



A comprehensive review on neuropeptides: databases and computational tools

Wei Xu^{1,3,#}, Lei Wang^{2,#}, Mingxia Wang^{2,3,#}, Jingkang Jiang², Zhidong Xue^{2,3,*}, Yan Wang^{1,3,*}

¹School of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, Hubei, China.

²School of Software Engineering, Huazhong University of Science and Technology, Wuhan 430074, Hubei, China.

³Institute of Medical Artificial Intelligence, Binzhou Medical University, Yantai 264003, Shandong, China.

#Authors contributed equally.

***Correspondence to:** Zhidong Xue, School of Software Engineering, Huazhong University of Science and Technology, Wuhan 430074, Hubei, China. E-mail: zdxue@hust.edu.cn

Yan Wang, School of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, Hubei, China. E-mail: yanw@hust.edu.cn

Received: December 31, 2024 **Accepted:** April 07, 2025 **Published:** April 10, 2025

Cite this article: Xu W, Wang L, Wang M, Jiang J, Xue Z, Wang Y. A comprehensive review on neuropeptides: databases and computational tools. *Comput Biomed.* 2026;1:202402. <https://doi.org/10.70401/cbm.2025.0001>

Abstract

Neuropeptides are crucial signaling molecules that regulate diverse physiological processes spanning growth, social behavior, learning, memory, metabolism, homeostasis, reproduction, and neural differentiation across both nervous and peripheral systems. Dysregulation of neuropeptides signaling is closely linked to various pathological conditions, such as neurological disorders, metabolic diseases, cardiovascular conditions, and even cancer, positioning them as potential therapeutic agents or targets for intervention. In recent years, research into neuropeptides has accelerated, with vast amounts of data continuously accumulating in multiple databases. However, the study of neuropeptides is often impeded by the need for extensive and time-consuming experimental investigations. As a result, computational tools have become essential for the rapid, large-scale identification of neuropeptides. This review systematically discusses neuropeptide-related databases and computational tools. These databases organize extensive data on neuropeptide sequences, structures, and functions. Among these, NeuroPep2.0, with 11,417 neuropeptide entries, is currently the most widely used dataset for neuropeptide prediction. Additionally, this review explores the application of computational approaches in neuropeptide prediction. While early methods predominantly relied on homologous sequence alignment and biochemical feature statistics, recent advances in machine learning have significantly enhanced prediction accuracy and efficiency. Tools such as NeuroPred-PLM and DeepNeuroPred, developed by our research group using protein language models, have substantially improved prediction performance. In conclusion, this review provides a comprehensive overview of current neuropeptide databases and computational tools, offering researchers a thorough survey of available resources and analytical methods, and emphasizing the necessity of continuous optimization to advance neuropeptide research and its therapeutic applications.

Keywords: Neuropeptides, databases, computational tools, machine learning, deep learning, prediction models

1. Introduction

1.1 Neuropeptides

Neuropeptides are signaling molecules composed of amino acid chains, synthesized and released by neurons, playing a pivotal role in modulating signals within the nervous system^[1]. They not only facilitate direct signal transmission but also regulate the release and activity of other neurotransmitters. As a result, neuropeptides influence a variety of physiological processes, including emotion, behavior, pain perception, stress responses, appetite, learning, memory, skeletal homeostasis, and metabolism^[2,3]. Unlike classical neurotransmitters such as acetylcholine and dopamine, neuropeptides are larger molecules with more complex structures, enabling them to convey richer chemical information and interact with a broader range of recognition sites. This structural complexity allows for more precise signal transmission and regulation. Furthermore, neuropeptides can be released not only at synaptic sites but also at various other locations outside of synaptic specializations, where they can diffuse over significant distances to exert their effects through G protein-coupled receptors (GPCRs)^[4]. While neuropeptide diffusion and binding occur more slowly than that of classical neurotransmitters, their interactions with receptors are typically more robust, leading to sustained and long-lasting regulatory effects.



The identification of neuropeptides dates back to the 19th century, when biological activity was observed in extracts derived from the brain tissues of various animals^[5]. In 1905, physiologist Ernest Starling coined the term “hormone” to describe chemical messengers such as vasopressin (VP) and oxytocin (OT), which are now classified as neuropeptides^[5,6]. By the 1950s, it was established that many hormones are peptide chains^[7]. Extensive extraction and purification efforts led to the discovery of additional peptides in brain regions such as the hypothalamus and brainstem, emphasizing their roles in behavior and memory^[8,9]. Starting in the 1950s, de Wied’s pioneering research demonstrated that adrenocorticotrophic hormone (ACTH), melanocyte stimulating hormone, and VP influence learning and memory^[10,11]. In the 1970s, he introduced the term “neuropeptide” to describe hormone-like peptides and their fragments that exhibit neural activity^[12]. The application of techniques such as mass spectrometry, immunostaining, and radiolabeled ligand binding has been crucial in exploring the distribution, synthesis pathways, and functional mechanisms of neuropeptides^[13-15]. In the 21st century, advancements in molecular biology, genomics, proteomics, and single-cell analysis have made neuropeptide research more precise and comprehensive^[16-18].

Neuropeptides play a dual role in the regulation and transmission signals within the nervous system, while also fulfilling a diverse range of functions throughout the body^[19]. They regulate emotional and behavioral processes^[20], pain perception^[21,22], as well as learning and memory^[23]. Neuropeptides are also essential for metabolic regulation, particularly in appetite control and energy homeostasis^[24]. Additionally, they contribute to neuroprotection and nerve regeneration^[25,26]. Dysregulation or dysfunction of neuropeptides is often associated with various neurological disorders, including Alzheimer’s disease, Parkinson’s disease, depression, anxiety, and chronic pain^[27-30]. Furthermore, neuropeptide dysregulation can negatively impact metabolic and endocrine functions, contributing to conditions such as obesity, metabolic syndrome, and diabetes^[31,32]. Moreover, neuropeptide dysregulation has been linked to significant cardiovascular disorders^[33]. Neuropeptides have also been implicated in tumorigenesis, including in breast cancer and neuroblastoma^[3,34].

Research on neuropeptides has substantially advanced our understanding of the nervous system and holds broad applications in biomedicine and agriculture. Neuropeptides and their receptors have emerged as promising targets for drug development, with over 80 neuropeptide-related drugs by the U.S. Food and Drug Administration for clinical use^[35]. Various neuropeptides, along with their analogs or receptor agonists, show potential in treating a range of diseases, including autism, depression, anxiety, diabetes, Alzheimer’s disease, stroke, and Parkinson’s disease^[32,36-41]. Neuropeptides also show promise in pest control by regulating insect growth and behavior, offering environmentally friendly alternatives^[42]. As a key area in neuroscience, clinical research, and agriculture, neuropeptides present vast potential for future research and practical applications.

1.2 Biosynthesis and release of neuropeptides

Neuropeptides are synthesized and released by neurons initially as inactive precursor proteins, referred to as prepropeptides or prohormones. These precursors undergo a series of cleavage processes and post-translational modifications to generate biologically active, mature peptides^[43]. Prepropeptides contain several elements, including a signal peptide, the neuropeptide sequence, spacer peptides, and cleavage sites^[44]. Following their synthesis on ribosomes associated with the rough endoplasmic reticulum, the signal peptide directs the prepropeptides into the endoplasmic reticulum, where it is subsequently cleaved. Initial folding and glycosylation occur within the endoplasmic reticulum, resulting in the formation of propeptides. These propeptides are then transported to the Golgi apparatus, where they are cleaved into short-chain, mature neuropeptides by endopeptidases and carboxypeptidases. Additional chemical modifications, such as glycosylation, sulfonation, and methylation, further refine their structure and functionality. The processed neuropeptides are packaged into dense-core vesicles for storage and rapid release when required. During storage, further modifications may occur, including phosphorylation, N-terminal pyroglutamate formation, and C-terminal amidation, which finely regulate their activity upon release. Dense-core vesicles are transported throughout the neuron, enabling the release of neuropeptides from the synaptic cleft, cell body, or axons as needed.

Upon neuronal stimulation, neuropeptides stored in dense-core vesicles are released into the synaptic cleft or extracellular space via exocytosis^[4], often simultaneously with classical neurotransmitters. However, their release mechanisms differ: neuropeptides require a lower concentration of Ca^{2+} and are released at greater distances from Ca^{2+} entry sites compared to classical neurotransmitters. This allows neuropeptides to act on distant receptors through a process known as volume transmission, leading to broader and more prolonged regulatory effects. Unlike classical neurotransmitters, neuropeptides lack a reuptake mechanism and are degraded more slowly by peptidases, which further contributing to their sustained activity.

The process of neuropeptide synthesis and release involves multiple steps and complex physiological mechanisms, making the identification of neuropeptides quite challenging. As described in the following sections, computational methods for neuropeptides are primarily divided into precursor prediction, mature neuropeptide prediction, and neuropeptide precursor cleavage site prediction. These methods are categorized based on the key stages in the synthesis and release of neuropeptides, which facilitates a better understanding and identification of these molecules.

1.3 Diversity of neuropeptides

Despite the limited number of genes encoding neuropeptides, the diversity of neuropeptides produced is exceptionally high^[35]. This

diversity arises from the utilization of multiple mechanisms during neuropeptides synthesis^[45]. It also explains the presence of many similar sequences in neuropeptide databases and underscores the need for computational prediction methods to consider these features.

Typically, proteins are synthesized from messenger RNAs (mRNAs), which are derived from precursor mRNAs (pre-mRNAs) through a series of cellular processes, including splicing, capping, and polyadenylation within the nucleus. One significant mechanism employed by neurons to enhance neuropeptide diversity is alternative splicing. For example, the calcitonin gene can undergo variable splicing to produce mRNAs that encode the mature calcitonin peptide (pCal) and the calcitonin gene-related peptides (pCGRP1 and pCGRP2)^[46].

