

Article

In Silico Prediction of Anti-Infective and Cell-Penetrating Peptides from *Thalassophryne nattereri* Natterin Toxins

Gabrielle Lupeti De Cena, Bruna Vitória Scavassa and Katia Conceição *

Laboratory of Peptide Biochemistry, Universidade Federal de São Paulo (UNIFESP),
São José dos Campos 12231-280, Brazil* Correspondence: katia.conceicao@unifesp.br

Abstract: The therapeutic potential of venom-derived peptides, such as bioactive peptides (BAPs), is determined by specificity, stability, and pharmacokinetics properties. BAPs, including anti-infective or antimicrobial peptides (AMPs) and cell-penetrating peptides (CPPs), share several physicochemical characteristics and are potential alternatives to antibiotic-based therapies and drug delivery systems, respectively. This study used in silico methods to predict AMPs and CPPs derived from natterins from the venomous fish *Thalassophryne nattereri*. Fifty-seven BAPs (19 AMPs, 8 CPPs, and 30 AMPs/CPPs) were identified using the web servers CAMP, AMPA, AmpGram, C2Pred, and CellPPD. The physicochemical properties were analyzed using ProtParam, PepCalc, and DispHred tools. The membrane-binding potential and cellular location of each peptide were analyzed using the Boman index by APD3, and TMHMM web servers. All CPPs and two AMPs showed high membrane-binding potential. Fifty-four peptides were located in the plasma membrane. Peptide immunogenicity, toxicity, allergenicity, and ADMET parameters were evaluated using several web servers. Sixteen antiviral peptides and 37 anticancer peptides were predicted using the web servers Meta-iAVP and ACPred. Secondary structures and helical wheel projections were predicted using the PEP-FOLD3 and Heliquist web servers. Fifteen peptides are potential lead compounds and were selected to be further synthesized and tested experimentally in vitro to validate the in silico screening. The use of computer-aided design for predicting peptide structure and activity is fast and cost-effective and facilitates the design of potent therapeutic peptides. The results demonstrate that toxins form a natural biotechnological platform in drug discovery, and the presence of CPP and AMP sequences in toxin families opens new possibilities in toxin biochemistry research.

Keywords: bioactive peptides; antimicrobial peptides; cell-penetrating peptides; in silico prediction; ADMET; hydrophobicity



Citation: De Cena, G.L.; Scavassa, B.V.; Conceição, K. In Silico Prediction of Anti-Infective and Cell-Penetrating Peptides from *Thalassophryne nattereri* Natterin Toxins. *Pharmaceuticals* **2022**, *15*, 1141. <https://doi.org/10.3390/ph15091141>

Academic Editors: Jean Leandro dos Santos and Chung Man Chin

Received: 21 May 2022

Accepted: 6 September 2022

Published: 13 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Animal venoms contain a diverse and complex mixture of bioactive compounds that target various receptors to support the survival of venomous animals [1]. Several drugs derived from animal venoms have been approved by the FDA for human use, while other drugs are in clinical trials [2,3]. Recent advances in genomics and proteomics have improved the biochemical analysis of venoms [4]. The ability to rapidly screen venom compounds using high-throughput technologies and the prediction of new molecules encoded in toxins allows for harnessing the therapeutic potential of animal venoms.

Peptides are key role molecules found in all organisms and play a crucial role in many biological processes [5–8]. The large distribution and functional diversity of peptides increase their therapeutic potential [9–11]. Venom-derived peptides involved in defense and predation have long been exploited for medicinal, agricultural, and biotechnological applications [1,12]. Most of these peptides originate from a limited number of taxa of venomous terrestrial animals. However, several bioactive compounds from fish venoms have been isolated and characterized [13]. Prediction of new bioactive peptides (BAPs)

derived from natterins from the venomous fish *Thalassophryne nattereri* by in silico analysis is the aim of this study.

T. nattereri is responsible for cases of envenomation of fishermen and bathers in the north and northeast of Brazil [14,15]. The most common sites of envenomation are the palm of the hands or soles of the feet [16]. The natterin family of toxins contains five orthologs: natterin 1–4 and -P [17]. Natterins are tissue-kallikrein-like enzymes and aerolysin-like pore-forming toxins responsible for the main toxic effects of *T. nattereri* venom: local edema, excruciating pain, and necrosis [18–20]. The degree of amino acid homology between natterin 1 and 2 is 84%, and these orthologs have 40% identity with natterin 3 and 4 (Figure S1). Natterin P is the shortest ortholog (71 amino acids) and shows 84% identity with the first 55 amino acid residues in the N-terminus of natterin 4 [17,20]. We hypothesize that natterins should be a source of BAPs with antimicrobial and cell-penetrating activity based on their pharmacological profile.

The therapeutic potential of venom-derived BAPs is determined by specificity, stability, and pharmacokinetic properties [21]. Two classes of BAPs—anti-infective or antimicrobial peptides (AMPs) and cell-penetrating peptides (CPPs)—share several physicochemical characteristics and are potential alternatives to antibiotic-based therapies and drug delivery systems, respectively. Since the plasma membrane selectively controls the transport of bioactive substances across cells, there is increased interest in developing novel strategies to overcome this barrier and increase bioavailability. In this context, peptide-based transport systems, such as CPPs, have come into focus, and their efficiency has been demonstrated in multiple applications [22–25].

AMPs are a large class of naturally occurring peptides with antibacterial and/or antifungal activity and can help overcome microbial resistance to conventional antibiotics [26–28]. Fusion of CPPs and AMPs produces multifunctional peptides capable of treating infections, cancer, obesity, and other diseases [29–32]. Thus, concerted efforts are being made to design new AMPs or CPPs [33–37]. Nonetheless, these BAPs have failed clinical trials, underscoring the need to optimize these peptides. In this context, the computer-aided design of BAPs has generated crucial information on the physicochemical characteristics and biological activities of BAPs, allowing analyzing these properties and activities before peptide synthesis. Several methods have been developed to predict AMPs and CPPs and evaluate physicochemical properties [38,39].

AMPs and CPPs can be derived from known protein sequences. However, analyzing the physicochemical properties of proteins using experimental techniques is expensive and laborious. In silico approaches are faster, cheaper, and less laborious, enabling the large-scale screening and identification of BAPs with application in biomedicine and pharmacology [40].

Several BAP prediction tools have been developed using different data features and machine learning methods [34,40,41], and the performance of these tools varies depending on these features and the nature of the training technique. Most prediction methods use single classifier models such as support vector machine (SVM), discriminant analysis, fuzzy k-nearest neighbor, and deep learning. Other methods use decision tree classifiers such as ensemble models and random forests [42].

In silico approaches have facilitated the design of highly effective engineered peptides with cell-penetrating, antimicrobial, and anticancer activity [43–48]. However, as peptides gain ground over small molecule drugs [2,49], some disadvantages must be overcome, including chemical and physical instability [50], high susceptibility to proteolytic degradation [51], short half-life and high clearance [52], slow tissue penetration [53], and high cytotoxicity [53]. In this context, machine learning techniques have been used to screen peptide template libraries based on physicochemical properties and absorption, distribution, metabolism, excretion, and toxicity (ADMET) parameters. This study evaluated the physicochemical and ADMET profiles of newly predicted peptides derived from natterins.

2. Results and Discussion

There has been an increased interest in therapeutic peptides as potential drug candidates [54]. Several studies identified and characterized a wide range of therapeutic peptides, including tumor-homing peptides [55], CPPs [56], AMPs [57], and anticancer peptides (ACPs) [58–60], and used these peptides for treating cancer, diabetes, and cardiovascular diseases. As a result of these efforts, several peptides have entered clinical trials over the past two decades [54]. Nonetheless, only a few peptide-based drugs are used clinically. Therefore, many research groups have focused on computational design based on physicochemical and structural features to produce potent and broad-spectrum peptides [9]. Several computational tools have been used to design peptide-based drug candidates [41]. This study predicted and characterized novel and potent BAPs derived from *T. nattereri* natterins by in silico analysis.

2.1. Identification of Potential Natterin-Derived AMPs and CPPs

Fifty-seven natterin-derived BAPs were identified using the web servers AMPA, CAMP, AmpGram, C2Pred, and CellPPD. These peptides were named according to the original sequence (natterin 1, 2, 3, 4, or P) and the order in which they were identified. For instance, the first peptide derived from natterin 1 was named NATT1_1, the second was named NATT1_2, etc. Some peptide sequences were homologous to more than one natterin. In these cases, the numbering of the source natterin was added to the nomenclature. For instance, the peptide RTYRGGKKTQTTTKGVYRTTQV was the first to be identified as belonging to natterin 1 and 2 and thus was named NATT1.2_1.

All predicted AMPs and CPPs and their respective scores (SVM, RF, or artificial neural networks (ANNs)) and probability scores are listed in Table 1. Nineteen peptides were classified as AMPs, of which seven, three, six, and three belonged to natterin 1, 2, 3, and 4, respectively. Eight CPPs were found, of which one and seven belonged to natterin 2 and 4, respectively. Thirty sequences shared AMP and CPP characteristics, of which five, eleven, four, five, and five sequences belonged to natterin 1, 2, 3, 4, and P, respectively. In the C2Pred web server, peptides with scores of <0.5 and ≥ 0.5 are classified as non-CPPs and CPPs, respectively. Although some peptides were predicted to be CPPs by CellPPD, C2Pred classified them as non-CPPs. For instance, the natterin 3-derived peptide NATT3_10 was classified as CPP and non-CPP using CellPPD (SVM score of 0.1) and C2Pred (score of 0.48551), respectively.

Table 1. Predicted antimicrobial peptides (AMPs) and cell-penetrating peptides (CPPs) and their uptake efficiency.

Peptides		AMP Prediction				CPP Prediction			
Name	Sequence	AA (n)	CAMP	AMPA	AmpGram	C2Pred		CELL PPD	
						Prediction	Probability	Prediction	SVM Score
NATT1_01	TCKTNRIYVGKAY	14	0	AMP	0.750	Non-CPP	0.836	Non-CPP	−0.380
NATT1_02	MRKSTVNNKQCKEVTK	16	0	AMP	0.492	CPP	0.530	Non-CPP	−0.250
NATT1_03	VNKDVIEQTM	10	0.501	-	0.047	Non-CPP	0.942	Non-CPP	−0.780
NATT1_04	DVIEQTMKDV	10	0.549	-	0.005	Non-CPP	0.912	Non-CPP	−0.640
NATT1_05	TESQSYMVTV	10	0.547	-	0.000	CPP	0.756	Non-CPP	−0.820
NATT1.2_01	RTYRGGKKTQTTTKGVYRTTQV	22	0	AMP	0.531	Non-CPP	0.524	Non-CPP	−0.350
NATT1.2_02	STNDETNLHW	10	0.524	-	0.000	Non-CPP	0.732	Non-CPP	−0.780
NATT1.2_03	CKTNRIYVGK	10	0.603	-	0.921	Non-CPP	0.657	Non-CPP	−0.100
NATT1.2_04	KTNRIYVGK	10	0.544	-	0.561	Non-CPP	0.784	Non-CPP	−0.120
NATT1.2_05	LIRTYRGGKK	10	0.699	-	0.544	CPP	0.882	CPP	0.300
NATT1.2_06	IRTYRGGKKT	10	0.613	-	0.541	CPP	0.864	CPP	0.010
NATT1.2_07	RTYRGGKKTQ	10	0.526	-	0.413	CPP	0.537	CPP	0.000
NATT2_01	TCKTNKIYVGKAY	14	0	AMP	0.996	Non-CPP	0.835	Non-CPP	−0.460
NATT2_02	RTYRGGKKTQTTTKGVYRTTQV	22	0	AMP	0.530	CPP	0.655	Non-CPP	−0.340

Table 1. Cont.

