SJP 7(1) (2026)

Sasambo Journal of Pharmacy

https://jffk.unram.ac.id/index.php/sjp

Computational intelligence in cell-penetrating peptide discovery: emerging paradigms and predictive frameworks

Ebrahim Sadaqa[1], Muhammad Ikhlas Arsul[2], Nurul Hidayati[3], Arif Setiawansyah[3,4*]

[1] Research Center for Vaccine and Drugs, National Research and Innovation Agency (BRIN), Bogor, Indonesia
[2] Department of Pharmacy, Faculty of Medicine and Health Sciences, Universitas Islam Negeri Alauddin, Gowa, Indonesia
[3] Pharmacy Diploma Program, Akademi Farmasi Cendikia Farma Husada, Bandar Lampung, Indonesia
[4] PharmaSynthix Indonesia, Bandar Lampung, Indonesia

DOI: https://doi.org/10.29303/sjp.v7i1.677

Article Info
Received : 2026-05-07
Revised : 2026-06-03
Accepted : 2026-06-05

Abstract: Researchers are increasingly fascinated by cell penetrating peptides (CPPs) owing to their ability to deliver biomolecules within cells. This unique capability makes CPPs an incredibly valuable asset in fields such as drug delivery, gene therapy and imaging where precision is paramount. However, the design and prediction of CPPs with optimal attributes and augmented cellular uptake still pose a challenge. With this literature review piece, we hope to shed light on significant progress that has been made within the field of in silico prediction techniques for CPPs. In particular our focus lies on discussing how these methodologies can assist with discovering and optimizing promising CPP candidates more efficiently. To accomplish this goal respectfully and comprehensively, we will examine computational methods like machine learning algorithms, sequence-based techniques, and structure-based modeling. These techniques employ large-scale databases, extensive peptide libraries, and advanced algorithms to scrutinize peptide properties, predict CPP activity, and optimize CPP sequences. Within this analysis, we find a significant emphasis placed on integrating physicochemical features, sequence motifs, and structural information into predictive models relating to CPP development. These factors are essential to predicting successful peptide uptake within cells while presenting various challenges during this process. Furthermore, the review highlights considerable strides in silico prediction techniques for CPPs, demonstrating great promise in accelerating their identification for further development. Incorporating computational tools and experimental validation holds tremendous promise in facilitating the design of CPPs with enhanced properties, thus advancing the field of intracellular delivery and therapeutics.

Keywords: cell-penetrating peptides; cellular uptake; in silico prediction techniques; machine learning; toxicity prediction.

Citation: Sadaqa, E., Arsul, M.I., Hidayati, N., & Setiawansyah, A. (2026). Computational intelligence in cell-penetrating peptide discovery: emerging paradigms and predictive frameworks. *Sasambo Journal of Pharmacy*, *7*(1), 33-51. doi: https://doi.org/10.29303/sjp.v7i1.677

Introduction

Cell-penetrating peptides (CPPs) are short amino acid sequences, typically ranging from 4 to 40 residues, that possess the ability to efficiently cross cellular membranes (Park et al., 2023). The field of drug delivery has been deeply intrigued by CPPs because they possess a unique ability to facilitate the transportation of diverse cargo molecules into cells (Heitz et al., 2009). Natural proteins, synthetic peptides, and peptide mimetics are among the various sources that produce these promising peptides. The fundamental characteristics of CPPs include their capacity to penetrate the cell membrane without causing substantial toxicity (Li et al., 2015; Xie et al., 2020). The potential utility of CPPs in drug delivery is extensive.

Email: arif12.setiawansyah@gmail.com *(*Corresponding Author)*

CPPs are characterized by their magnificent ability to ease intracellular transport of different therapeutic cargoes, such as chemotherapeutic agents, antimicrobial agents, and nucleic acids. This effective approach holds significant promise in enhancing the intracellular delivery and bioavailability of these therapeutic agents (Heitz et al., 2009). CPPs have demonstrated significant potential; however, multiple obstacles must be overcome before widespread clinical utilization. Some of these obstacles entail enhancing peptide design for improved efficiency, stability, and cargo release properties, while simultaneously reducing potentially damaging cytotoxic and immunogenic reactions.

Validating and optimizing CPPs through experimentation can be demanding and arduous. As a result, computational methods have been devised as a substitute to accelerate the discovery process (Kardani & Bolhassani, 2021; Porosk, Põhako, et al., 2021; Schissel et al., 2021). These methods employ diverse peptide properties, such as amino acid composition, biochemical characteristics, and expression patterns, to achieve prediction accuracies exceeding 80% (Park et al., 2023). Forecasting with certainty the behavior and effectiveness of CPPs is a demanding undertaking, given the intricate nature of interactions between peptides and cell membranes, combined with the multitude of available CPP sequences (Gräslund et al., 2011). To surmount these obstacles, it is imperative to attain a comprehensive understanding and thorough investigation of the fundamental mechanisms that regulate the interactions between CPPs and cell membranes.

In recent years, computational approaches have emerged as invaluable tools for facilitating the design and evaluation of CPPs (H. M. Lee et al., 2021). State-of-the-art techniques employ advanced algorithms and extensive databases to examine the characteristics of peptides, predict their cellular internalization, and assess their potential effectiveness. By utilizing computational modelling and machine learning methods researchers can expedite the identification and optimization of CPP candidates. This strategy is time efficient and economically savvy leading to considerable savings.

This review examines computational approaches for faster prediction of CPPs, aiming to expand understanding and speed up the identification and development of novel CPPs. Researchers can create and improve CPPs with enhanced cellular uptake and therapeutic effectiveness by incorporating computational methods. This review provides valuable insights for researchers and practitioners in peptide design, advancing CPP research and fostering more efficient and successful therapeutic strategies.

Several databases or computer-based electronic searches were used to access all journals relating to the prediction of cell penetrating peptides, including Pubmed, Science Direct, and Google Scholar. These searches used the keywords "cell penetrating peptides," "prediction," "in silico," "toxicity," and "molecular dynamic simulation."

Materials and Methods

Literature search strategy

A systematic literature search was conducted across PubMed, Scopus, Web of Science, and Google Scholar, covering publications from 1990 to 2023. The search combined MeSH terms and free-text keywords across three conceptual domains: peptide biology ("cell-penetrating peptide," "CPP," "protein transduction domain," "membrane-active peptide"), computational methodology ("machine learning," "deep learning," "artificial intelligence," "molecular dynamics," "quantitative structure-activity relationship," "QSAR," "transformer," "graph neural network"), and functional descriptors ("intracellular delivery," "cargo transport," "endosomal escape," "uptake prediction"). Boolean operators (AND, OR, NOT) linked terms across domains. Searches were restricted to peer-reviewed articles published in English. Preprints deposited in bioRxiv and arXiv were not included.

Inclusion and exclusion criteria

Studies were included if they described or applied a computational model for CPP identification, classification, or design; reported quantitative performance metrics; or introduced a publicly accessible dataset, benchmark, or web server relevant to CPP prediction. Reviews, meta-analyses, and primary experimental studies that incorporated computational validation pipelines were also retained when they provided sufficient methodological detail for extraction.

Studies were excluded if they addressed antimicrobial, antifungal, or antiviral peptides without explicit CPP characterization; reported computational methods applied exclusively to non-peptide delivery systems; lacked accessible performance data; or were conference abstracts without full-text availability.

Study selection and data extraction

Duplicate records were removed using Rayyan. Two reviewers independently screened titles and abstracts; disagreements were resolved by consensus. Full texts of candidate articles were retrieved and assessed against the inclusion criteria. For each included study, the following data were extracted: (1) year of publication and journal; (2) CPP dataset used (name, size, positive/negative ratio, source); (3) computational

approach (feature representation, model architecture, training strategy); (4) reported performance metrics and validation scheme (cross-validation, independent test set, holdout); (5) availability of source code or web server; and (6) biological or experimental validation, if performed.

Results and Discussion

Advancements of CPPs in Biomedical Delivery

CPPs have garnered significant interest in the pharmaceutical sciences for diverse reasons. Among its benefits lies cell-penetrating peptides' ability to deliver many therapeutic agents (Futaki et al., 2002). The mechanism allows compounds such as small molecules alongside large entities like proteins or nucleic acids to access cells with utmost precision. Such targeted intracellular transport proves crucial for progressive fields, including gene therapy, as seen from emerging methodologies utilizing protein-based therapeutics (El-Andaloussi et al., 2005; Veldhoen et al., 2008). Ongoing studies may disclose new scientific pathways within this area, warranting more exploration into this promising strategy still early in its discovery stage. CPPs function as versatile carriers that can overcome cellular barriers, such as the plasma membrane, which often restrict the intracellular availability and efficacy of therapeutic molecules (Y. Liu et al., 2022). CPPs have witnessed extensive utilization across varied domains of medicinal application, encompassing cancer therapy, anti-inflammatory therapy, and vaccine development (Khairkhah et al., 2023). CPPs can be an invaluable asset for targeted drug delivery systems, owing to their ability to precisely target specific intracellular compartments or organelles. By ingeniously incorporating organelle-targeting signals into CPP sequences, such as nuclear localization signals (NLS) or mitochondrial targeting sequences (MTS), cargo molecules can be delivered with exceptional specificity to the desired cellular compartments (Kardani et al., 2019). This targeted delivery is particularly significant for treatments that necessitate precise localization and action within specific cellular compartments. In addition, CPPs hold promise for resolving obstacles linked to dispensing drugs, including inadequate cellular uptake or restricted bioavailability. Their ability to facilitate cellular internalization through mechanisms like endocytosis or direct translocation across the plasma membrane provides a valuable approach for enhancing the efficacy of therapeutic agents (Erazo-Oliveras et al., 2012).