As previously noted, proteolytic processing occurs within the Golgi apparatus and dense-core vesicles, where precursor peptides are cleaved at specific sites. This selective cleavage, along with post-translational modifications, leads to the generation of a wide array of neuropeptides. A neuropeptide precursor containing a single bioactive peptide can undergo various post-translational modifications, resulting in the production of neuropeptides of differing lengths that retain identical C-terminal or N-terminal sequences^[47]. For instance, the cholecystokinin (CCK) gene encodes a propeptide that is enzymatically cleaved into six CCK peptides of varying lengths (83 to 8 amino acids), all sharing a common C-terminal bioactive octapeptide sequence. Similarly, the secretin gene produces three bioactive peptides (secretin-27, -28, and -30) through variable cleavage and trimming.

Selective precursor cleavage also involves precursors containing multiple distinct neuropeptides. Alternative splicing can alter cleavage sites, influencing which neuropeptide is ultimately released. Furthermore, dense-core vesicles may contain proteases that recognize different cleavage sites or generate diversity through post-translational modifications. For example, carbohydrate side chains can spatially hinder access to specific cleavage sites, preventing proteolysis. The pro-opiomelanocortin precursor protein illustrates this phenomenon, as it can be cleaved to produce ACTH and β -lipotropin, which can further γ -lipotropin and β -endorphin. Depending on the specific processing events, the same precursor protein can be modified to produce neuropeptides with markedly different biological activities^[48].

1.4 Mechanisms of neuropeptide action

Neuropeptides play a crucial role in regulatory processes by binding to specific receptors, thereby forming a complex network of receptor-mediated signaling pathways. The receptors are primarily located on the cell membrane and are widely distributed throughout the nervous system and various other tissues. Most neuropeptide receptors are GPCRs, which are classified into two main families: the rhodopsin-like family and the secretin family. Typically, a neuropeptide activates a specific GPCR; however, certain neuropeptides have the capacity to activate multiple GPCRs^[49]. For instance, neuropeptide Y (NPY) activates several GPCRs (e.g., Y1R, Y2R, Y4R, and Y5R)^[39], leading to diverse physiological responses such as appetite regulation, anxiety relief, and tumor metastasis^[50-52]. Additionally, some GPCRs can be activated by various neuropeptides. For instance, both VP and OT interact with the vasopressin receptor, influencing vascular contraction and social behaviors^[53,54]. Beyond GPCRs, neuropeptides can also target other receptor types, including ligand-gated ion channels, transmembrane receptor tyrosine kinases, and guanylate cyclase receptors. For example, the FMRF-amide-gated sodium channel (FaNaC) responds to neuropeptides such as FMRF-amide in invertebrates, leading to membrane depolarization and regulating functions like muscle contraction and feeding behavior^[55]. Similarly, insulin-like peptides in *Drosophila* and atrial natriuretic peptide in mammals interact with their respective receptors to facilitate signal transduction^[56,57].

Upon binding of a neuropeptide to its receptor, the receptor undergoes a conformational change, thereby activating downstream signaling cascades that increase levels of second messengers such as cAMP, IP₃, cGMP, Na⁺, tyrosine kinases, and others^[58]. These messengers activate multiple pathways, enabling cells to modulate neuronal excitability, hormone secretion, or gene expression. Furthermore, certain neuropeptide receptors can activate multiple pathways or interact with receptors for other neurotransmitters. For example, NPY receptors can synergize with gamma-aminobutyric acid (GABA) receptors to regulate mood and appetite, while VIP receptors can modulate cAMP or IP₃ signaling pathways^[59,60].

The number of receptors on the cell surface is dynamic, regulated by mechanisms such as downregulation, desensitization, internalization, and feedback inhibition regulating cellular responsiveness^[61]. For instance, prolonged opioid use leads to receptor downregulation, contributing to addiction and tolerance. Receptor desensitization and internalization help prevent overstimulation, while feedback inhibition plays a key role in maintaining homeostasis. Positive feedback mechanisms can also amplify responses in specific contexts, such as during stress when corticotropin-releasing factor (CRF) enhances receptor sensitivity^[62].

Neurotransmitters and neuropeptides often collaborate to form a multi-level regulatory network^[4,63,64]. While neurotransmitters such as glutamate, GABA, and acetylcholine mediate rapid, transient signals, neuropeptides exert slower, more prolonged effects through diffuse signaling. This complementary relationship enhances the regulatory capacity and flexibility of the nervous system. For example, during stress, CRF works with norepinephrine and serotonin to enhance the nervous system's adaptive response. In pain signaling, glutamate mediates rapid pain transmission, while substance P regulates pain persistence and intensity. This coordinated mechanism is essential for maintaining the dynamic balance of the nervous system and regulating complex physiological and psychological functions.

The mechanism of action of neuropeptides is not only the intrinsic basis for their ability to regulate physiological processes but also a key factor in predicting their effects and developing drugs. Understanding these mechanisms is essential for uncovering the roles of neuropeptides in health and disease, as well as for designing targeted therapeutic interventions.

1.5 Neuropeptide detection methods

In 1931, Euler discovered the first neuropeptide, substance P (SP), but its sequence was not determined until four decades later^[65]. Currently, researchers have cataloged thousands of distinct neuropeptides; however, the vast diversity within this class of biomolecules suggests that many remain unidentified. Consequently, ongoing investigations aim to discover and characterize new neuropeptides through various detection methodologies, thereby elucidating their essential biological functions.

Neuropeptide detection typically involves separation and purification techniques such as chromatography, capillary electrophoresis, and capillary electrochromatography, followed by identification using methods like UV absorbance, fluorescence, electrochemical detection, or antibody-based assays (e.g., radioimmunoassays, ELISA, and immunohistochemistry)^[66-68]. However, these methods have limitations, including potential cross-reactivity due to structural similarities among neuropeptides. Additionally, RNA-based techniques, such as Northern blotting and in situ hybridization, provide indirect measures of neuropeptide expression but are often time-consuming and unable to differentiate between similar peptides^[69].

Historically, biological assays were used to identify neuropeptides based on their activity in cellular or tissue contexts. More recently, biochemical characteristics (e.g., C-terminal amidation and precursor protein cleavage sites) have been utilized to guide neuropeptide discovery^[70]. Mass spectrometry (MS) has emerged as a pivotal tool for identifying unknown neuropeptides, offering high sensitivity and accuracy. However, challenges related to sample size and throughput still persist.

The rapid advancement of omics technologies, particularly genomics and transcriptomics, has generated extensive datasets for neuropeptide research. Bioinformatics tools, including genome annotation, gene prediction, and sequence alignment (e.g., BLAST and HMMER), facilitate the identification of potential neuropeptide-encoding genes by comparing nucleic acid sequences with known neuropeptides in databases like UniProt^[71] and NeuroPep^[72]. For instance, Koziol *et al.* used comparative genomics to identify 34 neuropeptide families from the genomes of parasitic flatworms^[73]. However, reliance on nucleic acid homology alone may overlook neuropeptides with complex post-translational modifications.

Peptidomics, particularly when coupled with MS, is a powerful approach for neuropeptide analysis, enabling both quantification and identification of unknown peptides. Two main approaches are employed: matching mass spectra against established databases, or utilizing de novo sequencing algorithms (e.g., PEAKS) to analyze MS data and identify peptides based on characteristic features, including signal peptides, cleavage sites, and evolutionary conservation^[17,74,75].

Despite their considerable potential in medical applications, the clinical use of neuropeptides remains limited due to their low endogenous concentrations and high susceptibility to enzymatic degradation. Traditional detection techniques often struggle with issues of sensitivity and specificity, and large-scale detection remains costly and technically challenging. Consequently, there is an urgent need for more efficient, sensitive, and scalable neuropeptide detection technologies.

To overcome these limitations, computational approaches have been developed to predict and design potential neuropeptides. These methods rely on neuropeptide databases that validate and support the accuracy of these predictions.

This article aims to provide a comprehensive overview of the databases and computational tools relevant to neuropeptide research, aiming to support the development of neuropeptide-based drugs. The structure of the paper is as follows: Section 1 outlines the biosynthesis, release, and mechanisms of action of neuropeptides; Section 2 summarizes current neuropeptide databases; Section 3 reviews prediction methodologies based on machine learning or deep learning; and Section 4 concludes the paper and outlines future research directions, as shown in [Figure 1](#).

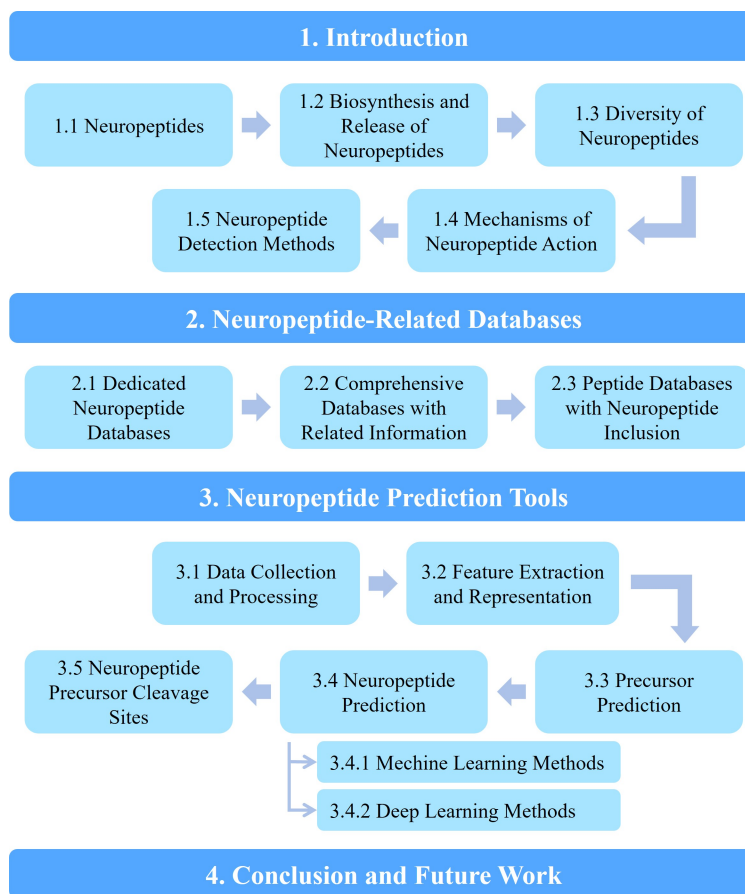


Figure 1. The organization of the study.