Peptides		AMP Prediction				CPP Prediction			
Name	Sequence	AA (n)	CAMP	AMPA	AmpGram	C2Pred		CELL PPD	
						Prediction	Probability	Prediction	SVM Score
NATT2_03	TLRPKLSKPPAK	13	0	AMP	1000	CPP	0.947	CPP	0.630
NATT2_04	TETQSYMVTV	10	0.684	-	0.000	Non-CPP	0.238	Non-CPP	-0.810
NATT2_05	ETQSYMVTVS	10	0.542	-	0.000	Non-CPP	0.238	Non-CPP	-0.710
NATT2_06	TTLRPKLSK	10	0.505	-	0.945	CPP	0.978	CPP	0.300
NATT2_07	TLRPKLSK	10	0.602	-	0.991	CPP	0.952	CPP	0.540
NATT2_08	LRPKLSKPP	10	0.565	-	0.987	CPP	0.952	CPP	0.460
NATT2_09	RPKLSKPPA	10	0.533	-	0.975	CPP	0.929	CPP	0.290
NATT2_10	PKLSKPPAK	10	0.638	-	1000	CPP	0.929	CPP	0.510
NATT2_11	KLKSKPPAK	10	0.627	-	1000	CPP	0.929	CPP	0.510
NATT2_12	LKSKPPAKPA	10	0.573	-	1000	CPP	0.908	CPP	0.100
NATT2_13	KSKPPAKPAG	10	0.529	-	1000	CPP	0.682	CPP	0.200
NATT2_14	SKKPPAKPAGK	10	0.556	-	1000	CPP	0.682	CPP	0.150
NATT2_15	LRPKLSKPPAKPAGK	16	0	-	1000	CPP	0.878	CPP	0.180
NATT3_01	VYVGKKNKYGLGKVHTKHE	18	0	AMP	0.996	Non-CPP	0.186	Non-CPP	-0.520
NATT3_02	MTRTYRNGQKRRTSITGTYRAIQ	23	0	AMP	0.015	CPP	0.838	Non-CPP	-0.220
NATT3_03	YVCSGCSSG	10	0.574	-	0.577	Non-CPP	0.184	Non-CPP	-0.680
NATT3_04	CSCGCSSGFY	10	0.548	-	0.406	Non-CPP	0.204	Non-CPP	-0.650
NATT3_05	HYAYGETEKT	10	0.501	-	0.001	CPP	0.508	Non-CPP	-0.510
NATT3_06	KYGLGKVHTK	10	0.546	-	0.993	Non-CPP	0.294	Non-CPP	-0.120
NATT3_07	PPNHYPVTM	10	0.582	-	0.949	Non-CPP	0.198	Non-CPP	-0.550
NATT3_08	PNHYCPVTMV	10	0.538	-	0.885	Non-CPP	0.246	Non-CPP	-0.410
NATT3_09	TRTYRNGQKR	10	0.531	-	0.168	CPP	0.843	CPP	0.190
NATT3_10	RTYRNGQKRT	10	0.528	-	0.166	Non-CPP	0.486	CPP	0.100
NATT4_01	LYVAKKNKYGLGKL	13	0.772	-	0.989	Non-CPP	0.089	Non-CPP	-0.270
NATT4_02	KACRDLYVAK	10	0	-	0.443	Non-CPP	0.144	CPP	0.030
NATT4_03	KITNVRYNMK	10	0	-	0.045	Non-CPP	0.406	CPP	0.070
NATT4_04	IPFTGRLTRK	10	0	-	0.418	Non-CPP	0.494	CPP	0.140
NATT4_05	PFTGRLTRKY	10	0	-	0.442	CPP	0.751	CPP	0.750
NATT4_06	FTGRLTRKYS	10	0	-	0.358	CPP	0.751	CPP	0.010
NATT4_07	TGRLTRKYSN	10	0	-	0.361	CPP	0.524	CPP	0.030
NATT4_08	GRLTRKYSNG	10	0.519	-	0.406	CPP	0.746	CPP	0.040
NATT4_09	RLTRKYSNGK	10	0	-	0.412	CPP	0.830	CPP	0.160
NATT4_10	KNKYGLGKLHQSS	12	0	AMP	0.989	CPP	0.604	Non-CPP	-0.160
NATT4_11	KANIPFTGRLTRK	13	0.516	-	0.449	CPP	0.702	CPP	0.050
NATT4_12	GRLTRKYSNGKVT	13	0.519	-	0.432	CPP	0.804	Non-CPP	-0.110
NATT4_13	KVTSSSVKGIYKK	13	0.601	-	0.908	Non-CPP	0.231	Non-CPP	-0.050
NATT4_14	VTSSSVKGIYKKV	13	0.508	-	0.971	Non-CPP	0.231	Non-CPP	-0.430
NATT4_15	VKGIYKKVQVGEI	13	0.746	-	0.919	Non-CPP	0.186	Non-CPP	-0.620
NATTP_01	LGQALIPRCRKMP	13	0.609	-	0.986	Non-CPP	0.468	CPP	0.150
NATTP_02	RCRKMPGVKM	10	0	-	0.634	CPP	0.767	CPP	0.010
NATTP_03	QALIPRCRKMPGV	13	0.526	-	0.990	CPP	0.547	Non-CPP	-0.090
NATTP_04	ALIPRCRKMPGVK	13	0.771	-	0.990	CPP	0.547	CPP	0.280
NATTP_05	LIPRCRKMPGVKM	13	0.645	AMP	0.893	CPP	0.563	CPP	0.050
Inference/Reference range			>0.5: AMP	-	>0.5: AMP	>0.5: CPP		SVM score >0: CPP	
			<0.5: non-AMP		<0.5: non-AMP	<0.5: non-CPP		SVM score <0: non-CPP	

The length of the predicted peptides varied from 10 to 23 amino acid residues. In the 1980s, most peptides entering clinical development were less than 10 amino acids long. However, the length of engineered peptides increased over the years due to improvements in chemical synthesis and manufacturing technologies [61,62]. In the current decade, candidate peptides have up to 40 amino acids, suggesting that length is no longer a limitation. Nonetheless, most drug candidates have 10 amino acid residues are still the majority for peptide drug development. In the present study, 63.8% of the peptides presented 10 amino acids.

2.2. Physicochemical Properties and Membrane-Binding Potential

The following physicochemical characteristics were analyzed: net charge, pI, molecular weight (MW), amphipathicity, water solubility, hydrophobicity, hydrophobicity ratio, and charge. CPPs and AMPs are rich in particular amino acids, such as Arg, Trp, Pro, Gly, Cys, and His. The hallmark of these two classes of peptides is an abundance of basic (Arg and Lys) residues and/or Trp. The charge and pI values are shown in Figure 1. Forty-six (81%) peptides were cationic, nine (16%) were anionic, and two (3%) were neutral. The modes of action are determined by the physicochemical features of amino acid residues [63]. The net positive charge and amphipathicity significantly influence the bioactivity of AMPs and most CPPs [27]. The net positive charge affects initial electrostatic interactions with anionic phospholipids and lipopolysaccharides in the plasma membranes of certain pathogens [64]. In turn, mammalian cells, such as red blood cells, are composed primarily of zwitterionic phospholipids in the outer leaflet of their membranes, which are more strongly affected by peptide hydrophobicity than by positive charges [65]. Highly hemolytic peptides interact with phosphatidylcholine, an abundant component of zwitterionic membranes [66]. In contrast, cholesterol inhibits peptide binding in mammalian cell membranes [67].

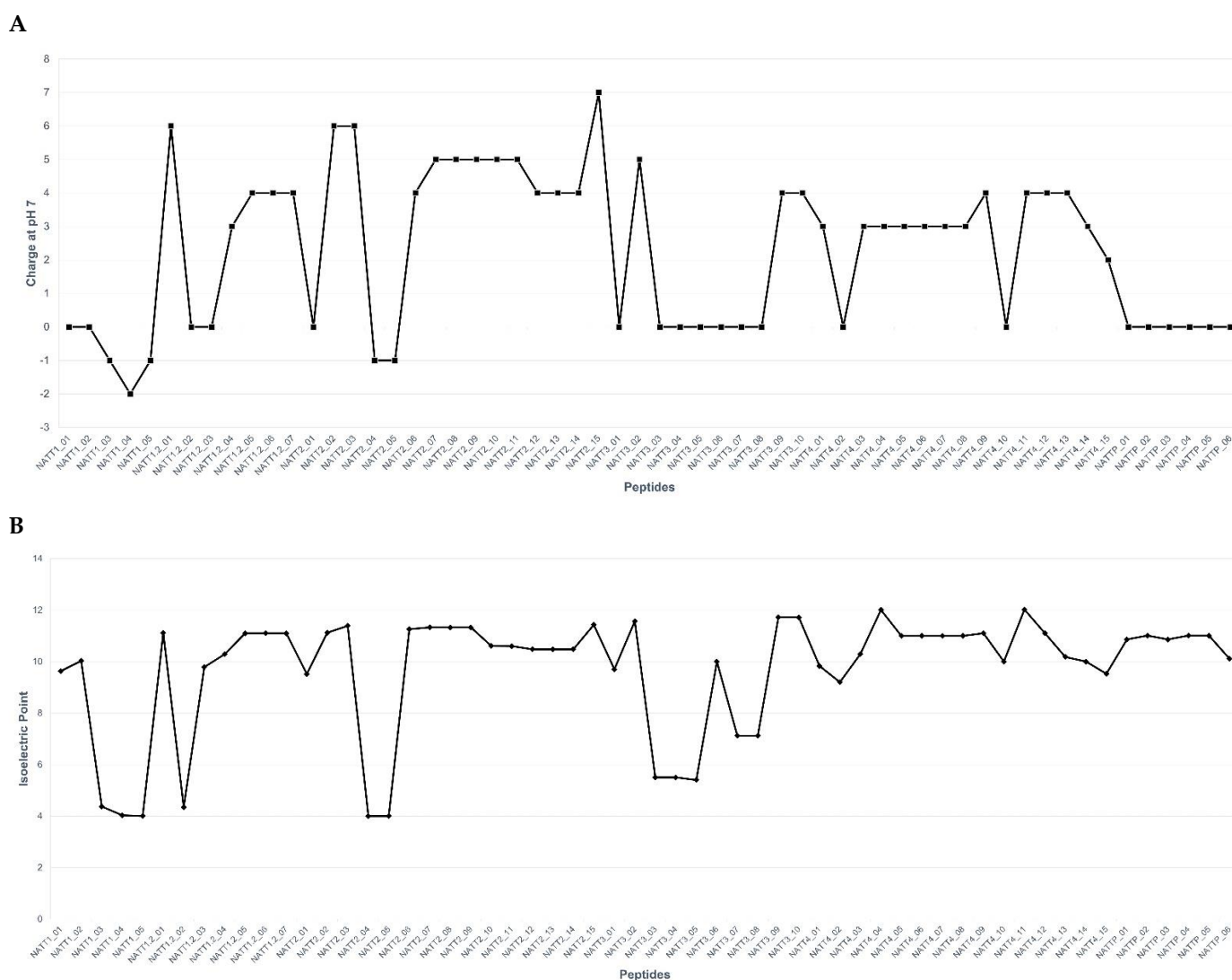


Figure 1. Graphical representation of the physicochemical properties of 57 peptides: (A) total net charge at pH 7.0 and (B) isoelectric point.