Decoding The Cellular Uptake Mechanisms of CPPs

The cellular uptake mechanisms of cell-penetrating peptides (CPPs) remain a subject of ongoing debate within the scientific community (Gräslund et al., 2011). Several pathways have been proposed to explain CPP internalization, including direct translocation, endocytosis, and membrane disruption (**Figure 1**) (Foged & Nielsen, 2008).

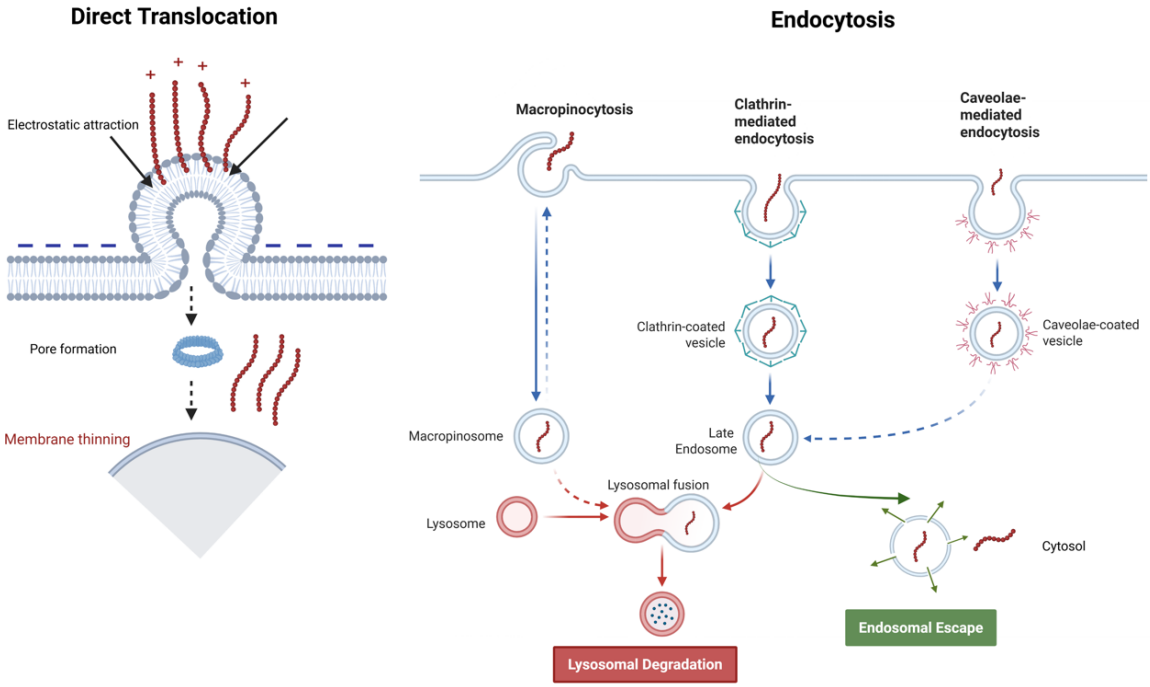

Figure 1. Mechanism of CPPs in cellular uptake

Direct translocation represents an energy-independent uptake mechanism in which CPPs cross the plasma membrane without vesicle formation. This process is primarily driven by electrostatic interactions between the positively charged residues of CPPs and the negatively charged components of the cell membrane. Such interactions may destabilize the lipid bilayer, leading to transient membrane perturbation or rupture that enables CPPs to directly penetrate the membrane (Henriques et al., 2006). In contrast, endocytosis is an energy-dependent pathway and constitutes one of the most frequently reported mechanisms for CPP internalization, particularly at lower peptide concentrations. CPPs may enter cells via multiple endocytic routes, including clathrin-mediated endocytosis, caveolin-mediated endocytosis, and macropinocytosis. In these processes, CPPs are encapsulated within membrane-bound vesicles that facilitate their transport into the cytoplasm (Richard et al., 2003). Additionally, CPPs may induce membrane disturbance or disruption upon interaction with the lipid bilayer. This interaction can result in the formation of transient pores or membrane defects, allowing peptides to pass directly into the cell interior (Futaki et al., 2020; Zorko & Langel, 2005).

Experimental evidence suggests that CPP uptake mechanisms are concentration-dependent. At low concentrations, many CPPs predominantly utilize endocytic pathways, whereas at higher concentrations, direct translocation becomes more prominent (Kauffman et al., 2015). For instance, increased concentrations of well-known CPPs such as HIV-1 Tat (GRKKRRQRRRPPQ) and octa-arginine (R8; RRRRRRRR) have been shown to shift their uptake mechanism from endocytosis to direct membrane translocation (Duchardt et al., 2007). Furthermore, changes in experimental conditions, such as temperature reduction, can influence CPP uptake pathways. Studies have demonstrated that HIV-1 Tat uptake persists even under conditions that inhibit clathrin- and caveolin-mediated endocytosis or at low temperatures (4 °C), suggesting a dominant role of direct translocation under these conditions. This observation indicates that membrane physical properties, including fluidity and flexibility, play a critical role in CPP-mediated penetration (Ouyang et al., 2022; Ter-Avetisyan et al., 2009). Several theoretical models have been proposed to explain direct CPP penetration across membranes, including the pore formation model (Gazit et al., 1994; Herce & Garcia, 2007), inverted micelle model (Derossi et al., 1996; Prochiantz, 1996), carpet model (Pouny et al., 1992; Taylor et al., 2000), and membrane thinning model (M. T. Lee et al., 2005).

It is important to emphasize that CPP uptake mechanisms are highly context-dependent and influenced by multiple factors, such as peptide sequence and structure, cell type, membrane lipid composition, and experimental conditions (Jones & Sayers, 2012; M. T. Lee et al., 2005). Moreover, individual CPPs may employ multiple uptake pathways simultaneously or switch mechanisms depending on environmental cues (Nakase et al., 2004).

Challenges in Predicting CPPs

Predicting the behavior and efficacy of CPPs poses several challenges due to the complex nature of peptide-membrane interactions and the diverse range of CPP sequences (Gräslund et al., 2011). Several challenges related to prediction CPPs were summarized as in **Figure 2**.

