMs) (Zadina et al., 1997).

Originally isolated from the bovine frontal cortex, EMs are a set of two tetrapeptides (EM-1:YPWFα and EM-2: YPFFα) that are shown to be present in most mammalian species, including primates and rodents (Mizoguchi et al., 2013), and are the first known endogenous peptide-ligands to target the μ-opioid receptor. Their antinociceptive properties, combined with the absence of major side effects, generated several follow up studies that further explored the usefulness of EM-1/EM-2 and their analogs as potential therapeutic agents (Liu and Wang, 2012). One such analog developed by the Zadina group (codenamed ZH853) (Peenan and Zadina, 2019) showed great promise, with a superior analgesic effect compared with morphine along with reduced time spent in the chronic pain state accompanied by lack of latent sensitization. This superior analgesic property combined with the significant reduction in the above-mentioned side effects of ZH853 could potentially make it a transformative drug for treating chronic pain.

Even though EM and its analogs have advantages over conventional alkaloid-based analgesics such as morphine, the biosynthetic route for EMs is still unclear (Terskiy et al., 2007). For a protein/peptide synthesized via the conventional ribosomal pathway in a model organism, the genome should contain the corresponding protein-encoding gene. However, there is no evidence yet to support the presence of an EM-encoding gene. Currently, there are several hypotheses of the origin of EMs: (1) De Novo Synthesis of EMs from Smaller Constituent Peptide Building Blocks (Rónai et al., 2006) and (2) Biosynthesis of EM-1 from the Oxidative Modification of a Transcript-Encoding Analogous Peptide (Matsushima et al., 2019). Although these two hypotheses seem plausible, conclusive evidence can be obtained by leveraging the combined power of MS, cDNA cloning, and bioinformatics. Once confirmed, the hypothesis can further be validated by using isotopically labeled amino acids as the starting building blocks and scanning for the isotopically labeled EM-1 via high-sensitivity MS techniques (Prigge et al., 1997; Rónai et al., 2006; Dai et al., 2018; Matsushima et al., 2019).

Overall, while EMs are potent natural analgesics, a better understanding of their endogenous synthesis and regulation is needed to further elucidate their therapeutic potential.

C. Microproteins and Small ORF-Coded Peptides

The central dogma of molecular biology is the information flow from genes to protein (DNA → RNA → Protein). An open reading frame (ORF) of the mRNA has a set of codons that code for specific amino acids. The start codon, AUG, signals the beginning of
protein translation, and the stop codons (UAG, UAA, and UGA) signal its end. Since a long ORF typically has a higher chance of coding for a protein, most ORF-finding algorithms set a cutoff at 300 nucleotides (100 amino acids) for the lower limit of detection. Moreover, mRNA stretches that do not have start codons and/or lack evolutionary similarity to other known ORFs are also disregarded. Consequently, several stretches of RNA have been incorrectly classified as noncoding. However, recent studies have demonstrated that several of these small ORFs (smORFs) and “noncoding” RNA sequences are translated and thus encode peptides (Hsu and Benfey, 2018; Orr et al., 2020). These bioactive smORF-encoded polypeptides (SEPs), also called microproteins or microproteins, are produced directly from ribosomal translation instead of post-translational proteolytic processing of the precursor protein. Combinations of computational, genomic and proteomic approaches can be used to identify and characterize these peptides (Ma et al., 2016; Martinez et al., 2020).

Despite the discovery of several SEPs over the past few decades, their exact functions are often unclear. However, the functional roles of a few SEPs have been ascertained. Some of the first to be characterized are from the model organism Drosophila melanogaster. These SEPs, derived from a polycistronic tarsal-less(tal)-gene that was previously thought to be noncoding, have been shown to play a crucial role in the embryonic development of Drosophila (Galindo et al., 2007). In E. coli, an smORF called SgrS has been shown to encode a 43-residue SEP named SgrT, which inhibits glucose transport (Wadler and Vanderpool, 2007). More recently, a bioactive SEP named myoregulin was discovered in rodents and shown to regulate muscle relaxation by controlling Ca$^{2+}$ influx pumps (Anderson et al., 2016). SEPs have also been shown to be involved in DNA repair (Zhu et al., 2018), inhibition of tumorigenesis (Zhu et al., 2019), and regulation of endoplasmic reticulum stress (Chu et al., 2019). smORFs and SEPs represent a new and emerging area of focus in the study of bioactive signaling peptides that are still largely unexplored (Saghatelian and Couso, 2015; Jackson et al., 2018). With the latest advancements in bioinformatics, we can expect to see even more smROFs and SEPs being uncovered. This new class of bioactive peptides could provide deeper insights into eukaryotic peptide signaling mechanisms.