In vitro and in vivo studies need controlled and accurate peptide concentration; hence, peptide solubilization is a critical step for successful assays. Consequently, poor peptide

solubilization can introduce experimental errors and lead to experimental failure [21]. In this respect, the solubility of bioactive peptides depends on the molecular length and the number of hydrophobic amino acids (Table 2) [68]. Peptides with a high percentage ($\geq 50\%$) of hydrophobic amino acids are generally partially soluble in aqueous solutions [20,69]. Our results showed that 22% of the peptides were stable (Table S1). The stability of drug candidates is critical for manufacturing the active pharmaceutical ingredient and for enabling formulation of a stable compound. Further, these properties enable producing peptides with different routes of administration, including topical, subcutaneous, and fast intravenous push preparations [70].

Table 2. Amino acid characteristics of predicted bioactive peptides.

Name	Peptides Sequence	MW (g/mol)	Polar Residues + GLY (n/%)	Uncharged Residues + GLY	Charged Residues	Non-Polar Residues (n/%)
NATT1_01	TCKTNRIYVGKGAY	1573.83	8/57.14	THR 2, ASN 1, GLY 2	LYS 2, ARG 1,	6/42.86
NATT1_02	MRKSTVNNKQCKEVTK	1894.24	12/75.00	GLN 1, SER 1, THR 2, ASN 2, GLY 0	LYS 4, ARG 1, GLU 1	4/25.00
NATT1_03	VNKDVIEQTM	1176.35	6/60.00	GLN 1, THR 1, ASN 1, GLY 0	LYS 1, GLU 1, ASP 1,	4/40.00
NATT1_04	DVIEQTMKDV	1177.34	6/60.00	GLN 1, THR 1, GLY 0	LYS 1, GLU 1, ASP 2,	4/40.00
NATT1_05	TESQSYMVTV	1144.26	6/60.00	GLN 1, SER 2, THR 2, GLY 0	GLU 1,	4/40.00
NATT1.2_01	RTYRGGKKTQTTTKGVYRTIQV	2530.87	18/81.82	GLN 2, THR 7, GLY 3	LYS 3, ARG 3,	4/18.18
NATT1.2_02	STNDETNLHW	1216.23	8/80.00	HIS 1, SER 1, THR 2, ASN 2, GLY 0	GLU 1, ASP 1,	2/20.00
NATT1.2_03	CKTNRIYVGK	1181.42	6/60.00	THR 1, ASN 1, GLY 1	LYS 2, ARG 1,	4/40.00
NATT1.2_04	KTNRIYVGKG	1135.33	7/70.00	THR 1, ASN 1, GLY 2	LYS 2, ARG 1,	3/30.00
NATT1.2_05	LIRTYRGGKK	1191.44	7/70.00	THR 1, GLY 2	LYS 2, ARG 2,	3/30.00
NATT1.2_06	IRTYRGGKKT	1179.39	8/80.00	THR 2, GLY 2	LYS 2, ARG 2,	2/20.00
NATT1.2_07	RTYRGGKKTQ	1194.36	9/90.00	GLN 1, THR 2, GLY 2	LYS 2, ARG 2,	1/10.00
NATT2_01	TCKTNKIYVGKGAY	1545.82	8/57.14	THR 2, ASN 1, GLY 2	LYS 3,	6/42.86
NATT2_02	RTYRGGKKTQTTTKGVYRTIQV	2542.92	17/77.27	GLN 2, THR 6, GLY 3	LYS 3, ARG 3,	5/22.73
NATT2_03	TLRPKLKSCKPAK	1494.89	8/61.54	SER 1, THR 1, GLY 0	LYS 5, ARG 1,	5/38.46
NATT2_04	TETQSYMVTV	1158.29	6/60.00	GLN 1, SER 1, THR 3, GLY 0	GLU 1,	4/40.00
NATT2_05	ETQSYMVTVS	1144.26	6/60.00	GLN 1, SER 2, THR 2, GLY 0	GLU 1,	4/40.00
NATT2_06	TTLRPKLKSCK	1171.45	7/70.00	SER 1, THR 2, GLY 0	LYS 3, ARG 1,	3/30.00
NATT2_07	TLRPKLKSCK	1198.52	7/70.00	SER 1, THR 1, GLY 0	LYS 4, ARG 1,	3/30.00
NATT2_08	LRPKLKSCKPAK	1194.53	6/60.00	SER 1, GLY 0	LYS 4, ARG 1,	4/40.00
NATT2_09	RPKLKSCKPA	1152.45	6/60.00	SER 1, GLY 0	LYS 4, ARG 1,	4/40.00
NATT2_10	PKLKSCKPAK	1124.44	6/60.00	SER 1, GLY 0	LYS 5,	4/40.00
NATT2_11	KLKSCKPAK	1124.44	6/60.00	SER 1, GLY 0	LYS 5,	4/40.00
NATT2_12	LKSCKPAKPA	1067.34	5/50.00	SER 1, GLY 0	LYS 4,	5/50.00
NATT2_13	KSKKPAKPAK	1011.23	6/60.00	SER 1, GLY 1	LYS 4,	4/40.00
NATT2_14	SKKPAKPAK	1011.23	6/60.00	SER 1, GLY 1	LYS 4,	4/40.00
NATT2_15	LRPKLKSCKPAKPAK	1747.20	9/56.25	SER 1, GLY 1	LYS 6, ARG 1,	7/43.75
NATT3_01	VYVGKNKYGLGKVHTKHE	2057.38	12/66.67	HIS 2, THR 1, ASN 1, GLY 3	LYS 4, GLU 1,	6/33.33
NATT3_02	MTRTYRNGQKRRTTSITGTYRAIQ	2704.06	17/73.91	GLN 2, SER 1, THR 6, ASN 1, GLY 2	LYS 1, ARG 4,	6/26.09
NATT3_03	YVCSGCGSSG	965.08	5/50.00	SER 3, GLY 2	-	5/50.00
NATT3_04	CSCGCGSSGFY	1013.12	5/50.00	SER 3, GLY 2	-	5/50.00
NATT3_05	HYAYGETEKT	1198.25	7/70.00	HIS 1, THR 2, GLY 1	LYS 1, GLU 2,	3/30.00
NATT3_06	KYGLGKVHTK	1130.36	7/70.00	HIS 1, THR 1, GLY 2	LYS 3,	3/30.00
NATT3_07	PPNHYPVMT	1158.36	3/30.00	HIS 1, THR 1, ASN 1, GLY 0	-	7/70.00
NATT3_08	PPNHYPVMTV	1160.37	3/30.00	HIS 1, THR 1, ASN 1, GLY 0	-	7/70.00
NATT3_09	TRTYRNGQKR	1279.42	9/90.00	GLN 1, THR 2, ASN 1, GLY 1	LYS 1, ARG 3,	1/10.00
NATT3_10	RTYRNGQKRT	1279.42	9/90.00	GLN 1, THR 2, ASN 1, GLY 1	LYS 1, ARG 3,	1/10.00
NATT4_01	LYVAKNKYGLGKL	1466.79	6/46.15	ASN 1, GLY 2	LYS 3,	7/53.85
NATT4_02	KACRDLYVAK	1166.4	4/40.00	GLY 0	LYS 2, ARG 1, ASP 1,	6/60.00
NATT4_03	KITNVRYNMK	1233.25	6/60.00	THR 1, ASN 2, GLY 0	LYS 2, ARG 1,	4/40.00
NATT4_04	IPFTGRLTRK	1188.44	6/60.00	THR 2, GLY 1	LYS 1, ARG 2,	4/40.00
NATT4_05	PFTGRLTRKY	1238.46	6/60.00	THR 2, GLY 1	LYS 1, ARG 2,	4/40.00
NATT4_06	FTGRLTRKYS	1228.42	7/70.00	SER 1, THR 2, GLY 1	LYS 1, ARG 2,	3/30.00
NATT4_07	TGRLTRKYSN	1195.34	8/80.00	SER 1, THR 2, ASN 1, GLY 1	LYS 1, ARG 2,	2/20.00
NATT4_08	GRLTRKYSNG	1151.29	8/80.00	SER 1, THR 1, ASN 1, GLY 2	LYS 1, ARG 2,	2/20.00
NATT4_09	RLTRKYSNGK	1222.41	8/80.00	SER 1, THR 1, ASN 1, GLY 1	LYS 2, ARG 2,	2/20.00
NATT4_10	KNKYGLGKLHQ	1372.59	9/75.00	GLN 1, HIS 1, SER 1, ASN 1, GLY 2	LYS 3,	3/25.00
NATT4_11	KANIPFTGRLTRK	1501.8	8/61.54	THR 2, ASN 1, GLY 1	LYS 2, ARG 2,	5/38.46
NATT4_12	GRLTRKYSNGKVT	1479.7	10/76.92	SER 1, THR 2, ASN 1, GLY 2	LYS 2, ARG 2,	3/23.08
NATT4_13	KVTSSSVKGIYKK	1424.7	9/69.23	SER 3, THR 1, GLY 1	LYS 4,	4/30.77
NATT4_14	VTSSSVKGIYKKV	1395.66	8/61.54	SER 3, THR 1, GLY 1	LYS 3,	5/38.46
NATT4_15	VKGIYKKVQVGEI	1460.78	7/53.85	GLN 1, GLY 2	LYS 3, GLU 1,	6/46.15
NATTP_01	LGQALIPRCRKM	1482.87	5/38.46	GLN 1, GLY 1	LYS 1, ARG 2,	8/61.54
NATTP_02	RCRKMPCGVKM	1205.56	5/50.00	GLY 1	LYS 2, ARG 2,	5/50.00
NATTP_03	QALIPRCRKMPCGV	1468.84	5/38.46	GLN 1, GLY 1	LYS 1, ARG 2,	8/61.54
NATTP_04	ALIPRCRKMPCGVK	1468.89	5/38.46	GLY 1	LYS 2, ARG 2,	8/61.54
NATTP_05	LIPRCRKMPCGVKM	1259.0	5/38.46	GLY 1	LYS 2, ARG 2,	8/61.54

The hydrophobic properties for all peptides were calculated, and a plot representing hydrophobicity vs. hydrophobic moment vs. GRAVY of peptides allowed us to visualize the differences in terms of hydrophobicity between each peptide (Figure 2). The hydrophobic plot can indicate that diminution of hydrophobicity and amphipathicity of the natterin peptides decreases their cellular uptake and that the substantial increase in these parameters can lead to an increase in their cytotoxicity. This suggests that carefully controlling these parameters can enhance peptide internalization and that above this threshold value it is expected that unwanted toxicity starts to appear. The nature of hydrophobic residues, positioning, and aromaticity are harmful mainly to CPPs' fate in terms of the reversibility of the membrane interaction and final membrane crossing. Studies with Trp-rich peptides revealed that less hydrophobic residues and more interfacial ones can contribute to the peptides establishing more transitory interactions with the membrane in part due to a less deep membrane insertion. This type of flexible membrane interaction is important to prevent the peptide from being locked in the membrane interior and to trigger translocation into membranes [71].