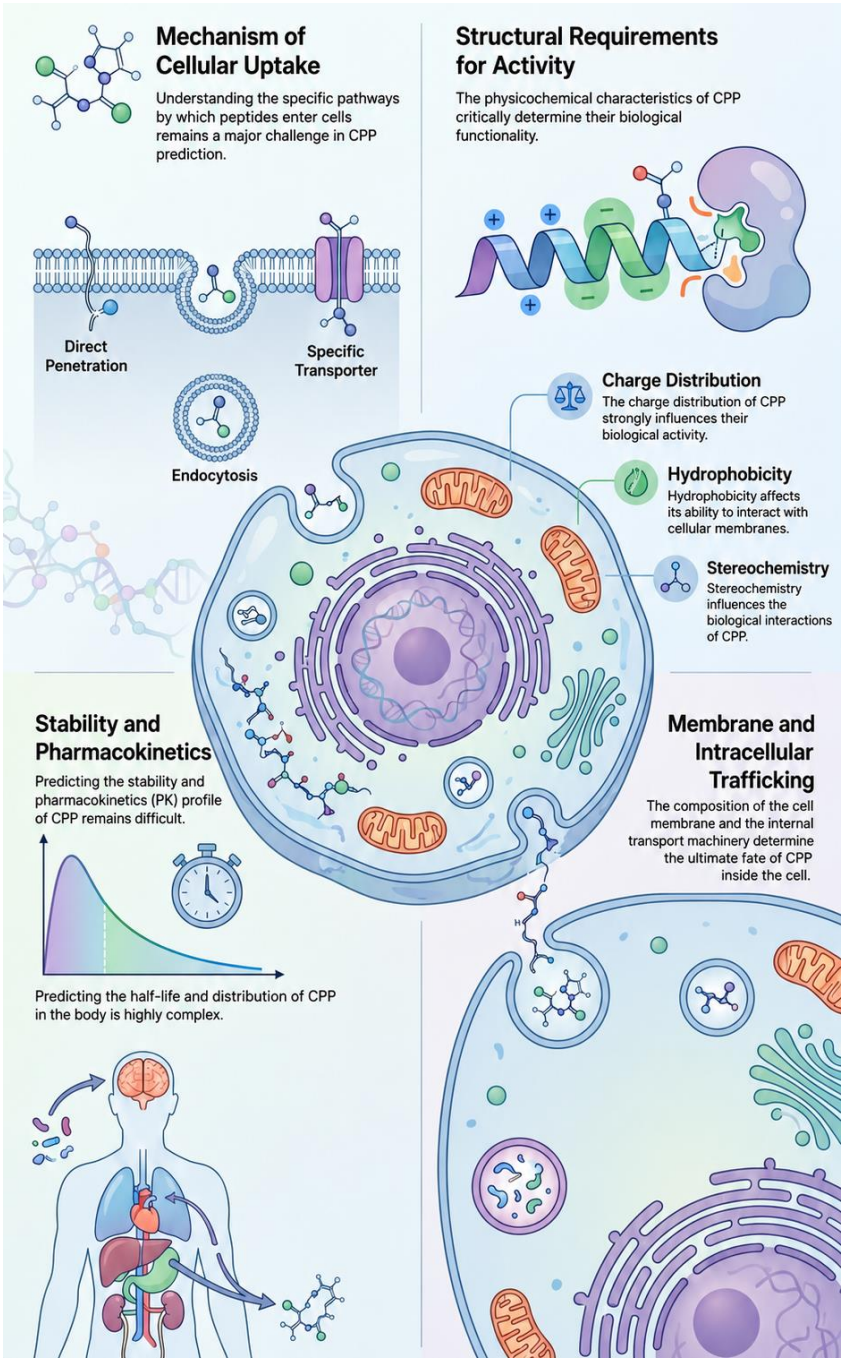

Figure 2. Challenges in Predicting the Behavior and Efficacy of CPPs

One of the foremost challenges in the prediction of CPPs lies in comprehending the mechanisms that govern their cellular internalization. CPPs employ diverse mechanisms for internalization, including endocytosis and direct translocation through the plasma membrane (Langel, 2021). The specific mechanism can vary depending on the CPP sequence, cargo, cell type, and experimental conditions. Consequently, accurately predicting the cellular uptake mechanism of a CPP remains challenging. An additional obstacle is the lack of a clear consensus on the structural requirements for CPP activity. According to several research findings, it is customary for CPPs to possess substantial levels of fundamental amino acids, predominantly arginine, and lysine. Exploration into these residues' most effective sequence and spatial arrangement for successful cell penetration is still incomplete (Deshayes et al., 2008). Additionally, the presence of hydrophobic residues in certain CPPs suggests their contribution to membrane interactions and translocation, further complicating the prediction process. Furthermore, CPPs often exhibit diverse cellular behaviors and can display cell-type specificity. The same CPP sequence may demonstrate varying efficiency or selectivity in different cell types

due to differences in membrane composition, receptor expression, and intracellular trafficking machinery (S. Patel et al., 2019; S. G. Patel et al., 2019). These variations make it challenging to accurately predict the efficacy and specificity of CPPs in different biological contexts. Moreover, it is very difficult to anticipate the stability and pharmacokinetic characteristics of CPPs. CPPs' in vivo stability and efficacy are constrained by proteases' propensity for degradation (Sarko et al., 2010). Furthermore, the way peptides are sized up in terms of their charge and hydrophobicity can have a bearing on their biodistribution patterns as well as tissue penetration capabilities while also influencing systemic clearance. Predicting these features accurately for a particular CPP sequence remains a difficult challenge. Improved computational models, a deeper understanding of the underlying mechanisms, and experimental validation would be necessary to get through these barriers and to have more accurate predictions.

Advances in Predicting and Understanding CPPs

Feature-based methods for predicting CPPs

There are several future-based methods for accurate prediction of CPPs that have been proposed in recent studies, as summarized in **Table 1**.

These methods use machine learning algorithms and sequence-based predictors to accurately identify CPPs (Fu et al., 2019, 2020). Each approach has its own set of advantages and disadvantages, and their performance may vary based on the dataset and application (Porosk, Gaidutšik, et al., 2021). These approaches give useful tools for future CPP prediction, helping researchers to save time and costs by empirically testing predictions (Porosk, Põhako, et al., 2021).

Table 1. Methods for Predicting CPPs using Future-Based Methods

Method	Description	References
Support Vector Machines (SVM)	supervised learning algorithms that use sequence characteristics and training data to categorize CPPs.	(Manavalan et al., 2018; Sanders et al., 2011; Singh et al., 2013)
Random Forest (RF)	An ensemble learning approach that utilizes multiple decision trees to predict the CPP activity through an examination of diverse sequences.	(P. Liu et al., 2022; Wei, Xing, et al., 2017)
Artificial Neural Networks (ANN)	computational models that draw inspiration from biological neural networks. These models have the ability to learn from given examples and predict CPP behavior.	(de Oliveira et al., 2021; Feger et al., 2020)
Hidden Markov Models (HMM)	They are capable of predicting various biological phenomena such as CPPs activity, transmembrane protein detection, and topology prediction based on patterns and statistical properties. This has been demonstrated in multiple studies.	(Bagos et al., 2004; Kahsay et al., 2005; Panuccio et al., 2002)
Deep Learning	Utilizes deep neural networks with multiple layers to automatically learn hierarchical representations of CPP sequences and predict their activity.	(Porosk, Gaidutšik, et al., 2021; Qiang et al., 2018)

Key criteria for effective CPPs

In the realm of projecting efficacious CPPs, it is absolutely crucial to apprehend the standards that carry significant leverage over their effectiveness (**Figure 3**). The net charge of CPPs influences their interaction with negatively charged cell membranes, and positively charged CPPs effectively interact with negatively charged cell membrane phospholipids (Derakhshankhah & Jafari, 2018). Certain structural elements, such as alpha-helices, have been found to enhance the cell-penetrating properties of CPPs. Incorporating helical motifs, either naturally or by design, can improve membrane penetration and cellular uptake (Kalafatovic & Giralt, 2017).

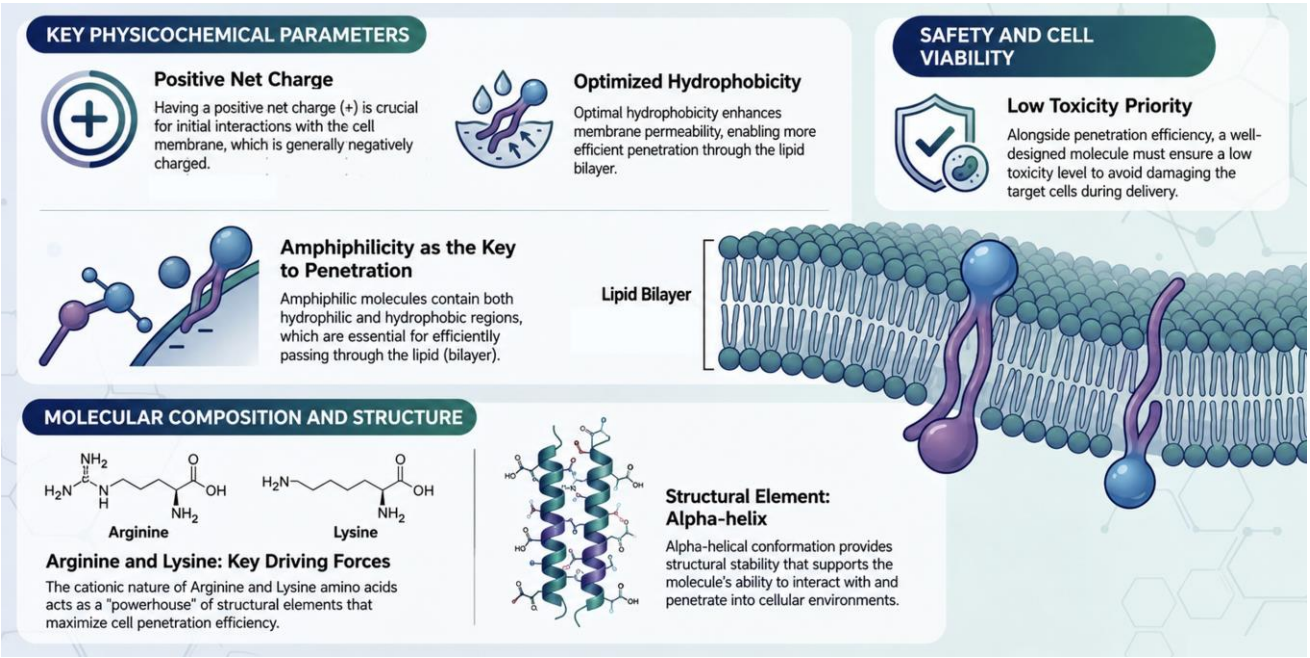

Figure 3. Key Factors Influencing Efficient CPPs

The hydrophobicity of CPPs affects their ability to traverse lipid-rich cell membranes. A balance between hydrophobic and hydrophilic residues is crucial for optimal membrane penetration. Moderate hydrophobicity, characterized by a hydrophobic moment, aids in membrane insertion while avoiding excessive cell toxicity (Allen & Pellois, 2022).

The characteristic of amphiphilicity pertains to the possession of both hydrophobic and hydrophilic domains, which can significantly contribute to the efficient cellular internalization facilitated by CPPs. The hydrophobic regions present in CPPs enable their interaction with the hydrophobic lipid bilayer of the cellular membrane, thereby facilitating their cellular internalization. Simultaneously, the hydrophilic domains of CPPs engage in interactions with the aqueous milieu, thereby contributing to the preservation of solubility and stability (Kang et al., 2019). CPPs tend to assume an alpha-helical structure during their interaction with cellular membranes. The stable helix formation is facilitated by the amphiphilic properties of the peptides, wherein the hydrophobic residues tend to aggregate towards the lipid bilayer, while the hydrophilic residues remain exposed to the aqueous surroundings. The helical configuration of CPPs has been observed to augment their membrane-binding capacity and facilitate their effective internalization into cells, as reported by Klein et al. (2017).