2. Neuropeptide-Related Databases

Neuropeptides are small, protein-like molecules that are integral to a variety of physiological functions. Neuropeptide databases serve as essential resources, providing extensive information that enables researchers to investigate a wide range of neuropeptide-related data. [Table 1](#) and [Table 2](#) enumerate these databases. Neuropeptide databases can be primarily classified into three categories: dedicated neuropeptide databases, comprehensive databases that include related information, and peptide databases that contain neuropeptide data. Dedicated neuropeptide databases focus exclusively on neuropeptide-related content, offering detailed and specific insights. In contrast, comprehensive databases, such as UniProt and GproteinDb^[76], incorporate neuropeptide information within a broader framework of proteins and associated biological elements. Additionally, bioactive peptide databases include neuropeptides as part of the broader bioactive peptide family, providing valuable information on their bioactivities and interrelationships with other peptides.

Table 1. A description of existing dedicated neuropeptide databases.

Database Name	Years	Species	Size	Sources	URL
NeuroPep 2.0	2024	924	11,417	UniProt, Literatures, NeuroPep	http://isyslab.info/NeuroPepV2/
DINeR	2017	425	4,782	NCBI, Literatures	https://www.neurostresspep.eu/diner/
NeuroPep	2015	493	5,945	UniProt, Literatures, www.neuropeptides.nl, Neuropedia	http://isyslab.info/NeuroPep/
NeuroPedia	2011	8	847	Mass Spectrometry Experiments	http://proteomics.ucsd.edu/Software/NeuroPedia.html
www.neuropeptides.nl	2010	-	-	NCBI, UCSC, Literatures	http://www.neuropeptides.nl/

DINeR: Database for Insect Neuropeptide Research; NCBI: National Center for Biotechnology Information; UCSC: University of California, Santa Cruz.

Table 2. A description of existing databases with neuropeptide inclusion.

Database Name	Years	Number of NPs	Size	Type	URL
IUPHAR/BPS guide to PHARMACOLOGY	2024	-	15,202	protein targets and ligand molecules	https://www.guidetopharmacology.org/
UniProt	2024	6,328	248,838,887	protein	https://www.uniprot.org/
GproteinDb	2024	-	333	G proteins	https://gproteindb.org/
Adult Zebrafish Brain Gene Expression Database	2024	38	38	brain gene expression	https://ssbd.riken.jp/azebex/
aSynPEP-DB	2023	35	123	biogenic peptides for inhibiting the aggregation of alpha-synuclein	https://asynpepdb.ppmclab.com/
BIOPEP-UWM	2021	-	5,098	bioactive peptide	https://biochemia.uwm.edu.pl/
EROP-Moscow	2020	-	26,750	endogenous regulatory oligopeptides	http://erop.inbi.ras.ru
StraPep	2018	39	3791	bioactive peptide	http://isyslab.info/StraPep/
SATPdb	2015	-	19,192	therapeutic peptide	http://crdd.osdd.net/raghava/satpdb/
Bioactive Peptide Database in Metazoa	2008	-	20,027	bioactive peptide	http://www.peptides.be/
Pepbank	2007	-	19,792	peptide	http://pepbank.mgh.harvard.edu/
SwePep	2006	-	4,180	endogenous peptide	www.swepep.org .
Database of biologically active peptide sequences	1999	90	527	bioactive peptide	-

NPs: neuropeptides.

2.1 Dedicated neuropeptide databases

The NeuroPep 2.0 database currently contains 11,417 unique neuropeptide entries, representing 924 organisms and spanning 81 distinct neuropeptide families^[35]. This marks nearly a twofold increase in the number of neuropeptides compared to the initial version of NeuroPep. As of now, it is the most comprehensive neuropeptide database available. All neuropeptides included in the NeuroPep database have been experimentally validated, with sources primarily derived from the MEDLINE and UniProtKB databases. In addition to detailed annotations for each entry, NeuroPep 2.0 predicts the three-dimensional structures of neuropeptides for which experimental data is lacking. The database also features a user-friendly interface and integrates several tools, including two neuropeptide discovery tools: DeepNeuropePred and NeuroPred-PLM^[35]. Currently, both Neuropep and Neuropep 2.0 are the most widely used dataset for neuropeptide prediction.

The Database for Insect Neuropeptide Research (DINeR) is a specialized resource focused on the search and retrieval of neuropeptide information specific to insects^[44]. It contains 4,782 records of insect neuropeptides from 50 neuropeptide families across 425 different insect species. DINeR provides detailed information on isoform sequences, physiological functions, images of receptor-binding sites, and comprehensive summaries for each neuropeptide family^[44].

NeuroPedia is another specialized neuropeptide database and mass spectrometry (MS/MS) spectral library, containing 847 neuropeptide sequences from eight organisms, all belonging to the phylum Chordata, with the exception of the medicinal leech. Additionally, NeuroPedia includes 3,401 identified spectra from various spectral libraries and provides annotated spectrum images for each library spectrum^[77].

The [Neuropeptides Database](#) serves as a single-table resource for neuropeptide gene families within mammalian genomes. It provides information on approximately 90 genes encoding both classical and candidate neuropeptides. For each gene, the database includes the gene symbol, chromosomal localization, precursors, active peptides, and brain expression data. Additionally, it offers hyperlinks to related bioinformatics databases containing information on genomes, transcripts, protein structures, and brain expression^[78].

2.2 Comprehensive databases with related information

UniProt is an extensive repository for protein sequence and annotation data, serving as a foundational resource for understanding the broader context of proteins associated with neuropeptides, although it is not exclusively focused on them^[74]. GproteinDb, an online research platform launched in 2022, specializes in G proteins—key signaling proteins involved in the action of pharmaceuticals, hormones, neurotransmitters, tastants, and odorants^[76]. While GproteinDb is not specifically dedicated to neuropeptides, it may provide relevant information on the signal transduction pathways in which neuropeptides participate are involved, considering the critical role of G proteins in numerous cellular signaling processes. The IUPHAR/BPS Guide to PHARMACOLOGY is an expert-curated resource cataloging pharmacological targets and their interacting substances^[79]. Although the database does not exhaustively detail neuropeptide-specific content, it is expected to contain significant valuable regarding the pharmacological properties and interactions of neuropeptides. Collectively, these three databases, despite their varied focuses, offer valuable insights into neuropeptides and are continuously updated.

2.3 Peptide databases with neuropeptide inclusion

As neuropeptides are classified as bioactive peptides, databases dedicated to bioactive peptides, such as BIOPEP-UWM, StraPep, and the Bioactive Peptide Database in Metazoa, also include neuropeptide data^[80-83]. Additionally, PepBank, a comprehensive peptide database, along with two endogenous peptide databases, EROP-Moscow and SwePep, also provide information on neuropeptides^[84-86]. Given the therapeutic potential of neuropeptides in addressing a variety of diseases, they are represented in SATPdb, a structurally annotated therapeutic peptide database that includes 19,192 unique experimentally validated sequences^[87]. Furthermore, two additional databases feature neuropeptide-related information: one focuses on neuropeptide gene expression in the adult zebrafish forebrain, while the other, aSynPEP-DB, is dedicated to biogenic peptides that inhibit α -synuclein aggregation^[88,89].

3. Neuropeptide Prediction Tools

3.1 Data collection and processing

In the field of neuropeptide prediction research, data collection and processing are essential steps in the development of predictive models. Researchers commonly rely on extensive databases and meticulously curated datasets to construct, train, and validate these models. The initial phase of most neuropeptide prediction projects involves the aggregation of datasets from neuropeptide-specific databases, as shown in Table 3. Prominent databases in this field include NeuroPep, NeuroPep 2.0, UniProt, SwissProt, and DIneR.

Table 3. Databases related to neuropeptide prediction tools.

Tools	Benchmark datasets	Evaluation
iNP_ESM	NeuroPep 2.0	10-fold cross-validation, ACC, MCC, AUC, F1, SEN, SPE, PRE, REC
NeuroPred-PLM	NeuroPep 2.0	10-fold cross-validation, ACC, MCC, F1, SEN, SPE, PRE, REC
DeepNeuropePred	UniProt	ACC, MCC, AUC, F1, PRE, REC, AUPR, FPR
PredNeuroP	NeuroPep and Swiss-Prot	10-fold cross-validation, ACC, MCC, AUC, SEN, SPE
NeuroPIpred	DIneR	5-fold cross-validation, ACC, MCC, AUC, SEN, SPE

ACC: accuracy; MCC: matthews correlation coefficient; AUC: area under the curve; F1: F1-score; SEN: sensitivity; SPE: specificity; PRE: precision; REC: recall; AUPR: area under the precision-recall curve; FPR: false positive rate.

Notably, NeuroPep and NeuroPep 2.0 are recognized as large-scale repositories of neuropeptide data and are frequently employed by various predictive tools, such as iNP_ESM^[90] and NeuroPred-PLM^[91]. These tools utilize the most current entries from NeuroPep 2.0 for training and the random selection of independent test sets, facilitating the generalization of models to novel data. Additionally, the UniProt and SwissProt databases provide comprehensive protein-related information, with tools like NeuroPred-ResSE^[92], NeuroCNN_GNB^[93], and Target-ensC_NP^[94] relying both NeuroPep and SwissProt for their training processes. Furthermore, models such as positive predictive value (PPV)^[95] and DeepNeuropePred^[96] leverage the resources from UniProt and SwissProt to enhance the accuracy and reliability of predictions. For example, DeepNeuropePred selected 31 precursor data entries from UniProt to establish an independent test set, ensuring a rigorous evaluation of model efficacy. Moreover, the DIneR database is utilized in studies like NeuroPIpred^[97], which incorporates unique methodologies for partitioning training and testing sets. Additionally, several other databases, including SATPdb, SwePep, APD3^[98], BioDADPep^[99], BIOPEP-UWM, and CancerPPD^[100], contribute valuable data resources for neuropeptide prediction.