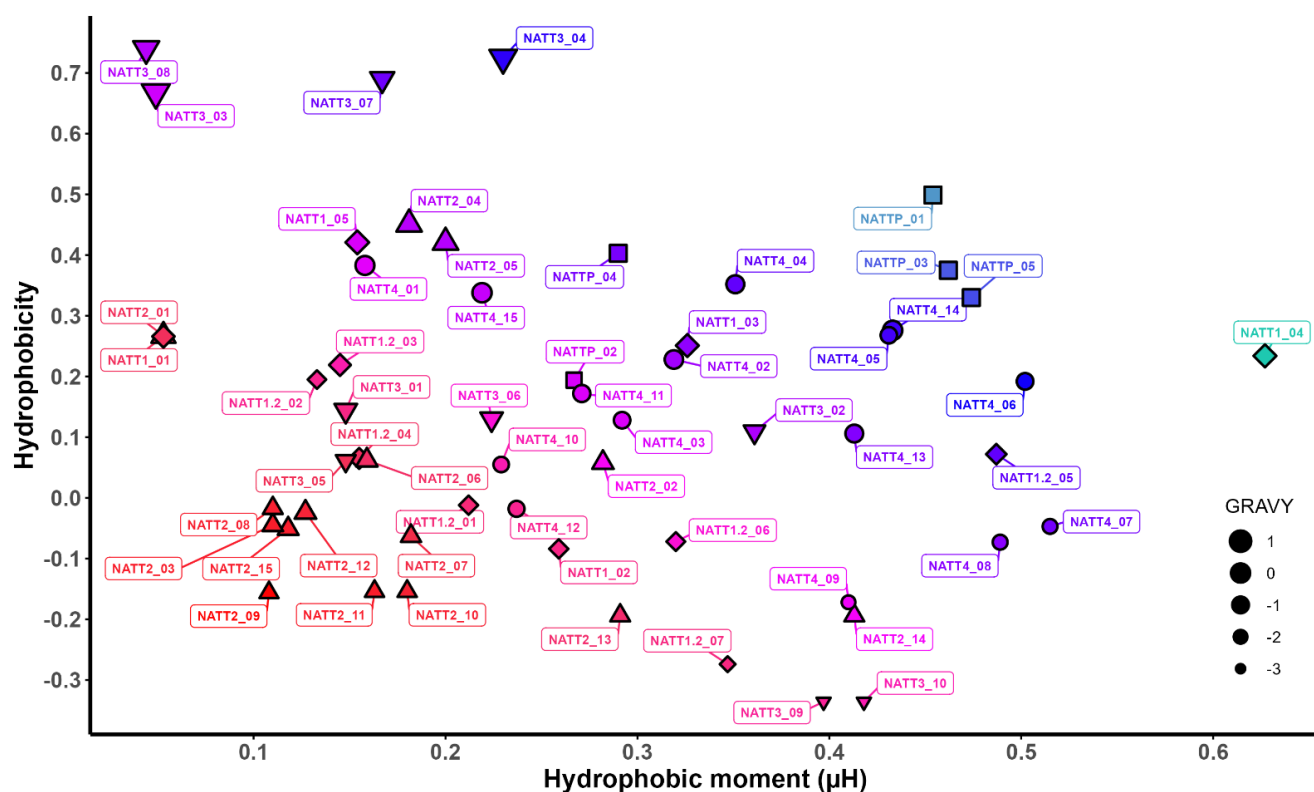


Figure 2. Hydrophobic plot and hydrophobicity (GRAVY) of predicted natterin-derived antimicrobial and cell-penetrating peptides. The X-axis represents the helical hydrophobic moment (μH), and the Y-axis represents the corresponding hydrophobicity (H). The peptides in blue and red are the most and least internalized, respectively. NATT1_04 (green) had high cell internalization and cytotoxicity.

The Boman index estimates protein-binding potential and is calculated on the basis of the cyclohexane-to-water partition coefficient of the respective amino acid side chains divided by the total number of amino acid residues within the peptide [72]. A high index (>2.48) indicates high binding potential (e.g., hormones), whereas a low index (≤ 1) indicates fewer side effects (e.g., lower toxicity to mammalian cells) [72]. Seven (12%) sequences had a Boman index below 1 (Figure 3). The sequences YVCSCGCSSG (NATT3_03) and LYVAKNKYGLGKL (NATT4_01) presented the best index (0.05 and 0.08, respectively) and will be further chemically synthesized and assayed in vitro and in vivo. These Boman values were expected as AMPs typically do not bind to other proteins but penetrate and

disrupt the plasma membrane. Given the amphiphilic nature of CPPs, ACPs, and antiviral peptides (AVPs), strong interaction with and deep penetration into the anionic lipid bilayers are expected for BAPs, making the plasma membranes prone to disruption, endocytosis, and/or direct translocation [73,74]. The Boman index of our peptides ranged from 4.0 to 6.7, which is higher than the range reported previously (~ 1.0 – 3.5) [73].

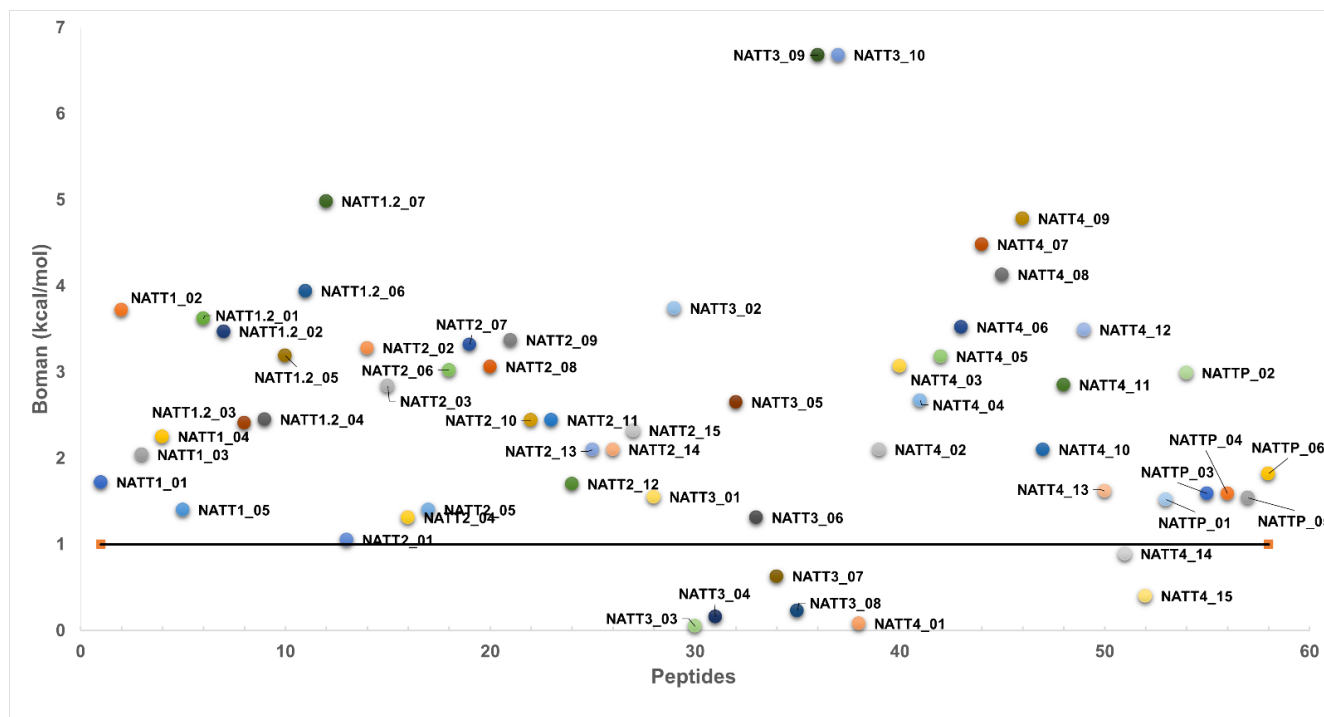


Figure 3. Boman indexes of predicted antimicrobial and cell-penetrating peptides.

The cellular localization of each peptide was evaluated using the TMHMM server to estimate the probability of peptide translocation across lipid membranes. The results showed that 90% of the predicted peptides were located in the cell membrane. The membrane-binding potential and cellular localization of CPPs are shown in Figure 4 and Table S2.

2.3. Prediction of Biological Activities

2.3.1. Immunogenicity, Allergenicity, and Toxicity

Immunogenicity assessment is a crucial step in the drug development process. The complexity of the immune system demands the use of multiple approaches to predict the immunogenicity of biopharmaceuticals. Experimental studies are straightforward, such as *in vitro*, *in vivo*, and *ex vivo*, but are sometimes expensive and time-consuming, and their results need to be confirmed [75]. Immunogenicity was analyzed using the Immune Epitope Database (IEDB), a database of epitopes and immune receptors [76] (Table 3). Higher scores indicated a higher probability of triggering an immune response. The immunogenicity of all predicted peptides was lower than 0.7, demonstrating that they did not cause immune responses [77,78].

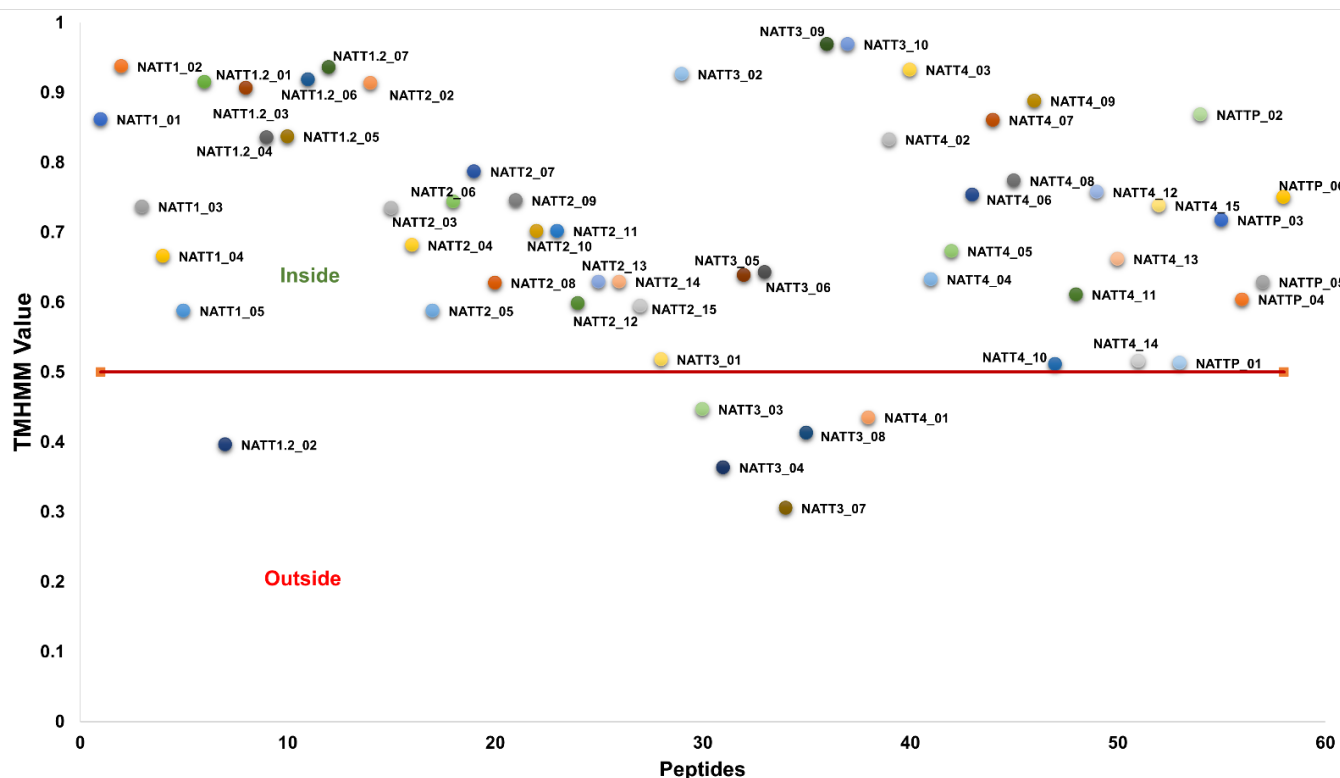


Figure 4. Cellular localization of predicted antimicrobial and cell-penetrating peptides using the TMHMM web server.