CPPs are characterized by a significant abundance of positively charged amino acid residues, such as arginine and lysine. This feature enables them to establish electrostatic interactions with the negatively charged phospholipids (Futaki, 2005). The importance of low toxicity in CPPs lies in their ability to effectively deliver therapeutic cargo into cells without causing significant harm or adverse effects. CPPs with low toxicity profiles are more likely to be safe for use in various biological systems, making them attractive candidates for drug delivery and other biomedical applications (Tünnemann et al., 2008).

Bioinformatics tools and websites for predicting CPPs

Over recent years, peptide science has undergone dramatic change due to emerging in silico methodologies. Future-oriented techniques have gained great attention for their ability to generate novel therapeutic peptides and accurately forecast functional performance of existing ones (Angell et al., 2018; Fosgerau & Hoffmann, 2015; Plisson et al., 2020). Research studies have proven the effectiveness of machine learning instruments for forecasting and designing CPPs. To add to this body of knowledge, we conducted an in-depth exploration of various machine learning tools utilized for CPP prediction and design between 2013 and the present, ultimately resulting in the identification and summary of popular prediction tools as outlined in **Table 2**.

Table 2. Summary of Prominent Tools for Cell-Penetrating Peptide Prediction

Tool/Website	Description	Reference
CPPpred	web server employs an N-to-1 neural network to predict CPPs with high accuracy. It includes a sequence analysis tool for determining the possibility of a peptide being CPP.	(Holton et al., 2013)
CellPPD	A database and web server that predicts CPPs based on multiple sequence properties using a support vector machine (SVM).	(Gautam et al., 2013)
C2Pred	A web server that uses a support vector machine to distinguish cell-penetrating peptides from non-cell-penetrating peptides.	(Tang et al., 2016)
CPPsite 2.0	A web server hosting a large database of experimentally validated CPPs.	(Agrawal et al., 2016)
CPPred-RF	A web server that predicts CPPs using a random forest algorithm and sequence-derived features. It incorporates a two-layer prediction framework.	(Wei, Xing, et al., 2017)
SkipCPP-Pred	An improved sequence predictor that uses the adaptive k-skip-n-gram algorithm to gather residue intrinsic correlation informatio. It employs a random forest classifier and demonstrates promising results.	(Wei, Tang, et al., 2017)
StackCPPred	Highly robust tool that utilizes train stacking-powered machine learning techniques to accurately predict CPP and evaluate their efficiency in uptake	(Fu et al., 2020)
BChemRF-CPPred	A framework that employs machine learning techniques, specifically artificial neural networks, support vector machines, and Gaussian process classifiers, to effectively distinguish between CPPs and non-CPPs. The framework is adept at extracting descriptors that are based on both sequence and structure from PDB and FASTA formats. Moreover, it is conveniently available as a web server.	(de Oliveira et al., 2021)
MLCPP 2.0	An interpretable stacking model for identifying CPPs and assessing their strength of uptake efficiency.	(Manavalan & Patra, 2022)
AiCPP	A deep learning-based model that accurately predicts the propensity of peptides to act as CPPs. It significantly reduces false positive predictions and exhibits high accuracy.	(Park et al., 2023)

Another physics-based computationally versatile tool is CELLPM, whose web site has the potential to predict the ability of peptides to penetrate cells and cross the lipid bilayer through an energy-independent pathway (Lomize & Pogozheva, 2018). It used one model cell membrane lipid bilayer type, which is composed of dioleoyl phosphatidylcholine (DOPC) with adjustable parameters of PH and temperature. The input can be amino acid sequences or a PDB file. This tool has the ability to give the results of interaction properties with the calculation of the log of permeability coefficient, in which a low value suggests limited ability of the tested peptide to penetrate cell membrane and a high value suggests potential ability of the tested peptide to penetrate cell membrane, as well as an energy profile that shows how energy fluctuates at different points (Z-axis) throughout the peptide within the membrane. The energy levels at each place indicate the peptide's stability in that region.

For further prediction of peptide subcellular localization, DeepLoc 2.0 and CELLO II can be used. DEEPloc 2.0 is a kind of prediction algorithm that uses deep neural networks and can make predictions from input sequences of peptides (Almagro Armenteros et al., 2017; Thumuluri et al., 2022). CELLO II is a two-level SVM system in which the first level comprises a number of SVM classifiers, each based on a specific type of feature vector derived from sequences; the second level SVM classifier functions as the jury machine to generate the probability distribution of decisions for possible localizations (Yu et al., 2004).

After assessing the previous tools and the different tools outlined in Table 2, it is apparent that these resources offer an all-inclusive range of features and algorithms for the prognosis of CPPs. Experts in the domain of peptide science now have access to a diverse variety of tools that facilitate the creation and examination of CPPs. The scores provided by these prognostication tools serve as valuable indications of the probability or likelihood of a peptide being classified as a CPP. These scores are derived from specific prognostication algorithms and the integration of various sequence-based features. To put it generally, a peptide has more chances of having cell-penetrating characteristics with higher scores. Conversely, lower scores suggest reduced likelihoods. It is noteworthy that the interpretation of these scores may differ based on the tool employed and the specific thresholds established by its developers. To accurately establish whether a peptide is anticipated to be a CPP or not, it is crucial to take into account the suggested cutoffs or thresholds provided by each tool.

Recent research has highlighted the significance of predicting key parameters and physico-chemical

parameters that influence the functionality of CPPs, with increasing focus on optimizing factors such as hydrophobicity and helical structure. The design and enhancement of novel CPPs often involve carefully manipulating these parameters to achieve efficient cellular uptake. In this context, several prediction tools have been developed to estimate the hydrophobicity and helical content of peptides (De Cena et al., 2022; Kardani & Bolhassani, 2021).

For assessing hydrophobicity, commonly utilized tools include TMHMM and ProtParam. TMHMM is a bioinformatics tool that predicts transmembrane helices in protein sequences, indirectly providing insights into the hydrophobicity of CPPs (Krogh et al., 2001). On the other hand, ProtParam is a digital tool accessible via the internet that calculates several parameters of protein sequences, encompassing the hydrophobicity index. By employing different scales such as the Kyte-Doolittle or Hopp-Woods scale, ProtParam estimates the hydrophobicity based on the peptide's amino acid composition. Higher values on the hydrophobicity index indicate greater hydrophobicity (Gasteiger et al., 2005).

Regarding the prediction of helical structure, crucial tools include AGADIR, HELIquest web server, and PEP-FOLD (Fisinger et al., 2001; Gautier et al., 2008; Thévenet et al., 2012). AGADIR predicts the helical content of peptides by employing empirical thermodynamic parameters and the peptide's amino acid composition (Muñoz & Serrano, 1994, 1995). It calculates a score or energy value representing the stability of the predicted helical structure, with higher scores indicating a higher likelihood of helicity. Similarly, the HELIquest web server and PEP-FOLD utilize algorithms and machine learning methods to predict the secondary structure, including helicity, of peptides based on their amino acid sequence. These tools assign scores or probabilities to the predicted secondary structure elements, with higher scores indicating a higher likelihood of helical structure. The prediction of these parameters is of significant importance in understanding and enhancing the functionality of CPPs. By utilizing these prediction tools, researchers can efficiently design or optimize CPPs, saving valuable time and resources before experimentally validating the in-silico results. To develop more effective CPP based therapeutics, a combined approach of computational predictions and experimental validation offers valuable insights into peptides potential for cell penetration. Researchers may utilize guidelines provided to make informed decisions about their peptides' ability to penetrate cells.

Toxicity-informed prediction of CPPs

To gain a comprehensive understanding of predicting CPPs, it is imperative to consider the parallel prediction of toxicity parameters associated with CPPs (Gupta et al., 2013). By incorporating these toxicity assessments alongside functionality predictions, researchers can enhance the potential of accurately identifying CPPs with both high efficacy and a favorable safety profile. This integrated approach ensures a more holistic evaluation of CPP candidates, facilitating their practical application in various biomedical fields.

One of the drawbacks associated with CPPs pertains to their immunogenicity. Specifically, in vivo studies have demonstrated that peptides have the capacity to activate immunologic responses, which may lead to allergic reactions. The presence of peptides within the body may also stimulate the production of antibodies, thereby reducing the therapeutic benefits of CPPs and impeding their overall efficiency (Kuriakose et al., 2016; Shankar et al., 2014). Given these concerns, evaluating immunogenicity, allergenicity, general toxicity, and potential hemolytic effects is essential to ensure the safety profile of CPPs. Consequently, we have investigated the most commonly utilized in silico tools for this purpose, as outlined in **Table 3**.