Collectively, these datasets and databases serve as vital resources for advancing neuropeptide prediction models, facilitating systematic evaluations of model performance and driving progress in this research domain.

Upon collecting the preliminary data required for neuropeptide prediction from the database, it is crucial to proceed with the

necessary data processing. The data processing strategy can be outlined in several key steps. Initially, data partitioning plays a fundamental role in the development of predictive models. Typically, various tools adopt standardized data partitioning methodologies, which involve dividing the dataset into training, validation, and test subsets. These partitioning frameworks, such as the 8:1:1 or 8:2 ratios, are designed to optimize model training efficacy while ensuring the independence of test sets for subsequent evaluation.

Following data partitioning, the next phase involves the creation of independent test sets. These test sets are generated either through random sampling or by selecting distinct precursor data, as demonstrated by tools, such as NeuroPred-PLM and DeepNeuroPred. The primary goal of these independent test sets is to evaluate the model’s ability to generalize to previously unseen data.

Additionally, the construction of negative samples is particularly important in the context of NeuroPIpred. Researchers augment the model’s recognition accuracy by developing negative sample sets. These negative samples are created through similarity-based matching with positive peptides, thereby enhancing the model’s capacity to effectively distinguish between positive and negative samples.

3.2 Feature extraction and representation

In the field of neuropeptide prediction, the development of various analytical tools has significantly improved the accuracy of predictive models. These tools employ a variety of feature extraction and numerical encoding techniques to capture critical characteristics from protein or peptide sequences, thereby facilitating model analysis and enhancing predictive efficacy.

As shown in Table 4, some tools, such as NeuroPred-ResSE and MultiPep^[101], predominantly use one-hot encoding, a technique that converts each amino acid in a protein or peptide sequence into a binary feature vector. Although one-hot encoding is computationally efficient and easy to implement, it has limitations in representing interdependencies among amino acids within a sequence, which reduces its effectiveness in deep learning applications.

Table 4. Feature representation used in neuropeptide prediction tools.

Tools	Feature representation
NeuroPred-ResSE	one-hot coding, dipeptide deviation from expected mean and natural vector
MultiPep	one-hot coding
NeuroPpred-Fuse	AAC, DPC, GGAP, CTD, ASDC, PSAAC
PredNeuroP	AAC, DPC, binary profile-based features, AAIndex, GAAC, GDPC, CTD, AAE
NPpred (BERT-NeuroPred)	biLSTM, TCN, BERT
NeuroCNN_GNB	one-hot encoding, AAIndex, G-gap dipeptide encoding, word2vec
Target-ensC_NP	Bi-PSSM_DWT, KSB_DWT, one-hot encoding, SHAP Feature Selection
NeuroPIpred	physicochemical properties and structural features (AAC, DPC, binary profile feature extraction)

AAC: amino acid composition; DPC: dipeptide composition; GGAP: G-gap dipeptide composition; CTD: composition, transition, and distribution; ASDC: amphiphilic sum of distance to the center; PSAAC: pseudo amino acid composition; GDPC: generalized dipeptide composition; biLSTM: bidirectional long short-term memory; TCN: temporal convolutional networks; BERT: bidirectional encoder representations from transformers; SHAP: shapley additive explanations.

To address these limitations, tools such as NeuroPpred-Fuse^[102] and PredNeuroP^[103] employ amino acid composition (AAC) and dipeptide composition (DPC) methodologies. These approaches assess evaluate the frequency of amino acid occurrences within peptide sequences, thereby accurately reflecting their intrinsic chemical and structural properties, and serve as essential strategies for feature representation.

With the advancement of deep learning and language models, tools such as iNP_ESM leverage pre-trained language models, including evolutionary scale modeling (ESM) and UniRep. These models, trained on large-scale sequence datasets, are capable of capturing complex dependencies and semantic relationships among amino acid residues, thereby further improving the quality of feature representation.

For the extraction of higher-dimensional semantic information, tools such as NPpred (BERT-NeuroPred)^[104] and NeuroCNN_GNB utilize models like BERT and Word2Vec. These models embed peptide sequences into high-dimensional vector spaces, capturing rich semantic information and offering comprehensive feature representations. Additionally, NeuroCNN_GNB incorporates feature selection and structural feature extraction by using shapley additive explanations (SHAP) to quantify the contributions of individual features to model predictions, thereby enhancing the interpretability of model outcomes. Furthermore, techniques such as G-gap

dipeptide encoding and AAIndex are employed to extract the physical and chemical properties of peptides, enriching the representation of sequence features.

In terms of feature combination and optimization, tools such as Target-ensC_NP integrate multiple descriptors (e.g., Bi-PSSM_DWT and KSB_DWT) and utilize genetic algorithms to optimize feature combinations, thereby improving predictive performance. Similarly, NeuroPIpred combines features such as amino acid composition, dipeptide composition, and split composition to provide a comprehensive representation of the physicochemical properties of peptide sequences, offering robust feature support for neuropeptide prediction.

In conclusion, these tools utilize a wide range of feature extraction and encoding methodologies, including one-hot encoding, AAC/DPC, pre-trained language models, physicochemical descriptors, and feature optimization techniques. By capturing sequence features from multiple perspectives, these methods collectively contribute to the development of robust and accurate neuropeptide prediction models, driving continuous advancements in their reliability and precision.

3.3 Precursor prediction

As outlined in the introduction, neuropeptide precursors are unprocessed protein sequences that undergo specific post-translational modifications to produce bioactive neuropeptides. Typically, these precursors consist of signal peptide sequences, one or more neuropeptide segments, and recognition sites for processing enzymes. Through enzymatic cleavage and post-translational modifications such as glycosylation and phosphorylation, these precursors are processed to release and activate mature neuropeptides.

To enable the accurate identification of neuropeptide precursors and their functional components, various computational tools have been developed. These tools employ diverse methodologies and feature extraction techniques to analyze genomic or proteomic datasets, enabling the identification of precursor proteins and the prediction of essential sequence features.

As indicated in [Table 5](#), notable tools for identifying neuropeptide precursors and their functional components include NP-HMMer^[105], which use hidden markov models to search for specific neuropeptide families within protein databases. By constructing and applying family-specific models for 46 neuropeptide families, NP-HMMer effectively aligns sequences and identifies potential neuropeptide precursors within complex proteomes. Another tool, NeuroPP^[106], focuses on compositional features derived from peptide sequences, utilizing characteristics such as monopeptide, dipeptide, and tripeptide composition. It employs an ANOVA-based feature selection approach and using an SVM-based classification method, enabling the identification of significant sequence features critical for model training and prediction. Additionally, NeuroPID^[107] is a machine learning tool designed to identify neuropeptide precursors in metazoan proteomes. It extracts approximately 600 features from primary sequences, including biophysical, chemical, informational-statistical, and dibasic cleavage site characteristics. These features are then processed using SVM and ensemble decision tree classifiers to accurately identify neuropeptide precursors.

Table 5. Prediction tools.

Tools	Methods	ACC	PRE	REC	F1	MCC	AUC	AUPRC	Time
NP-HMMer	HMM	0.73	1.00	0.45	0.62	0.54	0.73	0.73	2.9
NeuroPP	SVM	0.84	0.82	0.87	0.84	0.68	0.84	0.78	-
NeuroPID	Ensemble	0.47	0.40	0.13	0.20	-0.09	0.47	0.49	152.92

ACC: accuracy; PRE: precision; REC: recall; F1: F1-score; MCC: matthews correlation coefficient; AUC: area under the curve; AUPRC: area under precision-recall curve.

To better compare the performance of these tools, we used precursor sequences from the test set in DeepNeuropePred and obtained the corresponding negative samples from UniProtKB following the NeuroPID method to construct a test set for comparison (Sequence ids listed in [Table S1](#)). The source codes for NP-HMMer and NeuroPID were obtained from their respective repositories ([NP-HMMer GitHub](#) and [NeuroPID GitHub](#)). After installation, both tools were executed to perform neuropeptide precursor predictions. Additionally, NeuroPP predictions were conducted using the web server at [NeuroPP Web Server](#). The prediction results are shown in the [Table 5](#). Overall, NeuroPP achieved the best performance, while NeuroPID performed relatively poorly, possibly due to significant differences between the test set and its training set. NeuroPID took 53 times longer than NP-HMMer, which required less than three seconds. NP-HMMer, relying on manually constructed HMM models, supports cross-species identification and offers high accuracy. However, it cannot predict unknown neuropeptides and is best suited for homologous sequence searches of known neuropeptides. NeuroPP can identify neuropeptides that meet specific sequence feature criteria but is less sensitive to atypical cleavage sites. NeuroPID, which utilizes a collection of machine learning models, exhibited relatively poor performance. Overall, NeuroPP is the most recommended tool due to its superior performance and user-friendly web server.

Moreover, alternative prediction methods have been explored, such as those utilizing local sequence conservation. These methods

rely on comparative genomics to identify conserved regions across related species, thereby inferring functional neuropeptide precursors^[73]. Similarly, binary logistic regression models have been employed to investigate the relationships between sequence features and precursor functionality, providing additional insights into neuropeptide precursor prediction^[108].

The advancement of these computational tools and methodologies has significantly improved the identification and functional annotation of neuropeptide precursors. By integrating machine learning techniques, compositional analyses, sequence conservation studies, and statistical models, these approaches enhance our understanding of neuropeptide precursor processing and improve prediction accuracy in genomic and proteomic research. Collectively, these tools and methodologies contribute to the advancement of knowledge on neuropeptide precursors, facilitating their identification and functional annotation in genomic and proteomic studies.