Given the risk of inducing an immediate type I (IgE-mediated) allergic response, the allergenic potential of druggable proteins and peptides should be determined before they are marketed. The allergenic potential was evaluated using the AllerTOP web server by applying auto-cross covariance transformation to build a dataset of known allergens and developing alignment-independent models for allergen recognition based on the physicochemical properties of proteins [79]. The tool uses five machine learning methods for protein classification, including partial least squares discriminant analysis, logistic regression, decision tree, naïve Bayes, and k-nearest neighbors. In addition, AllerTOP attempts to identify the most likely route of exposure. AllerTOP outperforms other allergen prediction models, with a sensitivity of 94% [79]. Of the 57 predicted sequences, 35 were classified as non-allergenic (Table 3).

Toxicity was assessed using ToxinPred software [80,81], which uses the following datasets to train and test SVM models: (1) a main dataset (1805 toxin sequences from experimentally validated peptides/proteins (positive examples) and 3593 non-toxin sequences from SwissProt (negative examples)), (2) a main independent dataset (303 toxin sequences and 300 SwissProt non-toxin sequences), (3) an alternative dataset (1805 toxin sequences (positive examples) and 12,541 non-toxin sequences from TrEMBL (negative examples)), (4) and an alternative independent dataset (303 toxin sequences from SwissProt and 1000 non-toxin sequences from TrEMBL). All identified peptide sequences were classified as non-toxic (data not shown).

Table 3. Biological activities of predicted antimicrobial and cell-penetrating peptides.

Peptides		Immunogenicity	Allergenicity	Hemolysis (%)	T1/2 <i>Escherichia coli</i>	T1/2 in Mammalian (in hours)	Antiviral		Anticancer	
Name	Sequence						Prediction	Probability	Prediction	Probability
NATT1_01	TCKTNRIYVGKGAY	6082	Non-allergen	0.48	>10 h	7.2	Non-AVP	0.344	ACP	0.982
NATT1_02	MRKSTVNNKQCKEVTK	−6046	Allergen	0.49	>10 h	30	Non-AVP	0	Non-ACP	0.414
NATT1_03	VNKDVIEQTM	14,129	Non-allergen	0.49	>10 h	100	Non-AVP	0.31	ACP	0.692
NATT1_04	DVIEQTMKDV	−26,539	Allergen	0.49	>10 h	1.1	Non-AVP	0.282	ACP	0.695
NATT1_05	TESQSYMVTV	−44,274	Allergen	0.49	>10 h	7.2	Non-AVP	0.112	Non-ACP	0.639
NATT1.2_01	RTYRGGKKTQTTTKGVYRTTQV	−3531	Allergen	0.49	2 min	1	Non-AVP	0.004	ACP	0.933
NATT1.2_02	STNDETNLHW	15,897	Non-allergen	0.49	>10 h	1.9	Non-AVP	0.068	Non-ACP	0.837
NATT1.2_03	CKTNRIYVGK	23,725	Non-allergen	0.48	>10 h	1.2	Non-AVP	0.008	ACP	0.983
NATT1.2_04	KTNRIYVGKG	11,744	Non-allergen	0.48	3 min	1.3	Non-AVP	0	ACP	0.947
NATT1.2_05	LIRTYRGGKK	3716	Non-allergen	0.46	2 min	5.5	AVP	0.964	ACP	0.911
NATT1.2_06	IRTYRGGKKT	−18,382	Non-allergen	0.49	>10 h	20	AVP	0.962	ACP	0.906
NATT1.2_07	RTYRGGKKTQ	−24,544	Non-allergen	0.49	2 min	1	AVP	0.668	ACP	0.686
NATT2_01	TCKTNKIYVGKGAY	−19,958	Non-allergen	0.49	>10 h	7.2	Non-AVP	0	ACP	0.994
NATT2_02	RTYRGGKKTQTTTKGVYRTTIQV	−27,354	Allergen	0.49	2 min	1	Non-AVP	0.068	ACP	0.944
NATT2_03	TLRPKLKSKKPAK	−98,576	Non-allergen	0.48	>10 h	7.2	Non-AVP	0.008	ACP	0.67
NATT2_04	TETQSYMVTV	−37,644	Allergen	0.49	>10 h	7.2	Non-AVP	0.068	Non-ACP	0.5
NATT2_05	ETQSYMVTVS	−28,293	Allergen	0.49	>10 h	1	Non-AVP	0.112	Non-ACP	0.653
NATT2_06	TTLRPKLKSK	−46,142	Non-allergen	0.49	>10 h	7.2	AVP	0.524	ACP	0.836
NATT2_07	TLRPKLKSKK	−68,378	Non-allergen	0.48	>10 h	7.2	AVP	0.616	ACP	0.703
NATT2_08	LRPKLKSKKP	−90,513	Non-allergen	0.49	2 min	5.5	AVP	0.696	Non-ACP	0.345
NATT2_09	RPKLKSKKPA	−84,374	Non-allergen	0.49	2 min	1	Non-AVP	0	Non-ACP	0.445
NATT2_10	PKLKSKKPAK	−7812	Non-allergen	0.49	ND	>20	Non-AVP	0	ACP	0.895
NATT2_11	KLKSKKPAKP	−75,989	Non-allergen	0.49	3 min	1.3	Non-AVP	0	ACP	0.894
NATT2_12	LKSKKPAKPA	−64,315	Allergen	0.49	2 min	5.5	Non-AVP	0.318	ACP	0.757
NATT2_13	KSKKPAKPAG	−4492	Allergen	0.49	3 min	1.3	AVP	0.654	ACP	0.919
NATT2_14	SKKPAKPAGK	−21,068	Allergen	0.49	10 h	1.9	AVP	0.654	ACP	0.992
NATT2_15	LRPKLKSKKPAKPAGK	−0.91	Non-allergen	0.48	2 min	5.5	Non-AVP	0	ACP	0.848
NATT3_01	VYVGKNKYGLGKVHTKHE	−59,206	Allergen	0.48	>10 h	100	Non-AVP	0.332	ACP	0.957
NATT3_02	MTRTYRNGQKRRTSITGTYRAIQ	16,556	Non-allergen	0.49	>10 h	30	Non-AVP	0.44	Non-ACP	0.444
NATT3_03	YVCSCGCSG	−4905	Allergen	0.49	2 min	2.8	Non-AVP	0.398	ACP	0.996
NATT3_04	CSCGCSGFY	−25,573	Non-allergen	0.49	>10 h	1.2	Non-AVP	0.104	ACP	0.998
NATT3_05	HYAYGETEKT	13,452	Allergen	0.49	>10 h	3.5	Non-AVP	0.006	ACP	0.918
NATT3_06	KYGLGKVHTK	−8.832	Allergen	0.48	3 min	1.3	Non-AVP	0.218	ACP	0.99
NATT3_07	PPNHYCPVTM	2.143	Non-allergen	0.49	ND	>20	Non-AVP	0	Non-ACP	0.873
NATT3_08	PNHYCPVTMV	875.0	Allergen	0.49	ND	>20	Non-AVP	0.126	Non-ACP	0.757

Table 3. Cont.

Peptides		Immunogenicity	Allergenicity	Hemolysis (%)	T1/2 <i>Escherichia coli</i>	T1/2 in Mammalian (in hours)	Antiviral		Anticancer	
Name	Sequence						Prediction	Probability	Prediction	Probability
NATT3_09	TRTYRNGQKR	−13.888	Non-allergen	0.48	>10 h	7.2	Non-AVP	0.46	Non-ACP	0.672
NATT3_10	RTYRNGQKRT	−18.322	Non-allergen	0.48	2 min	1	Non-AVP	0.46	Non-ACP	0.666
NATT4_01	LYVAKNKYGLGKL	−0.45197	Allergen	0.47	2 min	5.5	AVP	0.998	ACP	0.838
NATT4_02	KACRDLYVAK	996.0	Non-allergen	0.49	3 min	1.3	AVP	0.678	ACP	0.725
NATT4_03	KITNVRYNMK	−1.485	Non-allergen	0.49	3 min	1.3	Non-AVP	0.154	ACP	0.674
NATT4_04	IPFTGRLTRK	21.302	Allergen	0.49	>10 h	20	Non-AVP	0.044	ACP	0.787
NATT4_05	PFTGRLTRKY	0.4052	Non-allergen	0.50	ND	>20	Non-AVP	0.004	ACP	0.787
NATT4_06	FTGRLTRKYS	−4.536	Non-allergen	0.49	2 min	1.1	AVP	0.542	ACP	0.9
NATT4_07	TGRLTRKYSN	−0.20894	Non-allergen	0.49	>10 h	7.2	Non-AVP	0.028	Non-ACP	0.375
NATT4_08	GRLTRKYSNG	−0.27102	Allergen	0.49	>10 h	30	AVP	0.876	ACP	0.702
NATT4_09	RLTRKYSNGK	−0.29031	Non-allergen	0.49	2 min	1	Non-AVP	0.412	ACP	0.747
NATT4_10	KNKYGLGKLHQS	−0.27934	Non-allergen	0.49	3 min	1.3	AVP	0.998	ACP	0.703
NATT4_11	KANIPFTGRLTRK	0.40878	Non-allergen	0.48	3 min	1.3	AVP	0.696	Non-ACP	0.493
NATT4_12	GRLTRKYSNGKVT	−0.42112	Non-allergen	0.49	>10 h	30	Non-AVP	0.126	Non-ACP	0.387
NATT4_13	KVTSSSVKGIYKK	−0.61671	Allergen	0.51	3 min	1.3	Non-AVP	0	ACP	0.999
NATT4_14	VTSSSVKGIYKKV	−0.69995	Allergen	0.49	>10 h	100	Non-AVP	0	ACP	0.999
NATT4_15	VKGIYKKVQVGEI	−0.22532	Allergen	0.49	>10 h	100	AVP	1	ACP	0.997
NATTP_01	LGQALIPRCRKMP	−0.012821	Allergen	0.48	2 min	5.5	Non-AVP	0	Non-ACP	0.005
NATTP_02	RCRKMPGVKM	−0.44126	Non-allergen	0.49	2 min	1	Non-AVP	0.07	Non-ACP	0.434
NATTP_03	QALIPRCRKMPGV	−0.19704	Non-allergen	0.46	10 h	0.8	Non-AVP	0.268	Non-ACP	0.015
NATTP_04	ALIPRCRKMPGVK	−0.25838	Non-allergen	0.46	>10 h	4.4	Non-AVP	0.282	Non-ACP	0.098
NATTP_05	LIPRCRKMPGVKM	−0.40468	Non-allergen	0.47	2 min	5.5	AVP	0.506	Non-ACP	0.241
Inference/Reference Range		-	SVM method	>0.5: likely hemolytic <0.5: unlikely hemolytic	ND: not determined		<0.5: low probability >0.5: high probability		<0.5: low probability >0.5: high probability	

2.3.2. Antiviral and Anticancer Potential

The control of viral diseases is challenging because of increased resistance to antiviral drugs and the emergence of new viral pathogens. AVPs, a subset of AMPs, are a potential source of therapeutics useful for preventing and treating viral infections [82]. The ability of AVPs to target various stages of the viral lifecycle, ranging from their attachment to host cells to their ability to impair viral replication within the cells, has been the subject of multiple studies [83–85]. Sixteen sequences were predicted to be AVPs, of which four had a score above 90%. NATT1.2_05 and NATT1.2_06 presented the highest scores (0.964 and 0.962, respectively). AVPpred predicts AVPs based on experimentally validated positive and negative datasets.