Table 3. Prediction Tools for Assessing Immunogenicity, Allergenicity, Toxicity, and Hemolytic Effects of Peptides and Proteins

Tool/Website	Parameter	Description	Reference
Immune Epitope Database (IEDB)	Immunogenicity	This online resource provides an estimation of peptide immunogenicity, encompassing MHC binding, T-cell epitopes, and antigenicity.	(Vita et al., 2015)
NetMHCpan	Immunogenicity	An online tool designed to compute the probability of a peptide binding to MHC class I molecules, indicating its immunogenic potential.	(Nielsen et al., 2007)
VaxiJen	Immunogenicity	An online computational tool that utilizes physicochemical properties and amino acid composition to predict the antigenicity of protein sequences.	(Doytchinova & Flower, 2007)
ANTIGENpro	Immunogenicity.	A web server that is capable of predicting B-cell epitopes and antigenic regions in protein sequences.	(Magnan et al., 2010)
AllerTOP	Allergenicity	A web server for predicting allergenicity of proteins and peptides based on their amino acid sequence.	(Dimitrov, Bangov, et al., 2014)
AllergenFP	Allergenicity	A machine learning-based tool that predicts allergenicity by analyzing the protein's physicochemical properties and sequence motifs.	(Dimitrov, Naneva, et al., 2014)
ToxinPred	Toxicity	A web server for predicting the toxicity of peptides based on their physicochemical properties and amino acid composition.	(Gupta et al., 2013)
HemoPI	Hemolytic potency	SVM-based method to predict hemolytic potency.	(Chaudhary et al., 2016)
HemoPRED	Hemolytic effect	An internet-based server that utilizes machine learning models to prognosticate the hemolytic potential of peptides through an investigation of features derived from their sequence.	(Win et al., 2017)

These tools play a crucial role in the field of peptide and protein research by offering valuable insights into the potential functional properties and potential risks associated with these molecules (Solanki et al., 2021). In general, the scoring of results from these tools is based on specific algorithms and analysis of sequence-based features, physicochemical properties, and amino acid compositions. Higher scores or likelihood values obtained from these tools generally indicate a higher probability or potential for the parameter being predicted (Timmons & Hewage, 2020). However, it is important to note that the interpretation of scores may vary depending on the specific tool being used and the thresholds established by its developers. It is recommended to consider the recommended cutoffs or thresholds provided by each tool to accurately determine the predicted parameter.

Experimental Application of In-Silico Prediction Tools

Several research revealed magnificent features of using in-silico prediction tools in easing modification of existed CPPs or in discovering novel CPPs which play significant role in advancement efficient drug delivery system of different cargos as explained in **Table 4**.

Table 4. Experimental Application of In Silico Prediction Tools for CPPs

Study	Experimental Application	Key Findings	Reference
Identification of novel Peptide P1 from MARCKS Protein Phosphorylation Site domain as a new potential CPP candidate that can efficiently internalize into various cell lines and deliver plasmid DNA into cultured cells	Bioinformatic prediction and wet-lab validation	Peptide P1 derived from the MARCKS protein phosphorylation site domain identified as a potential CPP candidate. In-silico prediction tools and wet-lab validation supported its potential as a novel CPP.	(Chen et al., 2021)
Discovery of novel CPPs in SARS-CoV-2 Proteome for vaccine and drug delivery	In silico analyses	Wide range of novel and potent CPPs identified in the SARS-CoV-2 proteome using in silico analyses. This novel CPP can be further evaluated for DNA delivery in vitro and in vivo in future.	(Kardani & Bolhassani, 2021)
Translocating peptides of biomedical Interest obtained from the spike (S) Glycoprotein of the SARS-CoV-2	In-silico and vitro analysis	Peptides obtained from the SARS-CoV-2 Spike Glycoprotein can be used as novel CPPs due to their ability to translocate eukaryotic lipid bilayers. MD simulations provided insights into peptide-bilayer interactions. Identified three potential new sequences (AHB-1, AHB-2, and AHB-3) with superior internalization and endosome escape capacities while maintaining cell viability as well as validated with in vitro analysis	(Henao et al., 2022)
Molecular Mechanism and Energetics of Translocation of Arginine-Rich CPPs through Membranes	Free energy calculations and MD simulation	Translocation of arginine-rich CPPs is likely associated with water-pore formation. Lower free-energy barrier observed along pore path compared to pore-free path. Translocation rate estimated using pore-formation rate-estimate. Thermodynamic evidence for pore-assisted translocation model	(Huang & García, 2013)
Difference in Free Energy Profiles of Three CPPs with Model Lipid Bilayer	Molecular dynamics simulations	The interaction between CPPs and a model lipid bilayer (DOPC) to understand the mechanism of how they penetrate the cell membrane. The study revealed that peptide-lipid interaction at the lipid-water interface has a direct correlation with the penetration efficiency of peptides across the lipid bilayer.	(Her Choong & Keat Yap, 2021)
Investigation of Internalization Mechanism and Structure-Function Relationships of α-aminoisobutyric acid (Aib) Incorporated CPPs	Molecular dynamics simulations	Aib-incorporated CPPs adopted amphipathic folding, enabling efficient membrane penetration, which is beneficial for the design of CPPs for efficient intracellular delivery. Aib residues played a crucial role in increasing efficiency of intracellular delivery.	(Gimenez-Dejoz & Numata, 2022)

The research findings presented in **Table 4** collectively showcase the potential of using computer-based prediction tools to explore types of CPPs and gain a better understanding of how they interact with cell membranes. These computational approaches offer insights into designing CPPs, understanding their mechanisms of action and discovering applications in drug delivery. This could lead to time and cost savings as well as opportunities to create improved CPPs. Chen et al. (2021) successfully utilized bioinformatic tools to

predict the physical-chemical properties, structures and penetration capabilities of a novel peptide called P1, which is derived from the MARCKS protein phosphorylation site domain. This peptide has shown promise as a CPP candidate by efficiently internalizing into various cell lines in a concentration dependent manner. Similarly, Henao et al. (2022) and his team provided successful work involved using computer-based tools to design CPPs based on the Spike (S) Glycoprotein of SARS- CoV-2. These CPPs were experimentally tested by incorporating them into chitosan nanoparticles (CNPs) for drug delivery purposes. The results demonstrated penetration with the ability to escape endosomes, biocompatibility and non-toxicity, toward Vero cells. Moreover, using computational tools, Kardani and Bolhassani (2021) have made progress in identifying CPPs in the SARS-CoV-2 Proteome. This discovery holds promise for applications, in vaccine development and drug delivery.

Furthermore Huang and García (2013) and colleagues conducted an investigation into the mechanism and energy aspects of arginine rich CPPs while they translocate through lipid membranes. The researchers employed advanced umbrella MD simulations and the weighted histogram analysis method (WHAM) to calculate the energies involved in translocating a cyclic Arginine 9 peptide along two pathways: one involving the formation of a water pore and another without a pore. Their study indicated that translocation is likely associated with water pore formation since the free energy barrier along the pathway was significantly lower compared to the pore free pathway.

Another noteworthy study by Her Choong and Keat Yap (2021) and their team explored differences in energy profiles among three established CPPs (TAT, CPP1 and CPP9) which have been experimentally validated for their efficient cell penetration as demonstrated by Qian et al. (2016) Using a model lipid bilayer (DOPC) and enhanced MD pulling simulations. This study further highlighted the significance of understanding peptide lipid interactions and their impact on peptide permeation through bilayers. By employing dynamics pulling simulations and umbrella sampling techniques they gained insights into how peptide lipid interactions, at the interface shared by lipids and water affect peptide's ability to permeate through lipid bilayers. By examining CPPs, such, as TAT, CPP1 and CPP9 we were able to compare their behavior and assess their ability to penetrate membranes. This research adds to our understanding of how to design CPPs with properties for applications, in drug delivery.

Moreover Gimenez-Dejoz and Numata (2022) and his research team in their study used computer-based tools, like MD simulation to explore the impact of adding aminoisobutyric acid (Aib) to CPPs. Through their investigation they found that incorporating Aib led to an increase in the structure of these CPPs and improved their ability to penetrate cells.

Ultimately incorporating computer-based tools in CPP research has the potential to transform the field of drug delivery and biomedicine completely. As we continue to harness the power of predictions, we can make groundbreaking strides in discovering targeted and effective therapeutic approaches based on CPPs. Combining bioinformatics tools with validation in wet labs can pave the way for a future where CPPs play a crucial role in delivering life-changing treatments with unparalleled precision and efficiency. The collaboration between computer-based predictions and wet lab experiments will foster innovation and optimize resources. Accelerate the development of CPP-based therapies that will benefit patients and healthcare systems alike.

Future Outlook

As computational methods continue to advance the prediction of CPPs is expected to make progress. One exciting area for research involves integrating omics data, such as genomics, proteomics and metabolomics. This integration can improve the accuracy and specificity of CPP predictions by providing an understanding of how CPPs behave and interact with biological components. Ultimately this will lead to predictions regarding CPP activity and how they are taken up by cells. Additionally, the utilization of intelligence (AI) and machine learning algorithms shows promise in refining CPP prediction models. AI- driven approaches have the capability to analyze datasets and identify patterns. This facilitates the development of prediction tools that are sophisticated in their ability to predict CPP behavior. By leveraging AI's power, researchers can uncover sequence motifs, structural features and physicochemical properties that govern CPP behavior. Ultimately this advances our knowledge of how CPPs function. Furthermore, validating and refining in silico predictions is crucial by integrating dynamics (MD) simulations with validation. Combining insights from computational models with laboratory-based experiments helps bridge the gap between theory and reality. This synergy provides a foundation for designing and optimizing CPPs while enhancing prediction reliability. Moreover, it facilitates translating in silico findings into applications

Conclusion

In conclusion, advances in computational tools, bioinformatics, and molecular dynamics simulations

have significantly enhanced our understanding of CPP structural and functional properties. Machine learning-based web servers enable accurate prediction of key CPP parameters, including toxicity, immunogenicity, and hemolytic effects, streamlining candidate screening and reducing experimental burden. MD simulations further illuminate CPP translocation mechanisms across lipid bilayers, guiding the design of efficient delivery systems. Despite these advances, experimental validation remains essential to ensure the accuracy of in silico predictions. Overall, these computational approaches accelerate the discovery of novel, effective CPPs for drug delivery applications.