3.4 Neuropeptide prediction

The physiological responses of neuropeptides in cells are initiated and regulated through their binding to specific receptors. Computational tools designed for neuropeptide analysis aim to identify potential neuropeptides within protein sequences and optimize them for both research and therapeutic applications. These tools employ various methodologies, which can be broadly classified into the following categories:

3.4.1 Machine learning methods

Recent advancements in neuropeptides have led to the development of tools that incorporate a variety of machine learning models and feature encoding techniques to enhance both prediction accuracy and interpretability. As shown in Table 6, these tools utilize ensemble learning strategies and extensive datasets to achieve robust performance.

Table 6. Neuropeptide prediction based on machine learning methods.

Tools	Methods	Source code or online service website
iNP_ESM	SVM (SVM, GNB, LDA, LGBM, LR, KNN, LR)	Not provided
NeuroPred-FRL	SVM, RF, ERT, AB, NB, KNN	http://kurata14.bio.kyutech.ac.jp/NeuroPred-FRL/
PPV	Logistic Regression	https://github.com/jancr/ppv
Target-ensC_NP	Integration(XGB, ETC, SVM, ADA, FKNN, LGBM) genetic algorithm	https://github.com/shahidawkum/Target-ensC_NP
NeuroPpred-Fuse	Integration (RF, GBDT, XGBoost)	https://github.com/mingmingjiang1/NeuroPpred-Fuse

SVM: support vector machine; GNB: gaussian naive bayes; LDA: linear discriminant analysis; NB: naive bayes; RF: random forest; ERT: extremely randomized trees; XGB: XGBoost; LGBM: light gradient boosting machine; LR: logistic regression; KNN: K-nearest neighbors; ETC: extra trees classifier; ADA: AdaBoost; FKNN: fuzzy K-nearest neighbors; GBDT: gradient boosting decision tree; XGBoost: extreme gradient boosting.

Noteworthy examples include iNP_ESM and NeuroPred-FRL^[109], which implement several machine learning algorithms such as support vector machines (SVM), random forests (RF), naive bayes (NB), and extremely randomized trees. By integrating multiple feature representations, including ESM and UniRep, these tools improve predictive performance through model ensemble strategies.

Specifically, iNP_ESM employs large-scale language models (e.g., ESM and UniRep) to generate comprehensive sequence-based features. It applies several machine learning classifiers, including SVM, gaussian naive bayes (GNB), and linear discriminant analysis (LDA), for neuropeptide classification tasks. The performance of the model is assessed using 10-fold cross-validation, with evaluation metrics such as accuracy (ACC), F1 score, and matthews correlation coefficient (MCC). The NeuroPep 2.0 dataset is utilized as a benchmark for training and validation, ensuring the generalizability of the findings.

In a similar vein, NeuroPred-FRL combines various feature encoding techniques with machine learning methods such as SVM, RF, and ERT. To enhance the interpretability of the model, it employs SHAP, which quantifies the contribution of individual features to the predictions. The model's performance is rigorously evaluated using 5-fold cross-validation across multiple evaluation metrics, establishing a robust and interpretable prediction framework.

Additionally, PPV focuses on the analysis of peptidomic data obtained through mass spectrometry, utilizing logistic regression models to predict potential neuropeptide candidates. Other tools, such as Target-ensC_NP and NeuroPpred-Fuse, emphasize ensemble learning methodologies to enhance neuropeptide prediction. These tools integrate multiple machine learning models, including XGBoost (XGB) and SVM, while leveraging model diversity to improve robustness and overall predictive performance.

Machine learning methods in this section often integrate multiple algorithms to enhance prediction accuracy. These ensemble approaches are effective in improving both accuracy and robustness. However, they do have certain drawbacks. They may fail to

identify potential neuropeptides if the models do not capture novel or rare patterns. Additionally, the computational complexity is high, as training and integrating multiple models is time-consuming. These methods also heavily depend on high-quality features, and poor feature selection can significantly reduce performance. Moreover, their generalization ability is limited, as they may not perform well on data that differs substantially from the training data. Nevertheless, these machine learning methods are well-suited for analyzing large-scale neuropeptide datasets and applications that demand high accuracy and stability, such as drug development and disease diagnosis, particularly when rich feature information is available.

3.4.2 Deep learning methods

In recent years, substantial progress in deep learning and machine learning technologies has significantly advanced the field of neuropeptide prediction, resulting in the development of various robust and precise tools. Table 7 enumerates several of these deep learning instruments. Notably, NPpred (BERT-NeuroPred)^[104] utilizes sophisticated deep learning architectures, including BERT, bidirectional long short-term memory (biLSTM) networks, and temporal convolutional networks (TCN). By employing ANOVA and Tukey’s HSD methods for feature selection, this tool effectively identifies critical features, thereby enhancing the model’s predictive performance.

Table 7. Neuropeptide prediction based on deep learning methods.

Tools	Methods	Source code or online service website
NPpred (BERT-NeuroPred)	ANOVA, HSD	https://neuropred.anvil.app/
NeuroPred-PLM	Global multi-head attention network	https://github.com/ISYSLAB-HUST/NeuroPred-PLM
NPpred (NSGA-NeuroPred)	NSGA-II	https://neuropred.anvil.app/
NeuroPred-CLQ	Convolutional neural TCN and multi-head attention mechanism	https://github.com/GEHAH/NeuroPred-CLQ
PredNeuroP	Ensemble of ERT, ANN, KNN, XGBoost	https://github.com/xialab-ahu/PredNeuroP

TCN: temporal convolutional network; ERT: extremely randomized trees; ANOVA: analysis of variance; HSD: honestly significant difference; NSGA-II: non-dominated sorting genetic algorithm II; ANN: artificial neural network; KNN: k-nearest neighbors; XGBoost: eXtreme gradient boosting.

Building on this foundation, NeuroPred-PLM introduces an innovative methodology that incorporates ESM, multi-scale convolutional networks, and a global attention mechanism. This integration broadens the model’s capabilities, enabling it to focus on significant sequence patterns and contextual relationships, which in turn improves the accuracy of neuropeptide prediction. Conversely, NSGA-NeuroPred (NPpred) is specifically tailored for the discovery and optimization of novel neuropeptides. This tool combines traditional machine learning techniques with deep learning frameworks and integrates genetic algorithms and evolutionary strategies to perform multi-objective optimization, yielding Pareto-optimal solutions. Through crossover and mutation operations, it further refines peptide designs across multiple objectives, including sequence stability, functionality, and synthetic cost.

In addition to optimization and prediction tools, classification-oriented methods have also exhibited exceptional efficacy. For example, NeuroPred-CLQ^[110] employs one-hot encoding, supervised embeddings, and Word2Vec features, combined with convolutional neural networks (CNNs) and multi-head attention mechanisms, to accurately classify peptides as neuropeptides or non-neuropeptides. Lastly, PredNeuroP broadens the application scope of neuropeptide prediction by assessing the neuropeptide potential of specific peptides. This tool extracts a diverse array of amino acid and dipeptide features and utilizes an ensemble framework that incorporates algorithms such as ERT and artificial neural networks.

Collectively, these tools demonstrate distinct strengths and functionalities, contributing not only to the precise prediction and classification of neuropeptides but also to the overall advancement of the field. These deep learning-based methods can automatically extract high-dimensional features and capture complex patterns in data, significantly improving prediction accuracy. However, they require large amounts of labeled data for training. Deep learning methods are often criticized for their “black-box” nature, with the lack of interpretability being a significant drawback. Additionally, these methods demand considerable computational resources to train effectively. Researchers should carefully consider both the size of the dataset and the available computing resources when selecting a deep learning method for predicting neuropeptides.

By utilizing these classifications, researchers can select appropriate methodologies and tools that align with the required analytical perspectives and the characteristics of the data, thereby facilitating the identification and examination of neuropeptides. These methodologies offer solutions that vary in granularity, covering a broad range of applications, from basic statistical analysis to advanced modeling techniques.

3.5 Neuropeptide precursor cleavage sites

Neuropeptide precursor cleavage sites are specific locations within unprocessed protein chains where proteolytic or processing enzymes recognize and cleave particular amino acid sequences, resulting in the formation of active neuropeptides. The accurate prediction and validation of these cleavage sites are essential for elucidating the mechanisms of neuropeptide biosynthesis and their functional implications. The integration of computational methodologies with experimental approaches has been pivotal instrumental in advancing in this area of research. Notable computational tools, such as DeepNeuropePred, NeuroPred, and ProP^[111], as detailed in Table 8, each offer unique methodologies for predicting neuropeptide precursor cleavage sites.

Table 8. Tools for neuropeptide precursor cleavage sites.

Tools	Methods	Source code or online service website
DeepNeuropePred	Convolutional Neural Networks	https://github.com/ISYSLAB-HUST/DeepNeuropePred
NeuroPred	Logistic Regression, Known Motif models	http://neuroproteomics.scs.uiuc.edu/neuropred.html
ProP	Artificial Neural Networks	https://services.healthtech.dtu.dk/services/ProP-1.0/

Among these tools, DeepNeuropePred exemplifies a cutting-edge approach that integrates pre-trained language models with CNNs to effectively capture intricate sequence patterns, thereby enhancing the accuracy of cleavage site identification. This tool has been trained and validated using 31 precursor datasets sourced from the UniProt database. Its performance has been rigorously evaluated using various metrics, including area under the curve (AUC), ACC, precision, recall, area under the precision-recall curve (AUPR), MCC, false positive rate (FPR), and F1 score, ensuring a robust and reliable assessment.

In contrast, NeuroPred adopts an ensemble methodology that combines logistic regression with established motif models to predict and identify cleavage sites. This tool is based on experimentally validated cleavage sites for model training and evaluation, employing performance metrics such as correct classification rate, sensitivity, specificity, PPV, negative predictive value, MCC, and AUC, thus providing a comprehensive and reliable framework for cleavage site prediction. On the other hand, ProP focuses on the direct prediction of cleavage sites from amino acid sequences. By utilizing ANNs, it learns and analyzes complex sequence patterns to accurately predict cleavage points based on sequence characteristics. Collectively, these tools offer complementary and robust computational strategies for predicting neuropeptide cleavage sites, thereby significantly advancing research within this field.