Cell membrane properties differ between cancer cells and healthy cells [86]. For instance, the membrane fluidity of cancer cells is higher than that of healthy cells [87]. In addition, the membrane of cancer cells has a higher negative charge, larger surface area due to the higher number of microvilli, and higher fluidity than that of healthy cells. ACPs, a subset of AMPs, are toxic to cancer cells [86]. ACPs have 5–30 cationic amino acid residues that adopt an α -helical or β -sheet structure but can assume a linear structure [88,89]. In the present study, 37 peptides were predicted to be ACPs. The physicochemical properties of ACPs determine electrostatic interactions with the anionic cell membrane of cancer cells and thus allow the selective killing of these cells [90]. ACPs have several advantages over small molecule cancer drugs. For instance, the shorter half-life decreases the probability of resistance. Moreover, ACPs have low toxicity, high specificity, high solubility, and good tumor penetration ability, demonstrating their great potential in cancer therapy [88–91]. The half-life of the predicted ACPs in mammalian cells varied from 1 to 100 h (Table 3). Compared to biologics, peptides have a much shorter circulatory half-life (days vs. weeks), resulting in the need for sub-optimal frequent drug administration [92].

2.3.3. Prediction of ADMET Properties

The analysis of biochemical processes from drug administration to elimination plays a crucial role in lead optimization. An ideal peptide drug should be quickly absorbed into the systemic circulation and eliminated without affecting pharmacological activity. Further, ideal candidates should be non-toxic. The analysis of ADMET parameters is essential in drug discovery. ADMET properties were predicted using the web server ADMETlab version 2.0 [93] (Table 4). The parameters analyzed were blood–brain barrier (BBB) penetration, Caco-2 permeability, volume of distribution (VD), plasma protein binding (PPB), human intestinal absorption (HIA), clearance (CL), half-life ($T_{1/2}$), skin sensitization, AMES toxicity, carcinogenicity, and synthetic accessibility (SA) score (Table 4). All compounds had positive HIA, indicating the high ability to cross the intestinal barrier. Higher BBB penetration is associated with higher lipophilicity profiles and higher uptake. The calculated value for the BBB was shown to have a high likelihood of being negative. PPB is an important parameter in drug safety assessments since compounds with high PPB (>90%) have a narrow therapeutic index, whereas compounds with low PPB are considerably safer. All analyzed peptides had low PPB, indicating a good therapeutic index. Caco-2 cells, derived from human colon adenocarcinoma cells, have permeability functions similar to those of intestinal enterocytes and are used to predict intestinal drug absorption *in vivo*. All analyzed compounds had the best scores (greater than -6.47) in Caco-2 cell permeability assays. Regarding carcinogenicity, none of the analyzed peptides showed potential to cause cancer. The results of the AMES test showed that none of the peptides were genotoxic. The analysis of other toxicity parameters, such as hERG inhibition, hepatotoxicity, and skin sensitization, revealed that all peptides were safe. The SA score estimates the ease of synthesis (Table S3). Approximately 38.5% of the peptides had an SA score of up to 6.0, indicating the feasibility of synthesis. All compounds had good ADMET properties.

Table 4. Cont.

Name	Peptides Sequence	Absorption		VD (L/Kg)	Distribution		Metabolism				Excretion		DILI Liver Injury	Toxicity		Skin Sensitization	
		HIA (%)	Caco-2 Permeability (cm/s)		BBB Penetration (%)	PPB(%)	CYP1A2-(I)	CYP1A2-(S)	CYP3A4-(I)	CYP3A4-(S)	CL (mL/min/Kg)	Half-Life		hERG Blockers	AMES		Carcinogenicity
NATTP_04	ALIPRCRKMPGVK	0.995	-6.759	0.488	0.011	20.18	0	0	0.002	0	1.078	0.7	0.015	0.004	0.011	0.033	0.159
NATTP_05	LIPRCRKMPGVKM	0.995	-6.721	0.467	0.009	22.53	0	0	0.004	0	1.119	0.742	0.014	0.003	0.013	0.038	0.153
Inference/Reference Range		HIA > 0.3: HIA positive HIA < 0.3: HIA negative	Optimal: higher than -5.15	Optimal: 0.04-20 L/Kg	≥0.1: BBB positive and <0.1: BBB negative	PPB < 90%: optimal PPB > 90%: low therapeutic index	>0.5: inhibitor <0.5: non inhibitor	>0.5: substrate <0.5: non substrate	>0.5: inhibitor <0.5: non-inhibitor	>0.5: substrate <0.5: non-substrate	High: >15 mL/min/kg Moderate: 5-5 mL/min/kg Low: <5 mL/min/kg	Long half-life: >3 h Short half-life: <3 h	>0.5: blocker <0.5: non-blocker	>0.5: hepatotoxic <0.5: non-hepatotoxic	>0.5: positive <0.5: negative	>0.5: carcinogen <0.5: non-carcinogen	>0.5: sensitizer <0.5: non-sensitizer

2.4. Medicinal Chemistry Studies

Small molecules defined as “drug-like” need to satisfy Lipinski’s rule of five (Ro5): MW <500 Da, ≤5 H-bond donors, ≤10 H-bond acceptors, and 1–octanol/water partition coefficient (LogP) <5. Molecules that satisfy these criteria are likely to be orally bioavailable. Several studies have demonstrated that the physicochemical and structural properties of peptides are outside the traditional chemical space of approved drugs [94–96] based on Ro5 criteria [97]. Medicinal chemistry parameters such as MW, topological polar surface area (tPSA), LogP, fraction of sp³-hybridized carbon atoms (Fsp³), number of rotatable bonds (NRB), number of hydrogen bond acceptors (HBAs), number of hydrogen bond donors (HBDs), and number of aromatic rings (NARs) were evaluated (Table 5).

Table 5. Chemical spaces of predicted peptides and marketed drugs (modified from Oliveira et al., 2021).

Molecular Properties	Oral Drugs				Peptides	
	Lipinski, 2001 and Veber, 2002	Doak et al., 2014	Santos et al., 2016 *	Diaz-Eufracio et al., 2018 **	De Oliveira et al., 2021 #	Our Study
MW	≤500	≤1,000	≤700	27.03 ≤ MW ≤ 5036.65	331.48 ≤ MW ≤ 3750.51	965.08 ≤ MW ≤ 2704.06
LogP	≤5	−2 ≤ LogP ≤ 10	≤7.5	−17.87 ≤ LogP ≤ 39.89	−42.12 ≤ LogP ≤ 2.97	−7.387 ≤ LogP ≤ 0.562
tPSA	≤140	≤250	≤200	≤2064.83	101.29 ≤ tPSA ≤ 1782.83	405.88 ≤ tPSA ≤ 1288.48
Fsp ³	—	—	≤0.55	—	0.37 ≤ Fsp ³ ≤ 0.84	0.45 ≤ Fsp ³ ≤ 0.80
NRB	≤10	≤20	≤20	≤209	9 ≤ NRB ≤ 137	37 ≤ NRB ≤ 117
HBA	≤10	≤15	≤10	≤71	5 ≤ HBA ≤ 55	25 ≤ HBA ≤ 75
NAR	—	—	—	—	≤10	≤5

Notes: * Investigated orally active peptides; ** investigated linear and cyclic pentapeptides; # investigated cell-penetrating peptides.

Santos et al. analyzed peptides approved by the FDA between 2012 and 2016 to allow comparison to the Ro5 [96]. The peptides with the highest oral availability had an MW of 1200 Da and a LogP of 5–8. Furthermore, these peptides had five times more H-bond donors and acceptors than what was considered acceptable by Ro5 for small molecules [96]. High MW, tPSA, and NRB limit passive transport across cell membranes because of increased molecular size and complexation with water molecules [98,99]. HBAs and HBDs are relevant factors for cell permeability by Ro5 [100]. Our results agreed with the number of HBAs and HBDs for linear and cyclic pentapeptides and two CPP libraries [44,95]. However, the number of HBAs and HBDs in predicted peptides differed from those of clinically approved drugs [100]. The NRB and Fsp³ are used to assess molecular flexibility and complexity. The NRBs of the predicted peptides (37 to 117) exceed the maximum value for oral drugs and peptides [95,96]. The Fsp³ correlates with solubility in the aqueous phase and melting point [101]. The Fsp³ of the predicted peptides was 0.45–0.80, similar to values of orally available peptides (90th percentile = 0.79). Lipophilicity was investigated using LogP and NAR. LogP values are positively correlated with lipophilicity and thus membrane penetration. The LogP of the evaluated peptides varied from −7.387 to 0.562, consistent with values for approved peptide drugs and small molecule drugs [96,100]. The addition of an aromatic ring can significantly increase LogP [102]. Our study found that the NAR varied from 2 to 5.

2.5. Prediction of Peptide Structures

After analyzing the physicochemical properties of the peptides, hydrophobicity, hydrophobic moment, GRAVY, Boman index, and ADMET parameters, fifteen BAP sequences of AMPs and CPPs with characteristics considered promising were selected for further studies. Among the 3D structures obtained, it was possible to observe the presence of a random coil, alpha helix, and a peptide sequence (NATT1.2_07) with a beta sheet structure.

The 3D structures were predicted using the PEP-FOLD3 web server. PEP-FOLD models for the peptides NATT1_04, NATT1.2_05, NATT1.2_06, NATT1.2_07, NATT2_06, NATT2_07, NATT2_13, NATT2_14, NATT3_03, NATT3_04, NATT4_01, NATT4_02, NATT4_06, NATT4_15, and NATTP_05 were recognized as the best with the lowest opti-

mized potential for efficient structure prediction (sOPEP) energy (-15.1734 to -1.97158). The models with sOPEP energy of -15.1734 and -14.3622 were considered the best and are presented in Figure 5. Ramachandran plot analysis indicated that these two models had 77.8% and 87.5% of the residues in the most favorable region and 0% and 22.2% of the residues in the favorable region, respectively. In addition, the helical wheel projection of these short peptides was obtained using the Heliquest web server (Figure 5). A hydrophobic face on a helical wheel is characterized by at least five adjacent hydrophobic residues (Leu, Ile, Ala, Val, Pro, Met, Phe, Trp, or Tyr) [103].

The pH-dependent conformational equilibrium of the peptides was predicted using DispHred [104]. Khandogin [105] showed that increasing pH increased the length of the helical segments of C peptide from ribonuclease, where the difference in the relative populations of unfolded states gave rise to the pH-dependent total helix content. Our results showed that at pH 1.5 and 7.0, 75% of the peptides are in the unfolded state, with data indicating the presence of partial helices. The results provided information on the pH-dependent distribution of folded and unfolded states of the peptides. However, further in vitro studies are necessary to corroborate these data.