VI. The Rise of Biologics: Therapeutic Peptides

A. Chemical Basis of Peptide Therapeutics

Peptide and protein therapeutics have experienced significant advances in their clinical utility over the past several decades, with over 200 Food and Drug Administration (FDA) approvals in the United States (Fosgerau and Hoffmann, 2015; Usmani et al., 2017). Monoclonal antibodies and soluble portions of receptors make up an increasing share of the therapeutic drug market (Elgundi et al., 2017; Urquhart, 2019). In addition to antibodies and receptor fragments, hormone-like peptides are attractive options due to their high affinity and specificity for protein targets, and the ability to use endogenous peptides as starting points for active drug molecules. The power of hormone biology as therapeutic molecules is nicely demonstrated by insulin. Pancreatic extracts containing insulin were first isolated and used to treat diabetes nearly 100 years ago (Banting et al., 1922); exogenous insulin therapy has proven essential for patients with diabetes, saving countless lives worldwide. In addition to insulin, other cell-cell signaling peptides, including oxytocin, vasopressin, and calcitonin, have been used therapeutically in their unaltered forms (Lau and Dunn, 2018). In some cases, bioactive fragments of endogenous peptide hormones can also be used therapeutically. For example, teriparatide, the N-terminal region of parathyroid hormone, increases bone mineral density and is used to treat osteoporosis (Orwell et al., 2003; Bri xen et al., 2004).

In many cases, nature has taken advantage of natural hormone-receptor signaling systems in the evolutionary development of toxins. Many components that make up the venom of diverse animals are prohormone-derived peptides, and several of these compounds have found successful therapeutic use. For example, peptides from cone snail venom (conotoxins) have proven to have remarkably diverse biologic activities (Lewis et al., 2012), and the evolutionary diversity from over 700 cone snail species provides an immense source of natural products for potential therapeutic use (Olivera, 2006; Gao et al., 2017). Demonstrating their potential, the conotoxin ziconotide, which inhibits N-type calcium ion channels, has been FDA-approved for the treatment of chronic pain (Garber, 2005). Another notable example is that of the GLP-1 receptor agonist exenatide, a peptide isolated from the venom of the Gila monster (Heloderma suspectum), which has shown success in the treatment of diabetes (Drucker and Nauck, 2006; Il tz et al., 2006).

B. Improving Stability and Function: Amino Acid Isomerization, Peptide Conjugates

Despite the numerous advantages of naturally derived peptide hormones in biomolecular interactions, unaltered peptides suffer from disadvantages that can limit their utility as drugs (Tang et al., 2004). Most notably, peptides often exhibit short halflives in blood due to a combination of proteolytic degradation and renal filtration. In addition, the high susceptibility of peptideic molecules to protease degradation in the stomach and low efficiency of absorption...
often lead to limited oral bioavailability for most peptides (Aguirre et al., 2016; Drucker, 2020). As a result, most peptide drugs must be injected intravenously, which is less desirable than oral administration for many patients.

A number of strategies have been employed to increase the plasma half-life of peptides for use as drugs (Fosgerau and Hoffmann, 2015). In some cases, minor modifications to a peptide’s primary amino acid sequence can be beneficial, especially for reducing cleavage by peptidases. Desmopressin, an analog of vasopressin lacking the N-terminal amine group and bearing a D-Arg substitution (Mannucci et al., 1977), is used as an antidiuretic and in the management of bleeding disorders, such as von Willebrand disease (Nichols et al., 2008; Cornu et al., 2012; Sharma and Flood, 2017). Similarly, afamelanotide is an analog of α-MSH, with engineered norleucine and D-phenylalanine substitutions approved to treat erythropoietic protoporphyria (Sawyer et al., 1980; Langendonk et al., 2015). In some cases, more drastic modifications from the primary sequence, including the addition of non-natural residues, have been advantageous. Octreotide, a modified analog of an active somatostatin fragment incorporating two D-amino acid residue substitutions, is much more potent and longer acting in vivo than native somatostatin (Bauer et al., 1982), and has shown efficacy in the treatment of acromegaly and certain neuroendocrine tumors (Lamberts and Hofland, 2019). The osteoporosis therapeutic abaloparatide, an analog of parathyroid hormone-related protein, contains eight amino acid residue substitutions to improve stability, including the incorporation of a non-natural aminoisobutyric acid residue (Hattersley et al., 2016).