References

Agrawal, P., Bhalla, S., Usmani, S. S., Singh, S., Chaudhary, K., Raghava, G. P. S., & Gautam, A. (2016). CPPsite 2.0: A repository of experimentally validated cell-penetrating peptides. *Nucleic Acids Research*, *44*(D1). https://doi.org/10.1093/nar/gkv1266

Allen, J., & Pellois, J. P. (2022). Hydrophobicity is a key determinant in the activity of arginine-rich cell penetrating peptides. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-20425-y

Almagro Armenteros, J. J., Sønderby, C. K., Sønderby, S. K., Nielsen, H., & Winther, O. (2017). DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics (Oxford, England)*, *33*(21). https://doi.org/10.1093/bioinformatics/btx431

Angell, Y., Holford, M., & Moos, W. H. (2018). Building on Success: A Bright Future for Peptide Therapeutics. *Protein & Peptide Letters*, *25*(12). https://doi.org/10.2174/0929866525666181114155542

Bagos, P. G., Liakopoulos, T. D., Spyropoulos, I. C., & Hamodrakas, S. J. (2004). A Hidden Markov Model method, capable of predicting and discriminating β-barrel outer membrane proteins. *BMC Bioinformatics*, *5*. https://doi.org/10.1186/1471-2105-5-29

Chaudhary, K., Kumar, R., Singh, S., Tuknait, A., Gautam, A., Mathur, D., Anand, P., Varshney, G. C., & Raghava, G. P. S. (2016). A web server and mobile app for computing hemolytic potency of peptides. *Scientific Reports*, *6*. https://doi.org/10.1038/srep22843

Chen, L., Guo, X., Wang, L., Geng, J., Wu, J., Hu, B., Wang, T., Li, J., Liu, C., & Wang, H. (2021). In silico identification and experimental validation of cellular uptake by a new cell penetrating peptide P1 derived from MARCKS. *Drug Delivery*, *28*(1). https://doi.org/10.1080/10717544.2021.1960922

De Cena, G. L., Scavassa, B. V., & Conceição, K. (2022). In Silico Prediction of Anti-Infective and Cell-Penetrating Peptides from Thalassophryne nattereri Natterin Toxins. *Pharmaceuticals*, *15*(9). https://doi.org/10.3390/ph15091141

de Oliveira, E. C. L., Santana, K., Josino, L., Lima e Lima, A. H., & de Souza de Sales Júnior, C. (2021). Predicting cell-penetrating peptides using machine learning algorithms and navigating in their chemical space. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-87134-w

Derakhshankhah, H., & Jafari, S. (2018). Cell penetrating peptides: A concise review with emphasis on biomedical applications. In *Biomedicine and Pharmacotherapy* (Vol. 108). https://doi.org/10.1016/j.biopha.2018.09.097

Derossi, D., Calvet, S., Trembleau, A., Brunissen, A., Chassaing, G., & Prochiantz, A. (1996). Cell internalization of the third helix of the antennapedia homeodomain is receptor-independent. *Journal of Biological Chemistry*, *271*(30). https://doi.org/10.1074/jbc.271.30.18188

Deshayes, S., Decaffmeyer, M., Brasseur, R., & Thomas, A. (2008). Structural polymorphism of two CPP: An important parameter of activity. *Biochimica et Biophysica Acta - Biomembranes*, *1778*(5). https://doi.org/10.1016/j.bbamem.2008.01.027

Dimitrov, I., Bangov, I., Flower, D. R., & Doytchinova, I. (2014). AllerTOP v.2 - A server for in silico prediction of allergens. *Journal of Molecular Modeling*, *20*(6). https://doi.org/10.1007/s00894-014-2278-5

Dimitrov, I., Naneva, L., Doytchinova, I., & Bangov, I. (2014). AllergenFP: Allergenicity prediction by descriptor fingerprints. *Bioinformatics*, *30*(6). https://doi.org/10.1093/bioinformatics/btt619

Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC*

Bioinformatics, *8*. https://doi.org/10.1186/1471-2105-8-4

Duchardt, F., Fotin-Mleczek, M., Schwarz, H., Fischer, R., & Brock, R. (2007). A comprehensive model for the cellular uptake of cationic cell-penetrating peptides. *Traffic*, *8*(7). https://doi.org/10.1111/j.1600-0854.2007.00572.x

El-Andaloussi, S., Holm, T., & Langel, U. (2005). Cell-Penetrating Peptides: Mechanisms and Applications. *Current Pharmaceutical Design*, *11*(28). https://doi.org/10.2174/138161205774580796

Erazo-Oliveras, A., Muthukrishnan, N., Baker, R., Wang, T. Y., & Pellois, J. P. (2012). Improving the endosomal escape of cell-penetrating peptides and their cargos: Strategies and challenges. In *Pharmaceuticals* (Vol. 5, Number 11). https://doi.org/10.3390/ph5111177

Feger, G., Angelov, B., & Angelova, A. (2020). Prediction of Amphiphilic Cell-Penetrating Peptide Building Blocks from Protein-Derived Amino Acid Sequences for Engineering of Drug Delivery Nanoassemblies. *Journal of Physical Chemistry B*, *124*(20). https://doi.org/10.1021/acs.jpcb.0c01618

Fisinger, S., Serrano, L., & Lacroix, E. (2001). Computational estimation of specific side chain interaction energies in α helices. *Protein Science*, *10*(4). https://doi.org/10.1110/ps.34901

Foged, C., & Nielsen, H. M. (2008). Cell-penetrating peptides for drug delivery across membrane barriers. In *Expert Opinion on Drug Delivery* (Vol. 5, Number 1). https://doi.org/10.1517/17425247.5.1.105

Fosgerau, K., & Hoffmann, T. (2015). Peptide therapeutics: Current status and future directions. In *Drug Discovery Today* (Vol. 20, Number 1). https://doi.org/10.1016/j.drudis.2014.10.003

Fu, X., Cai, L., Zeng, X., & Zou, Q. (2020). StackCPPred: A stacking and pairwise energy content-based prediction of cell-penetrating peptides and their uptake efficiency. *Bioinformatics*, *36*(10). https://doi.org/10.1093/bioinformatics/btaa131

Fu, X., Ke, L., Cai, L., Chen, X., Ren, X., & Gao, M. (2019). Improved Prediction of Cell-Penetrating Peptides via Effective Orchestrating Amino Acid Composition Feature Representation. *IEEE Access*, *7*. https://doi.org/10.1109/ACCESS.2019.2952738

Futaki, S. (2005). Membrane-permeable arginine-rich peptides and the translocation mechanisms. In *Advanced Drug Delivery Reviews* (Vol. 57, Number 4 SPEC.ISS.). https://doi.org/10.1016/j.addr.2004.10.009

Futaki, S., Arafiles, J. V. V., & Hirose, H. (2020). Peptide-assisted intracellular delivery of biomacromolecules. In *Chemistry Letters* (Vol. 49, Number 9). https://doi.org/10.1246/cl.200392

Futaki, S., Nakase, I., Suzuki, T., Youjun, Z., & Sugiura, Y. (2002). Translocation of branched-chain arginine peptides through cell membranes: Flexibility in the spatial disposition of positive charges in membrane-permeable peptides. *Biochemistry*, *41*(25). https://doi.org/10.1021/bi0256173

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*. https://doi.org/10.1385/1-59259-890-0:571

Gautam, A., Chaudhary, K., Kumar, R., Sharma, A., Kapoor, P., Tyagi, A., & Raghava, G. P. S. (2013). In silico approaches for designing highly effective cell penetrating peptides. *Journal of Translational Medicine*, *11*(1). https://doi.org/10.1186/1479-5876-11-74

Gautier, R., Douguet, D., Antonny, B., & Drin, G. (2008). HELIQUEST: A web server to screen sequences with specific α-helical properties. *Bioinformatics*, *24*(18). https://doi.org/10.1093/bioinformatics/btn392

Gazit, E., Shai, Y., Lee, W. J., & Brey, P. T. (1994). Mode of Action of the Antibacterial Cecropin B2: A Spectrofluorometric Study. *Biochemistry*, *33*(35). https://doi.org/10.1021/bi00201a016

Gimenez-Dejoz, J., & Numata, K. (2022). Molecular dynamics study of the internalization of cell-penetrating peptides containing unnatural amino acids across membranes. *Nanoscale Advances*, *4*(2). https://doi.org/10.1039/d1na00674f

Gräslund, A., Madani, F., Lindberg, S., Langel, Ü., & Futaki, S. (2011). Mechanisms of cellular uptake of cell-penetrating peptides. In *Journal of Biophysics*. https://doi.org/10.1155/2011/414729

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., & Raghava, G. P. S. (2013). In Silico Approach for Predicting Toxicity of Peptides and Proteins. *PLoS ONE*, *8*(9). https://doi.org/10.1371/journal.pone.0073957

Heitz, F., Morris, M. C., & Divita, G. (2009). Twenty years of cell-penetrating peptides: From molecular mechanisms to therapeutics. In *British Journal of Pharmacology* (Vol. 157, Number 2). https://doi.org/10.1111/j.1476-5381.2009.00057.x