These tools employ a variety of machine learning and deep learning techniques to detect cleavage sites within neuropeptide precursors. By integrating multiple models, such as neural networks, logistic regression, and support vector machines, alongside validated datasets, these tools excel at predicting the positions of cleavage sites^[112-114]. This functionality plays a critical role in advancing biological research by improving the understanding of neuropeptide biosynthesis and facilitating the design of subsequent experimental validation.

4. Conclusion and Future Work

Neuropeptides play a pivotal role in regulating various physiological processes, making them promising candidates for therapeutic development targeting a wide range of diseases. However, traditional experimental methods for discovering neuropeptides are often costly and time-consuming. Consequently, the development of efficient computational tools for identifying potential neuropeptides has become crucial. In this regard, bioinformatics databases and computational tools that support the analysis of neuropeptide data and their functional implications across multiple species are of significant value. Increasing attention is being devoted to the creation of computational strategies that facilitate the accurate prediction and comprehensive analysis of neuropeptides.

In the future, the development of neuropeptide databases and computational tools is expected to advance along several key directions. First, continuous efforts will likely focus on expanding and updating existing databases. This may involve increasing the coverage of neuropeptide data across a wider range of species and integrating more detailed functional and structural information, thereby enhancing the comprehensiveness and accuracy of the available data. Second, the performance of computational tools for neuropeptide analysis is expected to improve. This includes the development of more advanced algorithms capable of predicting neuropeptide sequences, structures, and functions with higher precision. For instance, future models may be further refined to better account for the complex interactions and modifications that neuropeptides undergo in physiological processes. Moreover, the user interfaces of these databases and tools may be further optimized to make them more accessible and user-friendly, even for researchers without extensive computational expertise. Finally, with the growing interest in the therapeutic potential of neuropeptides, future development is also expected to focus on practical applications in drug discovery and development. This may include the creation of specialized tools for screening neuropeptide-based drug candidates and predicting their efficacy and potential side effects, ultimately contributing to the development of novel treatments for various diseases.

Neuropeptides, owing their diverse biological functions and therapeutic potentials, have emerged as promising candidates for the

treatment of various neurological and inflammatory diseases. However, translating these predictive tools into clinical applications remains a considerable challenge.

Despite significant advancements in computational methods for neuropeptides prediction, several critical issues must be addressed to facilitate their clinical translation. Key concerns include toxicity, stability, and the feasibility of large-scale production. For example, the therapeutic utility of neuropeptides is often constrained by their susceptibility to enzymatic degradation and limited bioavailability. Additionally, the development of effective delivery systems that can overcome these challenges while minimizing adverse effects is crucial for their clinical success.

Future research should focus on integrating advanced computational tools with experimental approaches to better understand the pharmacokinetics and pharmacodynamics of neuropeptides. Such integration will not only enhance the predictive accuracy of computational models but also yield critical insights into the mechanisms underlying their therapeutic actions. In addition, overcoming challenges related to toxicity and stability, through innovative strategies such as nanocarrier-based delivery systems and bioconjugation techniques, could greatly improve the clinical applicability of neuropeptides. As the field continues to advance, the development of robust, high-precision computational tools, in conjunction with systematic experimental validation, is expected to facilitate the successful translation of neuropeptide-based therapeutics into clinical practice.

Future developments in the field are anticipated to involve more targeted predictions concerning emerging trends and associated challenges. One key direction will be the integration of multi-omics data to enhance the comprehensiveness and precision of neuropeptide analyses. Equally important is the establishment of standardized evaluation frameworks for benchmarking newly developed neuropeptide prediction tools. Such frameworks should incorporate unified performance metrics—including accuracy, precision, recall, F1-score, and MCC, to enable consistent and meaningful performance comparisons. Furthermore, the use of standardized datasets for cross-validation should be emphasized to improve the reliability and reproducibility of predictive models. In addition to predictive accuracy, future frameworks should also prioritize model interpretability, computational efficiency, and scalability for large-scale datasets, thereby supporting the practical deployment of these tools in diverse research and clinical settings.

In the future, research on neuropeptides is expected to advance through the integration of multi-omics data and the establishment of standardized evaluation frameworks. These developments will drive the expansion of databases, the optimization of algorithms, and the clinical translation of accurate prediction tools. Key challenges such as toxicity, stability, and drug delivery will remain central concerns. When combined with experimental validation, these efforts are expected to accelerate the development of novel therapeutic strategies and facilitate the transition from basic research to clinical applications.

Supplementary materials

The supplementary material for this article is available at: [Supplementary materials](#).

Declarations

Authors contribution

Xu W, Wang L, Wang M, Jiang J: Writing-original draft preparation.

Xue Z, Wang Y: Writing-review and editing, article conception and design.

All authors have given approval to the final version of the manuscript.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Funding

This work was supported by the National Natural Science Foundation of China under Grant 61772217 and 62172172, the Scientific Research Start-up Foundation of Binzhou Medical University under Grant BY2020KYQD01.

Copyright

© The Author(s) 2025.

References

1. Li C, Kim K. Neuropeptides. *WormBook*. 2008. [DOI]
2. Gomes I, Aryal DK, Wardman JH, Gupta A, Gagnidze K, Rodriguiz RM, et al. GPR171 is a hypothalamic G protein-coupled receptor for BigLEN, a neuropeptide involved in feeding. *Proc Natl Acad Sci*. 2013;110(40):16211-16216. [DOI]
3. Yu H, Wang Y, Gao J, Gao Y, Zhong C, Chen Y. Application of the neuropeptide NPVF to enhance angiogenesis and osteogenesis in bone regeneration. *Commun Biol*. 2023;6(1):197. [DOI]
4. Pol AN. Neuropeptide transmission in brain circuits. *Neuron*. 2012;76(1):98-115. [DOI]
5. Oliver G, Schäfer EA. On the physiological action of extracts of pituitary body and certain other glandular organs: Preliminary communication. *J Physiol*. 1895;18(3):277-279. [DOI]
6. Share L. Vasopressin and regulation of water homeostasis and cardiovascular function. In: McCann SM, editor. *Endocrinology: People and Ideas*. New York: Springer; 1988. p. 1-21. [DOI]
7. Vigneaud V, Lawler HC, Popenoe EA. Enzymatic cleavage of glycineamide from vasopressin and a proposed structure for this pressor-antidiuretic hormone of the posterior pituitary. *J Am Chem Soc*. 1953;75(19):4880-4881. [DOI]
8. Tatemoto K, Mutt V. Isolation of two novel candidate hormones using a chemical method for finding naturally occurring polypeptides. *Nature*. 1980;285:417-418. [DOI]
9. Bargmann W, Scharrer E. The site of origin of the hormones of the posterior pituitary. *Am Sci*. 1951;39(2):255-259. [PubMed]
10. Bohus B, de Wied D. Inhibitory and facilitatory effect of two related peptides on extinction of avoidance behavior. *Science*. 1966;153:318-320. [DOI]
11. de Wied D. Long term effect of vasopressin on the maintenance of a conditioned avoidance response in rats. *Nature*. 1971;232:58-60. [DOI]
12. de Wied D. Peptides and behavior. In: Zippel HP, editor. *Memory and Transfer of Information*. Boston: Springer; 1973. p. 373-389. [DOI]
13. Phetsanthead A, Vu NQ, Yu Q, Buchberger AR, Chen Z, Keller C, et al. Recent advances in mass spectrometry analysis of neuropeptides. *Mass Spectrom Rev*. 2021;42(2):706-750. [DOI]
14. Augustine JR, Mascagni F, McDonald AJ, Blake CA. Immunocytochemical staining of neuropeptide Y (NPY) in the insular lobe of the monkey: A light microscopic study. *Brain Res*. 1993;603(2):255-263. [DOI]
15. Gehlert DR, Gackenhaimer SL, Schober DA, Beavers L, Gadski R, Burnett JP, et al. The neuropeptide Y Y1 receptor selective radioligand, [¹²⁵I][Leu³¹,Pro³⁴]peptide YY, is also a high affinity radioligand for human pancreatic polypeptide 1 receptors. *Eur J Pharmacol*. 1996;318(2-3):485-490. [DOI]
16. Nathoo AN, Moeller RA, Westlund BA, Hart AC. Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species. *Proc Natl Acad Sci*. 2001;98(24):14000-14005. [DOI]
17. Fu Q, Li L. De novo sequencing of neuropeptides using reductive isotopic methylation and investigation of ESI QTOF MS/MS fragmentation pattern of neuropeptides with N-terminal dimethylation. *Anal Chem*. 2005;77(23):7783-7795. [DOI]
18. Smith SJ, Sümbül U, Graybuck LT, Collman F, Seshamani S, Gala R, et al. Single-cell transcriptomic evidence for dense intracortical neuropeptide networks. *eLife* 2019. [DOI]
19. Saklani P, Khan H, Gupta S, Kaur A, Singh TG. Neuropeptides: Potential neuroprotective agents in ischemic injury. *Life Sci*. 2022;288:120186. [DOI]
20. Baribeau DA, Anagnostou E. Oxytocin and vasopressin: linking pituitary neuropeptides and their receptors to social neurocircuits. *Front Neurosci*. 2015;9:335. [DOI]
21. Silva C, McNaughton N. Are periaqueductal gray and dorsal raphe the foundation of appetitive and aversive control? A comprehensive review. *Prog Neurobiol*. 2019;177:33-72. [DOI]
22. Sprouse-Blum AS, Smith G, Sugai D, Parsa FD. Understanding endorphins and their importance in pain management. *Haw Med J*. 2010;69(3):70. [PubMed]
23. Borbély É, Scheich B, Helyes Z. Neuropeptides in learning and memory. *Neuropeptides*. 2013;47(6):439-450. [DOI]
24. Sohn JW. Network of hypothalamic neurons that control appetite. *BMB Rep*. 2015;48(4):229-233. [DOI]
25. McGregor CE, English AW. The role of BDNF in peripheral nerve regeneration: Activity-dependent treatments and Val66Met. *Front Cell Neurosci*. 2018;12:522. [DOI]
26. Deussing JM, Chen A. The Corticotropin-releasing factor family: Physiology of the stress response. *Physiol Rev*. 2018;98(4):2225-2286. [DOI]
27. Beal MF, Martin JB. Neuropeptides in neurological disease. *Ann Neurol*. 1986;20(5):547-565. [DOI]
28. Chen XY, Du YF, Chen L. Neuropeptides exert neuroprotective effects in Alzheimer's disease. *Front Mol Neurosci*. 2019;11:493. [DOI]