Covalent conjugation of peptides to larger molecular weight species, such as polyethylene glycol, can increase plasma half-life, primarily by avoiding renal clearance due to the increase in hydrodynamic radius and some receptor-mediated mechanisms (Kontermann, 2011; Cavaco et al., 2017; Lau and Dunn, 2018). As prominent examples, the diabetes therapeutics liraglutide and semaglutide are analogs of GLP-1 conjugated to an extended fatty acid, which greatly enhances plasma half-life through noncovalent association to albumin (Knudsen et al., 2000; Drucker and Nauck, 2006; Lau et al., 2015). Several other GLP-1 analog conjugates, including dulaglutide (conjugated to Fc fragment) (Glaesner et al., 2010) and albigitide (conjugated to human serum albumin) (Baggio et al., 2004), have also been developed and show favorable pharmaceutical properties.

Progress has been made in developing hormone analogs for oral administration, with numerous clinical trials often relying on a combination of half-life-extending substitutions and careful formulations to protect the peptide and increase absorption (Aguirre et al., 2016; Drucker, 2020; Kong et al., 2020). At present, desmopressin (vasopressin analog) is currently one of a few orally available peptides, and its success is likely a result of its high stability as well as its high potency, which results in therapeutic efficacy despite relatively poor absorption. As an exciting recent advancement, in 2019 the FDA approved a semaglutide oral formulation for use in diabetes treatment (Aguirre et al., 2016; Drucker, 2020). Formulations of several peptides, including calcitonin, desmopressin, oxytocin, and analogs of gonadotropin-releasing hormone, have been approved for intranasal delivery (Al Bakri et al., 2018).

C. The Future of Peptide Therapeutics

The tremendous impact of small molecules in modulating neurotransmitter systems for therapy suggests that there is considerable opportunity for peptide transmitters to serve as selective modulators of cell-cell signaling pathways in the central nervous system; however, an additional challenge is the inefficiency of most peptides in spontaneously crossing the blood-brain barrier (Salameh and Banks, 2014). Ziconotide, an FDA-approved inhibitor of N-type calcium channels, must be injected through a pump implanted in the spinal cord to reach its molecular targets (Garber, 2005). In spite of these challenges, some peptides can cross the blood-brain barrier through various mechanisms (Van Dorpe et al., 2012; Banks, 2015). In particular, receptor-mediated transport can occur for a number of peptides, and preclinical and clinical studies using peptide-based shuttles acting through this mechanism have shown promise (Demeule et al., 2008; Oller-Salvia et al., 2016). Finally, intranasal administration has shown some promise as a route of administration for peptides as a direct route to bypass the blood-brain barrier (Lochhead and Thorne, 2012; Meredith et al., 2015; Al Bakri et al., 2018).

VII. Conclusions

Despite its relatively short history, MS-based neuropeptide discovery and characterization has been applied to the exploration of neuropeptide mechanisms of human disease processes and provided novel neuropeptide targets for drug development. Monitoring neuropeptides as biomarkers of pathology or drug responses has become instrumental in the expansion of translational research. It is already well established that neuropeptide-based therapies offer significant advantages over existing small-molecule drugs as they are more specific and seem to have few or no adverse effects. We believe that the future for medicinal neuropeptides is exciting for several reasons: (1) a substantial increase in our understanding of neuropeptide roles and routes of action has been achieved through the use of validated animal models and in vitro
research; (2) identification of neuropeptides powered by MS has led to unambiguous characterization of peptide complements at all organismal levels, from subcellular components to individual cells, cellular populations, tissue, organs, and entire organisms; (3) methods of rapid screening for receptors facilitates deorphanizing GPCRs with the increasing number of identified endogenous peptides. The contribution of MS-based research is synergistic to discoveries using the continuously advancing molecular toolsets and is becoming indispensable for drug discovery and development.

Authorship Contributions
Wrote or contributed to the writing of the manuscript: Anapindi, Checco, Romanova, Sweeder.

References


Mass Spectrometry for Neuropeptide Discovery


rodent hypothalamus using a rapid conductive sample heating system. Analyst (Lond) 142:4476–4485.