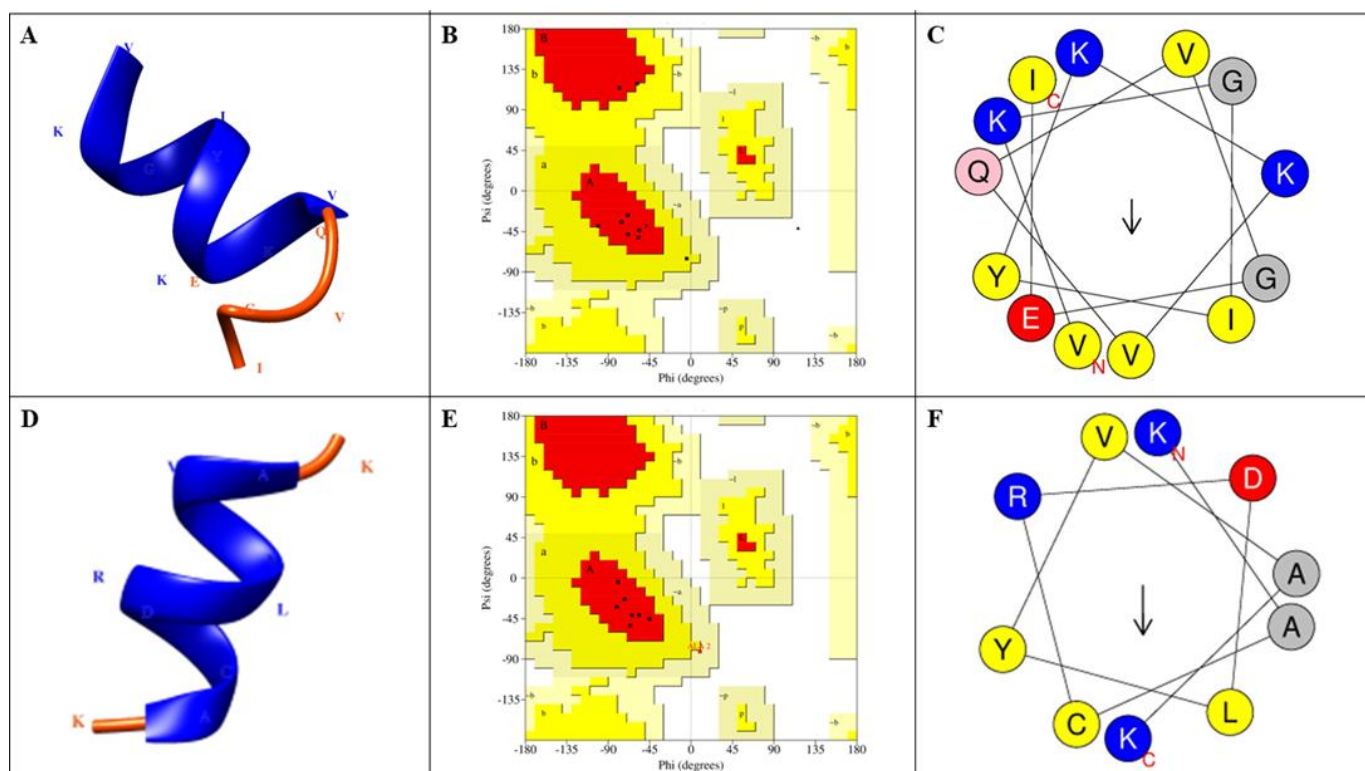


Figure 5. Selected PEP FOLD predicted 3D structure homology models, Ramachandran validation plots, and helical wheel projections. (A) NATT4_15 motif, (B) Ramachandran plot for the NATT4_15 motif, (C) NATT4_15 helical wheel projection, (D) NATT4_02 motif, (E) Ramachandran plot for the NATT4_02 motif, and (F) NATT4_02 helical wheel projection. NATT4_15 had nine amino acid sequences in the allowed region, whereas NATT4_02 had eight amino acids in the favorable region. These two peptides had no amino acid sequence in the unfavorable region. The graphical representations were rendered using UCSF Chimera [106]. Arrows indicate the direction of the hydrophobic moment (μH).

3. Materials and Methods

3.1. Study Design

The current study used several in silico approaches to find and design novel and potent AMPs and CPPs as a drug delivery system. The flowchart of peptide prediction and analysis is illustrated in Figure 6.

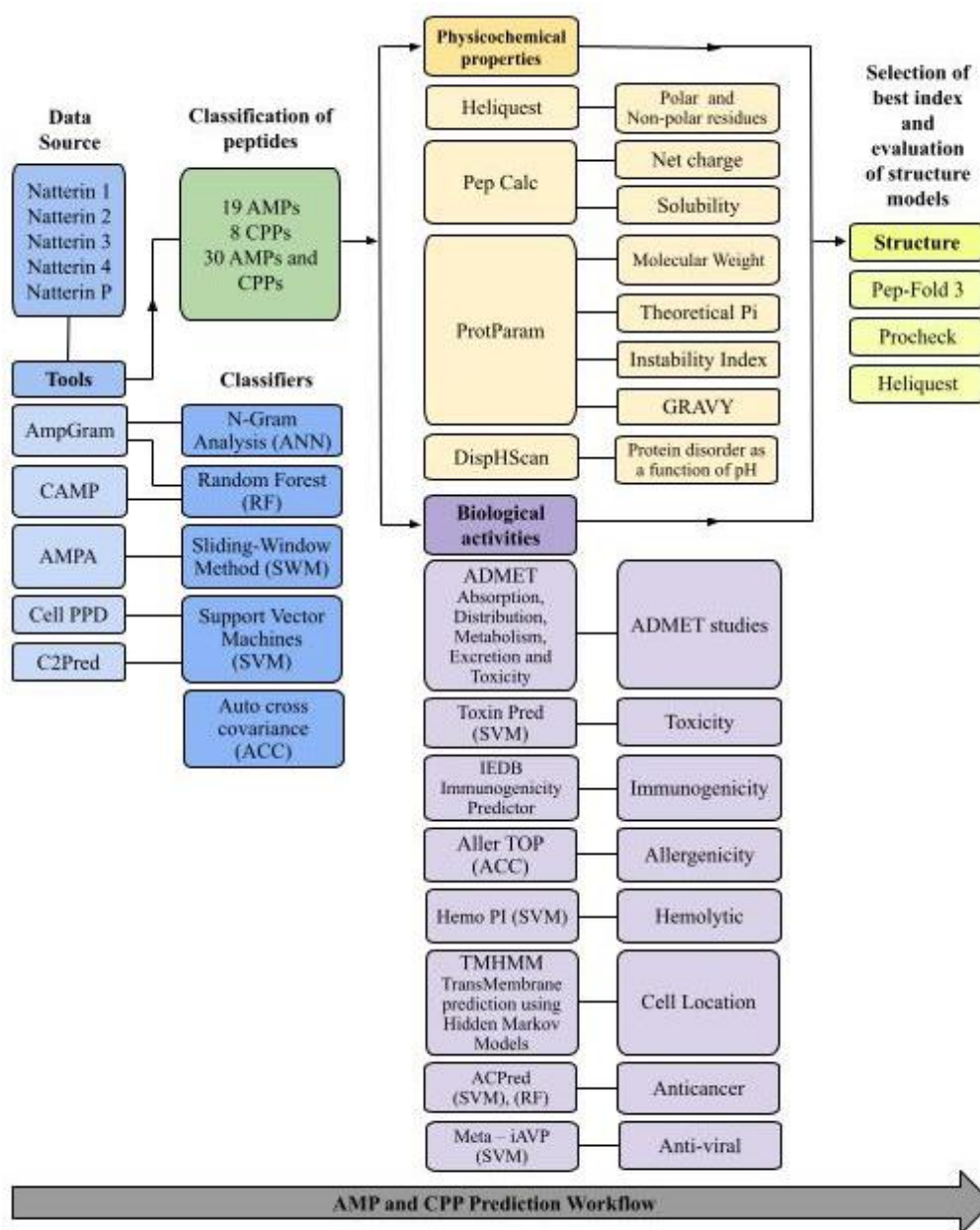


Figure 6. Flowchart of peptide prediction and analysis.

3.2. Prediction of BAPs

BAPs from natterin 1 (UniProt Q66S25), natterin 2 (UniProt Q66S21), natterin 3 (UniProt Q66S17), natterin 4 (UniProt Q66S13), and natterin P (UniProt Q66S08) were predicted using bioinformatics tools. AMPs were predicted using AMPA (<http://tcoffeecrg.cat/apps/ampa/do>, accessed on 5 January 2022), CAMP algorithm (<http://www.camp.bicnirrh.res.in>, accessed on 5 January 2022) (based on primary amino acid sequences), and AmpGram (<http://biongram.biotech.uni.wroc.pl/AmpGram>, accessed on 5 August 2022) [107], which employs n-grams (amino acid motifs) and random forests for prediction. All sequences were in FASTA format, and antimicrobial domains were detected for the design of AMPs [108,109]. CPPs were predicted using the SVM-based web server CellPPD (<https://webs.iitd.edu.in/raghava/cellppd/index.html>) and C2Pred (<http://lin-group.cn/server/C2Pred>, accessed both on 10 January 2022) [54–56]. The sequences were submitted to a protein scanning tool with the default threshold of the SVM-based prediction method [110].

3.3. Physicochemical Properties

The physicochemical parameters of peptide sequences were evaluated using different tools. MW, net charge, theoretical isoelectric point (pI), instability index, and GRAVY were estimated using ProtParam, available on the bioinformatics resource portal ExPASy of the Swiss Institute of Bioinformatics website (<http://web.expasy.org/protparam>, accessed on 19 October 2021). Peptide solubility and net charge at pH 7.0 were evaluated using PepCalc (<https://pepcalc.com/>, accessed on 24 January 2022). Hydrophobic moment (μ H), hydrophobicity (H), and amino acid charge were estimated using Heliquest (<https://heliquest.ipmc.cnrs.fr/cgi-bin/ComputParams.py>, accessed on 24 January 2022) and the Antimicrobial Peptide Database (APD3) (<https://aps.unmc.edu/prediction/predict>, accessed on 25 January 2022).

3.4. Evaluation of the Membrane-Binding Ability of BAPs

The Boman index and protein-binding potential were evaluated using APD3 (http://aps.unmc.edu/AP/prediction/prediction_main.php, accessed on 5 February 2022). The Boman index is the sum of solubility values for all amino acids in a peptide sequence and indicates the ability to bind to the cell membrane or other proteins [72]. The cellular localization of BAPs was assessed using the TMHMM web server (<http://www.cbs.dtu.dk/services/TMHMM>, accessed on 8 February 2022). TMHMM analyzes the probability of a peptide to bind to the negatively charged bacterial cell membranes.

3.5. Assessment of Immunogenicity, Toxicity, Allergenicity, and Anticancer and Antiviral Properties

Peptides can induce immune responses in vivo, resulting in allergic reactions. Neutralizing antibodies bind to proteins, reducing the therapeutic efficacy of these proteins [77,78]. Immunogenicity was evaluated using IEDB Immunogenicity Predictor (<http://tools.iedb.org/immunogenicity>, accessed on 9 February 2022) [57]. Toxicity and allergenicity were analyzed using ToxinPred (<https://webs.iiitd.edu.in/raghava/toxinpred/algo.php>, accessed on 10 February 2022) and AllerTop (<https://www.ddg-pharmfac.net/AllerTOP>, accessed on 10 February 2022) (<http://ddg-pharmfac.net/AllergenFP>, accessed on 11 February 2022) [61,62]. Antiviral and anticancer peptides were predicted using MetaIAP (<http://codes.bio/meta-iavp>, accessed on 12 February 2022) and ACPred (<http://codes.bio/acpred>, accessed on 12 February 2022), respectively [111,112].