29. Behl T, Madaan P, Sehgal A, Singh S, Makeen HA, Albratty M, et al. Demystifying the neuroprotective role of neuropeptides in Parkinson's disease: A newfangled and eloquent therapeutic perspective. *Int J Mol Sci.* 2022;23(9):4565. [DOI]
30. Rae M, Duarte ML, Gomes I, Camarini R, Devi LA. Oxytocin and vasopressin: Signalling, behavioural modulation and potential therapeutic effects. *Br J Pharmacol.* 2022;179(8):1544-1564. [DOI]
31. Huang Y, Lin X, Lin S. Neuropeptide Y and metabolism syndrome: An update on perspectives of clinical therapeutic intervention strategies. *Front Cell Dev Biol.* 2021;9:695623. [DOI]
32. Sanlioglu AD, Karacay B, Balci MK, Griffith TS, Sanlioglu S. Therapeutic potential of VIP vs PACAP in diabetes. *J Mol Endocrinol.* 2012;49(3):R157-R167. [DOI]
33. Widiapradja A, Chunduri P, Levick SP. The role of neuropeptides in adverse myocardial remodeling and heart failure. *Cell Mol Life Sci.* 2017;74(11):2019-2038. [DOI]
34. Tilan J, Kitlinska J. Neuropeptide Y (NPY) in tumor growth and progression: lessons learned from pediatric oncology. *Neuropeptides.* 2016;55:55-66. [DOI]
35. Wang M, Wang L, Xu W, Chu Z, Wang H, Lu J, et al. NeuroPep 2.0: An updated database dedicated to neuropeptide and its receptor annotations. *J Mol Biol.* 2024;436(4):168416. [DOI]
36. Yatawara CJ, Einfeld SL, Hickie IB, Davenport TA, Guastella AJ. The effect of oxytocin nasal spray on social interaction deficits observed in young children with autism: a randomized clinical crossover trial. *Mol Psychiatry.* 2016;21(9):1225-1231. [DOI]
37. Cagna FD, Fusar-Poli L, Damiani S, Rocchetti M, Giovanna G, Mori A, et al. The role of intranasal oxytocin in anxiety and depressive disorders: A systematic review of randomized controlled trials. *Clin Psychopharmacol Neurosci.* 2019;17(1):1-11. [DOI]
38. Valdez GR. Development of CRF1 receptor antagonists as antidepressants and anxiolytics: progress to date. *CNS Drugs.* 2006;20(11):887-896. [DOI]
39. Tang T, Tan Q, Han S, Diemar A, Löbner K, Wang H, et al. Receptor-specific recognition of NPY peptides revealed by structures of NPY receptors. *Sci Adv.* 2022;8(18):eabm1232. [DOI]
40. Gilotra NA, Russell SD. Arginine vasopressin as a target in the treatment of acute heart failure. *World J Cardiol.* 2014;6(12):1252-1261. [DOI]
41. Walker JM, Sandman CA, Bernston GG, McGivern RF, Coy DH, Kastin AJ. Endorphin analogs with potent and long-lasting analgesic effects. *Pharmacol Biochem Behav.* 1977;7(6):543-548. [DOI]
42. Hromníková D, Furka D, Furka S, Santana JAD, Ravingerová T, Klöcklerová V, et al. Prevention of tick-borne diseases: challenge to recent medicine. *Biologia.* 2022;77(6):1533-1554. [DOI]
43. Russo AF. Overview of neuropeptides: awakening the senses? *Headache.* 2017;57(S2):37-46. [DOI]
44. Yeoh JGC, Pandit AA, Zandawala M, Nässel DR, Davies SA, Dow JAT. DINer: Database for insect neuropeptide research. *Insect Biochem Mol Biol.* 2017;86:9-19. [DOI]
45. Albrechtsen NJW, Rehfeld JF. On premises and principles for measurement of gastrointestinal peptide hormones. *Peptides.* 2021;141:170545. [DOI]
46. Amara SG, Jonas V, Rosenfeld MG, Ong ES, Evans RM. Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. *Nature.* 1982;298(5871):240-244. [DOI]
47. Schubert ML, Rehfeld JF. Gastric peptides-gastrin and somatostatin. *Compr Physiol.* 2020;10(1):197-228. [DOI]
48. Spiess J, Mount CD, Nicholson WE, Orth DN. NH₂-terminal amino acid sequence and peptide mapping of purified human beta-lipotropin: comparison with previously proposed sequences. *Proc Natl Acad Sci USA.* 1982;79(16):5071-5075. [DOI]
49. Lundström L, Elmquist A, Bartfai T, Langel U. Galanin and its receptors in neurological disorders. *Neuromol Med.* 2005;7(1-2):157-180. [DOI]
50. Eva C, Serra M, Mele P, Panzica G, Oberto A. Physiology and gene regulation of the brain NPY Y₁ receptor. *Front Neuroendocrinol.* 2006;27(3):308-339. [DOI]
51. Tang T, Hartig C, Chen Q, Zhao W, Kaiser A, Zhang X, et al. Structural basis for ligand recognition of the neuropeptide Y₂ receptor. *Nat Commun.* 2021;12(1):737. [DOI]
52. Abualsaud N, Caprio L, Galli S, Krawczyk E, Alamri L, Zhu S, et al. Neuropeptide Y/Y₅ receptor pathway stimulates neuroblastoma cell motility through RhoA activation. *Front Cell Dev Biol.* 2021;8:627090. [DOI]
53. Carter CS. The oxytocin-vasopressin pathway in the context of love and fear. *Front Endocrinol.* 2017;8:356. [DOI]
54. Everett NA, McGregor IS, Baracz SJ, Cornish JL. The role of the vasopressin V_{1A} receptor in oxytocin modulation of methamphetamine primed reinstatement. *Neuropharmacology.* 2018;133:1-11. [DOI]
55. Cottrell GA. The first peptide-gated ion channel. *J Exp Biol.* 1997;200(18):2377-2386. [DOI]
56. Gorczyca M, Augart C, Budnik V. Insulin-like receptor and insulin-like peptide are localized at neuromuscular junctions in *Drosophila*. *J Neurosci.* 1993;13(9):3692-3704. [DOI]
57. Chang JC, Yang RB, Adams ME, Lu KH. Receptor guanylyl cyclases in *Inka* cells targeted by eclosion hormone. *Proc Natl Acad Sci USA.* 2009;106(32):13371-13376. [DOI]
58. Hökfelt T, Bartfai T, Bloom F. Neuropeptides: opportunities for drug discovery. *Lancet Neurol.* 2003;2(8):463-472. [DOI]
59. Broberger C. Brain regulation of food intake and appetite: molecules and networks. *J Intern Med.* 2005;258(4):301-327. [DOI]
60. Dejda A, Jozwiak-Bebenista M, Nowak JZ. PACAP, VIP, and PHI: effects on AC-, PLC-, and PLD-driven signaling systems in the primary glial cell cultures. *Ann N Y Acad Sci.* 2006;1070(1):220-225. [DOI]