Henao, M. C., Ocasion, C., Puentes, P. R., González-Melo, C., Quezada, V., Cifuentes, J., Yepes, A., Burgos, J. C., Cruz, J. C., & Reyes, L. H. (2022). Translocating Peptides of Biomedical Interest Obtained from the Spike (S) Glycoprotein of the SARS-CoV-2. *Membranes*, *12*(6). https://doi.org/10.3390/membranes12060600

Henriques, S. T., Melo, M. N., & Castanho, M. A. R. B. (2006). Cell-penetrating peptides and antimicrobial peptides: How different are they? In *Biochemical Journal* (Vol. 399, Number 1). https://doi.org/10.1042/BJ20061100

Her Choong, F., & Keat Yap, B. (2021). Cell-Penetrating Peptides: Correlation between Peptide-Lipid Interaction and Penetration Efficiency. *ChemPhysChem*, *22*(5). https://doi.org/10.1002/cphc.202000873

Herce, H. D., & Garcia, A. E. (2007). Molecular dynamics simulations suggest a mechanism for translocation of the HIV-1 TAT peptide across lipid membranes. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(52). https://doi.org/10.1073/pnas.0706574105

Holton, T. A., Pollastri, G., Shields, D. C., & Mooney, C. (2013). CPPpred: Prediction of cell penetrating peptides. *Bioinformatics*, *29*(23). https://doi.org/10.1093/bioinformatics/btt518

Huang, K., & García, A. E. (2013). Free energy of translocating an arginine-rich cell-penetrating peptide across a lipid bilayer suggests pore formation. *Biophysical Journal*, *104*(2). https://doi.org/10.1016/j.bpj.2012.10.027

Jones, A. T., & Sayers, E. J. (2012). Cell entry of cell penetrating peptides: Tales of tails wagging dogs. In *Journal of Controlled Release* (Vol. 161, Number 2). https://doi.org/10.1016/j.jconrel.2012.04.003

Kahsay, R. Y., Gao, G., & Liao, L. (2005). An improved hidden Markov model for transmembrane protein detection and topology prediction and its applications to complete genomes. *Bioinformatics*, *21*(9). https://doi.org/10.1093/bioinformatics/bti303

Kalafatovic, D., & Giralt, E. (2017). Cell-penetrating peptides: Design strategies beyond primary structure and amphipathicity. In *Molecules* (Vol. 22, Number 11). https://doi.org/10.3390/molecules22111929

Kang, Z., Ding, G., Meng, Z., & Meng, Q. (2019). The rational design of cell-penetrating peptides for application in delivery systems. In *Peptides* (Vol. 121). https://doi.org/10.1016/j.peptides.2019.170149

Kardani, K., & Bolhassani, A. (2021). Exploring novel and potent cell penetrating peptides in the proteome of SARS-COV-2 using bioinformatics approaches. *PLoS ONE*, *16*(2 February). https://doi.org/10.1371/journal.pone.0247396

Kardani, K., Milani, A., H. Shabani, S., & Bolhassani, A. (2019). Cell penetrating peptides: the potent multi-cargo intracellular carriers. In *Expert Opinion on Drug Delivery* (Vol. 16, Number 11). https://doi.org/10.1080/17425247.2019.1676720

Kauffman, W. B., Fuselier, T., He, J., & Wimley, W. C. (2015). Mechanism Matters: A Taxonomy of Cell Penetrating Peptides. In *Trends in Biochemical Sciences* (Vol. 40, Number 12). https://doi.org/10.1016/j.tibs.2015.10.004

Khairkhah, N., Namvar, A., & Bolhassani, A. (2023). Application of Cell Penetrating Peptides as a Promising Drug Carrier to Combat Viral Infections. In *Molecular Biotechnology*. https://doi.org/10.1007/s12033-023-00679-1

Klein, M. J., Schmidt, S., Wadhwani, P., Bürck, J., Reichert, J., Afonin, S., Berditsch, M., Schober, T., Brock, R., Kansy, M., & Ulrich, A. S. (2017). Lactam-Stapled Cell-Penetrating Peptides: Cell

Uptake and Membrane Binding Properties. *Journal of Medicinal Chemistry*, *60*(19). https://doi.org/10.1021/acs.jmedchem.7b00813

Krogh, A., Larsson, B., Von Heijne, G., & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology*, *305*(3). https://doi.org/10.1006/jmbi.2000.4315

Kuriakose, A., Chirmule, N., & Nair, P. (2016). Immunogenicity of Biotherapeutics: Causes and Association with Posttranslational Modifications. In *Journal of Immunology Research* (Vol. 2016). https://doi.org/10.1155/2016/1298473

Langel, Ü. (2021). Cell-penetrating peptides and transportan. In *Pharmaceutics* (Vol. 13, Number 7). https://doi.org/10.3390/pharmaceutics13070987

Lee, H. M., Ren, J., Tran, K. M., Jeon, B. M., Park, W. U., Kim, H., Lee, K. E., Oh, Y., Choi, M., Kim, D. S., & Na, D. (2021). Identification of efficient prokaryotic cell-penetrating peptides with applications in bacterial biotechnology. *Communications Biology*, *4*(1). https://doi.org/10.1038/s42003-021-01726-w

Lee, M. T., Hung, W. C., Chen, F. Y., & Huang, H. W. (2005). Many-body effect of antimicrobial peptides: On the correlation between lipid's spontaneous curvature and pore formation. *Biophysical Journal*, *89*(6). https://doi.org/10.1529/biophysj.105.068080

Li, H., Tsui, T. Y., & Ma, W. (2015). Intracellular delivery of molecular cargo using cell-penetrating peptides and the combination strategies. In *International Journal of Molecular Sciences* (Vol. 16, Number 8). https://doi.org/10.3390/ijms160819518

Liu, P., Ding, Y., Rong, Y., & Chen, D. (2022). Prediction of cell penetrating peptides and their uptake efficiency using random forest-based feature selections. *AIChE Journal*, *68*(9). https://doi.org/10.1002/aic.17781

Liu, Y., Zhao, Z., & Li, M. (2022). Overcoming the cellular barriers and beyond: Recent progress on cell penetrating peptide modified nanomedicine in combating physiological and pathological barriers. In *Asian Journal of Pharmaceutical Sciences* (Vol. 17, Number 4). https://doi.org/10.1016/j.ajps.2022.05.002

Lomize, A. L., & Pogozheva, I. (2018). Modeling of Peptide Folding and Translocation across Membranes. *Biophysical Journal*, *114*(3). https://doi.org/10.1016/j.bpj.2017.11.1545

Magnan, C. N., Zeller, M., Kayala, M. A., Vigil, A., Randall, A., Felgner, P. L., & Baldi, P. (2010). High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics*, *26*(23). https://doi.org/10.1093/bioinformatics/btq551

Manavalan, B., & Patra, M. C. (2022). MLCPP 2.0: An Updated Cell-penetrating Peptides and Their Uptake Efficiency Predictor. *Journal of Molecular Biology*, *434*(11). https://doi.org/10.1016/j.jmb.2022.167604

Manavalan, B., Subramaniyam, S., Shin, T. H., Kim, M. O., & Lee, G. (2018). Machine-Learning-Based Prediction of Cell-Penetrating Peptides and Their Uptake Efficiency with Improved Accuracy. *Journal of Proteome Research*, *17*(8). https://doi.org/10.1021/acs.jproteome.8b00148

Muñoz, V., & Serrano, L. (1994). Elucidating the folding problem of helical peptides using empirical parameters. *Nature Structural Biology*, *1*(6). https://doi.org/10.1038/nsb0694-399

Muñoz, V., & Serrano, L. (1995). Elucidating the folding problem of helical peptides using empirical parameters. II†. Helix macrodipole effects and rational modification of the helical content of natural peptides. *Journal of Molecular Biology*, *245*(3). https://doi.org/10.1006/jmbi.1994.0023

Nakase, I., Niwa, M., Takeuchi, T., Sonomura, K., Kawabata, N., Koike, Y., Takehashi, M., Tanaka, S., Ueda, K., Simpson, J. C., Jones, A. T., Sugiura, Y., & Futaki, S. (2004). Cellular uptake of arginine-rich peptides: Roles for macropinocytosis and actin rearrangement. *Molecular Therapy*, *10*(6). https://doi.org/10.1016/j.ymthe.2004.08.010

Nielsen, M., Lundegaard, C., Blicher, T., Lamberth, K., Harndahl, M., Justesen, S., Røder, G., Peters, B., Sette, A., Lund, O., & Buus, S. (2007). NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS ONE*,

2(8). https://doi.org/10.1371/journal.pone.0000796

Ouyang, J., Sheng, Y., & Wang, W. (2022). Recent Advances of Studies on Cell-Penetrating Peptides Based on Molecular Dynamics Simulations. In *Cells* (Vol. 11, Number 24). https://doi.org/10.3390/cells11244016

Panuccio, A., Bicego, M., & Murino, V. (2002). A hidden Markov model-based approach to sequential data clustering. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *2396*. https://doi.org/10.1007/3-540-70659-3_77

Park, H., Park, J.-H., Kim, M. S., Cho, K., & Shin, J.-M. (2023). In Silico Screening and Optimization of Cell-Penetrating Peptides Using Deep Learning Methods. *Biomolecules*, *13*(3). https://doi.org/10.3390/biom13030522