61. Williams JT, Ingram SL, Henderson G, Chavkin C, von Zastrow M, Schulz S, et al. Regulation of μ -Opioid receptors: desensitization, phosphorylation, internalization, and tolerance. *Pharmacol Rev.* 2013;65(1):223-254. [DOI]
62. Slominski AT, Zmijewski MA, Zbytek B, Tobin DJ, Theoharides TC, Rivier J. Key role of CRF in the skin stress response system. *Endocr Rev.* 2013;34(6):827-884. [DOI]
63. Hökfelt T, Barde S, Xu ZQD, Kuteeva E, Rüegg J, Le Maitre E, et al. Neuropeptide and small transmitter coexistence: Fundamental studies and relevance to mental illness. *Front Neural Circuits.* 2018;12:106. [DOI]
64. Eiden LE, Hernández VS, Jiang SZ, Zhang L. Neuropeptides and small-molecule amine transmitters: cooperative signaling in the nervous system. *Cell Mol Life Sci.* 2022;79(9):492. [DOI]
65. Euler US, Gaddum JH. An unidentified depressor substance in certain tissue extracts. *J Physiol.* 1931;72(1):74-87. [DOI]
66. Saban R, Gerard NP, Saban MR, Nguyen NB, DeBoer DJ, Wershil BK. Mast cells mediate substance P-induced bladder inflammation through an NK(1) receptor-independent mechanism. *Am J Physiol Renal Physiol.* 2002;283(4):F616-F629. [DOI]
67. Xu ZQD, Zheng K, Hökfelt T. Electrophysiological studies on galanin effects in brain-progress during the last six years. *Neuropeptides.* 2005;39(3):269-275. [DOI]
68. Christy NP. Radioimmunoassay of human plasma ACTH and the pathogenesis of Cushing's disease. *Mt Sinai J Med.* 1973;40(3):298-301. [PubMed]
69. Rezaeian AH, Isokane T, Nishibori M, Chiba M, Hiraiwa N, Yoshizawa M, et al. alphaCGRP and betaCGRP transcript amount in mouse tissues of various developmental stages and their tissue expression sites. *Brain Dev.* 2009;31(9):682-693. [DOI]
70. Tatemoto K. Neuropeptide Y: complete amino acid sequence of the brain peptide. *Proc Natl Acad Sci USA.* 1982;79(18):5485-5489. [DOI]
71. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res.* 2023;51(D1):D523-D531. [DOI]
72. Wang Y, Wang M, Yin S, Jang R, Wang J, Xue Z, et al. NeuroPep: a comprehensive resource of neuropeptides. *Database.* 2015;2015:bav038. [DOI]
73. Koziol U, Koziol M, Preza M, Costabile A, Brehm K, Castillo E. De novo discovery of neuropeptides in the genomes of parasitic flatworms using a novel comparative approach. *Int J Parasitol.* 2016;46(11):709-721. [DOI]
74. Akhtar MN, Southey BR, Andrén PE, Sweedler JV, Rodriguez-Zas SL. Evaluation of database search programs for accurate detection of neuropeptides in tandem mass spectrometry experiments. *J Proteome Res.* 2012;11(12):6044-6055. [DOI]
75. Corbière A, Vaudry H, Chan P, Walet-Balieu ML, Lecroq T, Lefebvre A, et al. Strategies for the identification of bioactive neuropeptides in vertebrates. *Front Neurosci.* 2019;13:948. [DOI]
76. Pándy-Szekeres G, Taracena Herrera LP, Caroli J, Kermani AA, Kulkarni Y, Keserű GM, et al. GproteinDb in 2024: new G protein-GPCR couplings, AlphaFold2-multimer models and interface interactions. *Nucleic Acids Res.* 2024;52(D1):D466-D475. [DOI]
77. Kim Y, Bark S, Hook V, Bandeira N. NeuroPedia: neuropeptide database and spectral library. *Bioinformatics.* 2011;27(19):2772-2773. [DOI]
78. Burbach JPH. Neuropeptides from concept to online database www.neuropeptides.nl. *Eur J Pharmacol.* 2010;626(1):27-48. [DOI]
79. Harding SD, Armstrong JF, Faccenda E, Southan C, Alexander SPH, Davenport AP, et al. The IUPHAR/BPS guide to PHARMACOLOGY in 2024. *Nucleic Acids Res.* 2024;52(D1):D1438-D1449. [DOI]
80. Minkiewicz P, Iwaniak A, Darewicz M. BIOPEP-UWM database of bioactive peptides: current opportunities. *Int J Mol Sci.* 2019;20(23):5978. [DOI]
81. Wang J, Yin T, Xiao X, He D, Xue Z, Jiang X, et al. StraPep: a structure database of bioactive peptides. *Database.* 2018;2018:bay038. [DOI]
82. Liu F, Baggerman G, Schoofs L, Wets G. The construction of a bioactive peptide database in Metazoa. *J Proteome Res.* 2008;7(9):4119-4131. [DOI]
83. Dziuba J, Minkiewicz P, Nałecz D, Iwaniak A. Database of biologically active peptide sequences. *Food/Nahrung.* 1999;43(3):190-195. [DOI]
84. Shtatland T, Guettler D, Kossodo M, Pivovarov M, Weissleder R. PepBank-a database of peptides based on sequence text mining and public peptide data sources. *BMC Bioinformatics.* 2007;8:280. [DOI]
85. Zamyatnin AA, Borchikov AS, Vladimirov MG, Voronina OL. The EROP-Moscow oligopeptide database. *Nucleic Acids Res.* 2006;34:D261-266. [DOI]
86. Fälth M, Sköld K, Norrman M, Svensson M, Fenyö D, Andren PE. SwePep, a database designed for endogenous peptides and mass spectrometry. *Mol Cell Proteomics.* 2006;5(6):998-1005. [DOI]
87. Singh S, Chaudhary K, Dhanda SK, Bhalla S, Usmani SS, Gautam A, et al. SATPdb: a database of structurally annotated therapeutic peptides. *Nucleic Acids Res.* 2016;44(D1):D1119-D1126. [DOI]
88. Hiraki-Kajiyama T, Miyasaka N, Ando R, Wakisaka N, Itoga H, Onami S, et al. An atlas and database of neuropeptide gene expression in the adult zebrafish forebrain. *J Comp Neurol.* 2024;532(6):e25619. [DOI]
89. Pintado-Grima C, Bárcenas O, Iglesias V, Santos J, Manglano-Artuñedo Z, Pallarès I, et al. aSynPEP-DB: a database of biogenic peptides for inhibiting α -synuclein aggregation. *Database.* 2023;2023:baad084. [DOI]
90. Li H, Jiang L, Yang K, Shang S, Li M, Lv Z. iNP_ESM: Neuropeptide identification based on evolutionary scale modeling and unified representation embedding features. *Int J Mol Sci.* 2024;25(13):7049. [DOI]
91. Wang L, Huang C, Wang M, Xue Z, Wang Y. NeuroPred-PLM: an interpretable and robust model for neuropeptide prediction by protein language model. *Brief Bioinform.* 2023;24(2):bbad077. [DOI]

92. Liang Y, Cao M, Zhang S. NeuroPred-ResSE: Predicting neuropeptides by integrating residual block and squeeze-excitation attention mechanism. *Anal Biochem.* 2024;695:115648. [DOI]
93. Liu D, Lin Z, Jia C. NeuroCNN_GNB: an ensemble model to predict neuropeptides based on a convolution neural network and Gaussian naive Bayes. *Front.Genet.* 2023;14:1226905. [DOI]
94. Akbar S, Mohamed HG, Ali H, Saeed A, Khan AA, Gul S, et al. Identifying neuropeptides via evolutionary and sequential based multi-perspective descriptors by incorporation with ensemble classification strategy. *IEEE Access.* 2023;11:49024-49034. [DOI]
95. Madsen CT, Refsgaard JC, Teufel FG, Kjærulff SK, Wang Z, Meng G, et al. Combining mass spectrometry and machine learning to discover bioactive peptides. *Nat Commun.* 2022;13:6235. [DOI]
96. Wang L, Zeng Z, Xue Z, Wang Y. DeepNeuroPred: A robust and universal tool to predict cleavage sites from neuropeptide precursors by protein language model. *Comput Struct Biotechnol J.* 2023;23:309-315. [DOI]
97. Agrawal P, Kumar S, Singh A, Raghava GPS, Singh IK. NeuroPIpred: a tool to predict, design and scan insect neuropeptides. *Sci Rep.* 2019;9:5129. [DOI]
98. Wang G, Li X, Wang Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res.* 2016;44(D1):D1087-1093. [DOI]
99. Roy S, Teron R. BioDADPep: A Bioinformatics database for anti diabetic peptides. *Bioinformatics.* 2019;15(11):780-783. [DOI]
100. Tyagi A, Tuknait A, Anand P, Gupta S, Sharma M, Mathur D, et al. CancerPPD: a database of anticancer peptides and proteins. *Nucleic Acids Res.* 2015;43(D1):D837-D843. [DOI]
101. Grønning AGB, Kacprowski T, Schéele C. MultiPep: a hierarchical deep learning approach for multi-label classification of peptide bioactivities. *Biol Methods Protoc.* 2021;6(1):bpab021. [DOI]
102. Jiang M, Zhao B, Luo S, Wang Q, Chu Y, Chen T, et al. NeuroPpred-Fuse: an interpretable stacking model for prediction of neuropeptides by fusing sequence information and feature selection methods. *Brief Bioinform.* 2021;22(6):bbab310. [DOI]
103. Bin Y, Zhang W, Tang W, Dai R, Li M, Zhu Q, et al. Prediction of neuropeptides from sequence information using ensemble classifier and hybrid features. *J Proteome Res.* 2020;19(9):3732-3740. [DOI]
104. Singh V, Singh SK, Sharma R. A novel framework based on explainable AI and genetic algorithms for designing neurological medicines. *Sci Rep.* 2024;14:12807. [DOI]
105. Zandawala M, Bilal Amir M, Shin J, Yim WC, Alfonso Yañez Guerra L. Proteome-wide neuropeptide identification using NeuroPeptide-HMMer (NP-HMMer). *Gen Comp Endocrinol.* 2024;357:114597. [DOI]
106. Kang J, Fang Y, Yao P, Li N, Tang Q, Huang J. NeuroPP: A tool for the prediction of neuropeptide precursors based on optimal sequence composition. *Interdiscip Sci Comput Life Sci.* 2019;11(1):108-114. [DOI]
107. Karsenty S, Rappoport N, Ofer D, Zair A, Linal M. NeuroPID: a classifier of neuropeptide precursors. *Nucleic Acids Res.* 2014;42(W1):W182-W186. [DOI]
108. Amare A, Hummon AB, Southey BR, Zimmerman TA, Rodriguez-Zas SL, Sweedler JV. Bridging neuropeptidomics and genomics with bioinformatics: Prediction of mammalian neuropeptide prohormone processing. *J Proteome Res.* 2006;5(5):1162-1167. [DOI]
109. Hasan MM, Alam MA, Shoombatong W, Deng HW, Manavalan B, Kurata H. NeuroPred-FRL: an interpretable prediction model for identifying neuropeptide using feature representation learning. *Brief Bioinform.* 2021;22(6):bbab167. [DOI]
110. Chen S, Li Q, Zhao J, Bin Y, Zheng C. NeuroPred-CLQ: incorporating deep temporal convolutional networks and multi-head attention mechanism to predict neuropeptides. *Brief Bioinform.* 2022;23(5):bbac319. [DOI]
111. Duckert P, Brunak S, Blom N. Prediction of proprotein convertase cleavage sites. *Protein Eng Des Sel.* 2004;17(1):107-112. [DOI]
112. Hummon AB, Hummon NP, Corbin RW, Li L, Vilim FS, Weiss KR, et al. From precursor to final peptides: a statistical sequence-based approach to predicting prohormone processing. *J Proteome Res.* 2003;2(6):650-656. [DOI]
113. Southey BR, Sweedler JV, Rodriguez-Zas SL. Prediction of neuropeptide cleavage sites in insects. *Bioinformatics.* 2008;24(6):815-825. [DOI]
114. Southey BR, Sweedler JV, Rodriguez-Zas SL. A python analytical pipeline to identify prohormone precursors and predict prohormone cleavage sites. *Front Neuroinform.* 2008;2:7. [DOI]

Publisher's Note

Science Exploration remains a neutral stance on jurisdictional claims in published maps and institutional affiliations. The views expressed in this article are solely those of the author(s) and do not reflect the opinions of the Editors or the publisher.