Patel, S. G., Sayers, E. J., He, L., Narayan, R., Williams, T. L., Mills, E. M., Allemann, R. K., Luk, L. Y. P., Jones, A. T., & Tsai, Y. H. (2019). Cell-penetrating peptide sequence and modification dependent uptake and subcellular distribution of green florescent protein in different cell lines. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-42456-8

Patel, S., Kim, J., Herrera, M., Mukherjee, A., Kabanov, A. V., & Sahay, G. (2019). Brief update on endocytosis of nanomedicines. In *Advanced Drug Delivery Reviews* (Vol. 144). https://doi.org/10.1016/j.addr.2019.08.004

Plisson, F., Ramírez-Sánchez, O., & Martínez-Hernández, C. (2020). Machine learning-guided discovery and design of non-hemolytic peptides. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-73644-6

Porosk, L., Gaidutšik, I., & Langel, Ü. (2021). Approaches for the discovery of new cell-penetrating peptides. In *Expert Opinion on Drug Discovery* (Vol. 16, Number 5). https://doi.org/10.1080/17460441.2021.1851187

Porosk, L., Põhako, K., Arukuusk, P., & Langel, Ü. (2021). Cell-Penetrating Peptides Predicted From CASC3, AKIP1, and AHRR Proteins. *Frontiers in Pharmacology*, *12*. https://doi.org/10.3389/fphar.2021.716226

Pouny, Y., Rapaport, D., Shai, Y., Mor, A., & Nicolas, P. (1992). Interaction of Antimicrobial Dermaseptin and its Fluorescently Labeled Analogs with Phospholipid Membranes. *Biochemistry*, *31*(49). https://doi.org/10.1021/bi00164a017

Prochiantz, A. (1996). Getting hydrophilic compounds into cells: Lessons from homeopeptides. Commentary. In *Current Opinion in Neurobiology* (Vol. 6, Number 5). https://doi.org/10.1016/S0959-4388(96)80095-X

Qian, Z., Martyna, A., Hard, R. L., Wang, J., Appiah-Kubi, G., Coss, C., Phelps, M. A., Rossman, J. S., & Pei, D. (2016). Discovery and Mechanism of Highly Efficient Cyclic Cell-Penetrating Peptides. *Biochemistry*, *55*(18). https://doi.org/10.1021/acs.biochem.6b00226

Qiang, X., Zhou, C., Ye, X., Du, P. F., Su, R., & Wei, L. (2018). CPPred-FL: A sequence-based predictor for large-scale identification of cell-penetrating peptides by feature representation learning. *Briefings in Bioinformatics*, *21*(1). https://doi.org/10.1093/bib/bby091

Richard, J. P., Melikov, K., Vives, E., Ramos, C., Verbeure, B., Gait, M. J., Chernomordik, L. V., & Lebleu, B. (2003). Cell-penetrating peptides: A reevaluation of the mechanism of cellular uptake. *Journal of Biological Chemistry*, *278*(1). https://doi.org/10.1074/jbc.M209548200

Sanders, W. S., Johnston, C. I., Bridges, S. M., Burgess, S. C., & Willeford, K. O. (2011). Prediction of Cell Penetrating Peptides by Support Vector Machines. *PLoS Computational Biology*, *7*(7). https://doi.org/10.1371/journal.pcbi.1002101

Sarko, D., Beijer, B., Boy, R. G., Nothelfer, E. M., Leotta, K., Eisenhut, M., Altmann, A., Haberkorn, U., & Mier, W. (2010). The pharmacokinetics of cell-penetrating peptides. *Molecular Pharmaceutics*, *7*(6). https://doi.org/10.1021/mp100223d

Schissel, C. K., Mohapatra, S., Wolfe, J. M., Fadzen, C. M., Bellovoda, K., Wu, C. L., Wood, J. A., Malmberg, A. B., Loas, A., Gómez-Bombarelli, R., & Pentelute, B. L. (2021). Deep learning to design nuclear-targeting abiotic miniproteins. *Nature Chemistry*, *13*(10). https://doi.org/10.1038/s41557-021-00766-3

Shankar, G., Arkin, S., Cocea, L., Devanarayan, V., Kirshner, S., Kromminga, A., Quarmby, V.,

Richards, S., Schneider, C. K., Subramanyam, M., Swanson, S., Verthelyi, D., & Yim, S. (2014). Assessment and reporting of the clinical immunogenicity of therapeutic proteins and peptides - Harmonized terminology and tactical recommendations. In *AAPS Journal* (Vol. 16, Number 4). https://doi.org/10.1208/s12248-014-9599-2

Singh, H., Ansari, H. R., & Raghava, G. P. S. (2013). Improved Method for Linear B-Cell Epitope Prediction Using Antigen's Primary Sequence. *PLoS ONE*, *8*(5). https://doi.org/10.1371/journal.pone.0062216

Solanki, V., Tiwari, M., & Tiwari, V. (2021). Immunoinformatic approach to design a multiepitope vaccine targeting non-mutational hotspot regions of structural and non-structural proteins of the SARS CoV2. *PeerJ*, *9*. https://doi.org/10.7717/peerj.11126

Tang, H., Su, Z. D., Wei, H. H., Chen, W., & Lin, H. (2016). Prediction of cell-penetrating peptides with feature selection techniques. *Biochemical and Biophysical Research Communications*, *477*(1). https://doi.org/10.1016/j.bbrc.2016.06.035

Taylor, C. T., Furuta, G. T., Synnestvedt, K., & Colgan, S. P. (2000). Phosphorylation-dependent targeting of cAMP response element binding protein to the ubiquitin/proteasome pathway in hypoxia. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(22). https://doi.org/10.1073/pnas.220211797

Ter-Avetisyan, G., Tünnemann, G., Nowak, D., Nitschke, M., Hermann, A., Drab, M., & Cardoso, M. C. (2009). Cell entry of arginine-rich peptides is independent of endocytosis. *Journal of Biological Chemistry*, *284*(6). https://doi.org/10.1074/jbc.M805550200

Thévenet, P., Shen, Y., Maupetit, J., Guyon, F., Derreumaux, P., & Tufféry, P. (2012). PEP-FOLD: An updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Research*, *40*(W1). https://doi.org/10.1093/nar/gks419

Thumuluri, V., Almagro Armenteros, J. J., Johansen, A. R., Nielsen, H., & Winther, O. (2022). DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Research*, *50*(W1). https://doi.org/10.1093/nar/gkac278

Timmons, P. B., & Hewage, C. M. (2020). HAPPENN is a novel tool for hemolytic activity prediction for therapeutic peptides which employs neural networks. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-67701-3

Tünnemann, G., Ter-Avetisyan, G., Martin, R. M., Stöckl, M., Herrmann, A., & Cardoso, M. C. (2008). Live-cell analysis of cell penetration ability and toxicity of oligo-arginines. *Journal of Peptide Science*, *14*(4). https://doi.org/10.1002/psc.968

Veldhoen, S., Laufer, S. D., & Restle, T. (2008). Recent developments in peptide-based nucleic acid delivery. In *International Journal of Molecular Sciences* (Vol. 9, Number 7). https://doi.org/10.3390/ijms9071276

Vita, R., Overton, J. A., Greenbaum, J. A., Ponomarenko, J., Clark, J. D., Cantrell, J. R., Wheeler, D. K., Gabbard, J. L., Hix, D., Sette, A., & Peters, B. (2015). The immune epitope database (IEDB) 3.0. *Nucleic Acids Research*, *43*(D1). https://doi.org/10.1093/nar/gku938

Wei, L., Tang, J., & Zou, Q. (2017). SkipCPP-Pred: An improved and promising sequence-based predictor for predicting cell-penetrating peptides. *BMC Genomics*, *18*. https://doi.org/10.1186/s12864-017-4128-1

Wei, L., Xing, P., Su, R., Shi, G., Ma, Z. S., & Zou, Q. (2017). CPPred-RF: A Sequence-based Predictor for Identifying Cell-Penetrating Peptides and Their Uptake Efficiency. *Journal of Proteome Research*, *16*(5). https://doi.org/10.1021/acs.jproteome.7b00019

Win, T. S., Malik, A. A., Prachayasittikul, V., S Wikberg, J. E., Nantasenamat, C., & Shoombuatong, W. (2017). HemoPred: A web server for predicting the hemolytic activity of peptides. *Future Medicinal Chemistry*, *9*(3). https://doi.org/10.4155/fmc-2016-0188

Xie, J., Bi, Y., Zhang, H., Dong, S., Teng, L., Lee, R. J., & Yang, Z. (2020). Cell-Penetrating Peptides in Diagnosis and Treatment of Human Diseases: From Preclinical Research to Clinical Application. In *Frontiers in Pharmacology* (Vol. 11). https://doi.org/10.3389/fphar.2020.00697

Yu, C.-S., Lin, C.-J., & Hwang, J.-K. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n -peptide compositions. *Protein Science*, *13*(5). https://doi.org/10.1110/ps.03479604

Zorko, M., & Langel, Ü. (2005). Cell-penetrating peptides: Mechanism and kinetics of cargo delivery. In *Advanced Drug Delivery Reviews* (Vol. 57, Number 4 SPEC.ISS.). https://doi.org/10.1016/j.addr.2004.10.010