3.6. Hemolytic Activity and Half-Life

Hemolytic activity was predicted using the SVM-based HemoPI (<https://webs.iiitd.edu.in/raghava/hemopi/design.php>, accessed on 12 February 2022). Peptide half-life in *E. coli* and mammalian cells was calculated using ProtParam (<https://web.expasy.org/protparam>, accessed on 12 February 2022).

3.7. Prediction of ADMET and Medicinal Chemistry Parameters

The Simplified Molecular Input Line Entry System (SMILES) structural format of 58 peptides was obtained using PepSMI (<https://www.novoprolabs.com/tools/convert-peptide-to-smiles-string>, accessed on 2 April 2022). PepSMI runs an algorithm that converts raw sequences into a string of texts and unambiguously describes each atom and molecular bond in a manner amenable to machine processing. ADMET parameters, including human intestinal absorption (HIA), mutagenicity, carcinogenicity, central nervous system penetration, drug-induced liver injury (DILI), cytochrome P450 enzyme inhibition, carcinogenicity, mutagenicity, clearance, half-life, and skin sensitization, were assessed using version 420 (released on July 2021) of the ADMETlab 2.0 platform (<https://admetmesh.scbdd.com/>, accessed on 4 April 2022) and a comprehensive database composed of 0.25 million entries from PubChem, Online Chemical Modeling Environment (OCHEM), DrugBank, ChEMBL, Toxicity Estimation Software Tools (developed by the U.S. Environmental Protection Agency), and peer-reviewed literature [93]. Pan-assay interference compounds

(PAINS) and undesirable reactive compounds were analyzed using the PAINS and Pfizer rules [113]. The ADMETlab 2.0 platform predicts the pharmacokinetic parameters based on basic information and experimental values of the respective entries.

3.8. Prediction of Peptide Structure

The three-dimensional (3D) structures of predicted BAPs were analyzed using PEP-FOLD3 (<https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/>, accessed on 3 April 2022), which predicts peptide structures de novo based on primary amino acid sequences. Peptides are described as a series of fragments of four amino acids, overlapping by three, and each fragment is associated with a geometric descriptor [114]. The quality of the best models was assessed. Peptide structures were validated using PROCHECK to measure the stereochemical properties of the modeled peptide motifs [115]. Furthermore, the helical wheel diagram of peptides was defined by Schiffer Edmundson wheel modeling using Heliquest (<https://heliquest.ipmc.cnrs.fr/cgi-bin/ComputParams.py>, accessed on 5 April 2022) [103]. pH-dependent folded and unfolded states were predicted using SVM-based DispHred (<https://ppmclab.pythonanywhere.com/DispHred>, accessed on 5 August 2022) [104].

4. Conclusions

Fifty-seven novel and potent AMPs and CPPs derived from natterins were predicted in silico from natterin toxins. Moreover, we predicted novel peptides that had high binding membrane indexes and localization inside cells. These peptide sequences can be further evaluated for antimicrobial, cell penetration, and anticancer activity in vitro and in vivo in advance. Generally, the predicted and engineered toxin-derived AMPs and CPPs with different properties can be applied to deliver different cargoes and drug development. Overall, the present study showed that using machine learning tools in peptide research can streamline the development of targeted peptide therapies.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ph15091141/s1>, Supplementary Figure S1. Multiple sequence alignment analysis. Natterins from *Thalassophryne nattereri* venom were aligned using ClustalW (UNIT 2.3 Multiple Sequence Alignment Using ClustalW and ClustalX, 2003). Accession numbers: SP|Q66S25| Natterin 1; SP|Q66S21| Natterin 2; SP|Q66S17| Natterin 3; SP|Q66S13| Natterin 4; SP|Q66S08| Natterin P. *Fully conserved residues; conservation of strongly similar groups (>0.5 in the Gonnet PAM 250 matrix); conservation of weakly similar groups (≤ 0.5 in the Gonnet PAM 250 matrix) [116]. Table S1. Physicochemical proprieties of predicted antimicrobial and cell-penetrating peptides. Table S2. Membrane-binding potential and cellular localization of predicted antimicrobial and cell-penetrating peptides. Table S3. Prediction of medicinal chemistry proprieties of peptides. Reference [116] are cited in the Supplementary Materials.

Author Contributions: Conceptualization, K.C.; methodology, K.C., G.L.D.C. and B.V.S.; software, K.C., G.L.D.C. and B.V.S.; validation, K.C., G.L.D.C. and B.V.S.; data analysis, K.C. and G.L.D.C.; investigation, K.C. and G.L.D.C.; resources, K.C.; data curation, K.C. and G.L.D.C.; writing of the first draft, K.C. and G.L.D.C.; manuscript review, K.C. and G.L.D.C.; funding acquisition, K.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by FAPESP (Grant Number 2021/04316-9).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article and supplementary material.

Acknowledgments: We are thankful to Bruno Mito and Carlos Cruz for technical assistance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pennington, M.W.; Czerwinski, A.; Norton, R.S. Peptide Therapeutics from Venom: Current Status and Potential. *Bioorg. Med. Chem.* **2018**, *26*, 2738–2758. [[CrossRef](#)] [[PubMed](#)]
2. Usmani, S.S.; Bedi, G.; Samuel, J.S.; Singh, S.; Kalra, S.; Kumar, P.; Ahuja, A.A.; Sharma, M.; Gautam, A.; Raghava, G.P.S. THPdb: Database of FDA-Approved Peptide and Protein Therapeutics. *PLoS ONE* **2017**, *12*, e0181748. [[CrossRef](#)] [[PubMed](#)]
3. Herzig, V.; Cristofori-Armstrong, B.; Israel, M.R.; Nixon, S.A.; Vetter, I.; King, G.F. Animal Toxins—Nature’s Evolutionary-Refined Toolkit for Basic Research and Drug Discovery. *Biochem. Pharmacol.* **2020**, *181*, 114096. [[CrossRef](#)] [[PubMed](#)]
4. Slagboom, J.; Kaal, C.; Arrahman, A.; Vonk, F.J.; Somsen, G.W.; Calvete, J.J.; Wüster, W.; Kool, J. Analytical Strategies in Venomics. *Microchem. J.* **2022**, *175*, 107187. [[CrossRef](#)]
5. Padhi, A.; Sengupta, M.; Sengupta, S.; Roehm, K.H.; Sonawane, A. Antimicrobial Peptides and Proteins in Mycobacterial Therapy: Current Status and Future Prospects. *Tuberculosis* **2014**, *94*, 363–373. [[CrossRef](#)]
6. Buchwald, H.; Dorman, R.B.; Rasmus, N.F.; Michalek, V.N.; Landvik, N.M.; Ikramuddin, S. Effects on GLP-1, PYY, and Leptin by Direct Stimulation of Terminal Ileum and Cecum in Humans: Implications for Ileal Transposition. *Surg. Obes. Relat. Dis.* **2014**, *10*, 780–786. [[CrossRef](#)]
7. Giordano, C.; Marchiò, M.; Timofeeva, E.; Biagini, G. Neuroactive Peptides as Putative Mediators of Antiepileptic Ketogenic Diets. *Front. Neurol.* **2014**, *5*, 63. [[CrossRef](#)]
8. Fosgerau, K.; Hoffmann, T. Peptide Therapeutics: Current Status and Future Directions. *Drug Discov. Today* **2015**, *20*, 122–128. [[CrossRef](#)] [[PubMed](#)]
9. Fjell, C.D.; Hiss, J.A.; Hancock, R.E.W.; Schneider, G. Designing Antimicrobial Peptides: Form Follows Function. *Nat. Rev. Drug Discov.* **2012**, *11*, 37–51. [[CrossRef](#)]
10. Diocot, S.; Baron, A.; Salinas, M.; Douguet, D.; Scarzello, S.; Dabert-Gay, A.S.; Debayle, D.; Friend, V.; Alloui, A.; Lazdunski, M.; et al. Black Mamba Venom Peptides Target Acid-Sensing Ion Channels to Abolish Pain. *Nature* **2012**, *490*, 552–555. [[CrossRef](#)]
11. Hancock, R.E.W.; Sahl, H.G. Antimicrobial and Host-Defense Peptides as New Anti-Infective Therapeutic Strategies. *Nat. Biotechnol.* **2006**, *24*, 1551–1557. [[CrossRef](#)] [[PubMed](#)]
12. Holford, N.; Channon, S.; Heron, J.; Jones, I. The Impact of Postpartum Psychosis on Partners 11 Medical and Health Sciences 1117 Public Health and Health Services. *BMC Pregnancy Childbirth* **2018**, *18*, 414. [[CrossRef](#)]
13. Ziegman, R.; Alewood, P. Bioactive Components in Fish Venoms. *Toxins* **2015**, *7*, 1497–1531. [[CrossRef](#)] [[PubMed](#)]
14. Facó, P.E.; Bezerra, G.P.; Ferreira Barbosa, P.S.; Maria Costa Martins, A.; Guimarães, J.A.; Ferreira, M.L.; Serra, H.; Monteiro, A. Epidemiology of the injuries caused by *Thalassophryne nattereri* (niquim) in Ceara State (1992–2002). *Rev. Soc. Bras. Med. Trop.* **2005**, *38*, 479–482. [[CrossRef](#)]
15. Haddad, V., Jr.; Pardal, P.P.O.; Cardoso, J.L.C.; Martins, I.A. Case Report the Venomous Toadfish *Thalassophryne Nattereri* (Niquim or Miquim): Report of 43 Injuries Provoked in Fishermen of Salinópolis (Pará State) and Aracaju (Sergipe State), Brazil. *Rev. Do Inst. De Med. Trop. De São Paulo* **2003**, *45*, 221–223.
16. Lopes-Ferreira, M.; Grund, L.Z.; Lima, C. *Thalassophryne Nattereri* fish Venom: From the Envenoming to the Understanding of the Immune System. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **2014**, *20*, 1–12. [[CrossRef](#)]
17. Magalhães, G.S.; Junqueira-de-Azevedo, I.L.M.; Lopes-Ferreira, M.; Lorenzini, D.M.; Ho, P.L.; Moura-da-Silva, A.M. Transcriptome Analysis of Expressed Sequence Tags from the Venom Glands of the Fish *Thalassophryne Nattereri*. *Biochimie* **2006**, *88*, 693–699. [[CrossRef](#)]
18. Lopes-Ferreira, M.; Emim, J.A.D.S.; Oliveira, V.; Puzer, L.; Cezari, M.H.; Araújo, M.D.S.; Juliano, L.; Lapa, A.J.; Souccar, C.; Moura-Da-Silva, A.M. Kininogenase Activity of *Thalassophryne Nattereri* Fish Venom. *Biochem. Pharmacol.* **2004**, *68*, 2151–2157. [[CrossRef](#)]
19. Magalhães, G.S.; Lopes-Ferreira, M.; Junqueira-De-Azevedo, I.L.M.; Spencer, P.J.; Araújo, M.S.; Portaro, F.C.V.; Ma, L.; Valente, R.H.; Juliano, L.; Fox, J.W.; et al. Natterins, a New Class of Proteins with Kininogenase Activity Characterized from *Thalassophryne Nattereri* Fish Venom. *Biochimie* **2005**, *87*, 687–699. [[CrossRef](#)]
20. Lima, C.; Falcao, M.A.P.; Andrade-Barros, A.I.; Seni-Silva, A.C.; Grund, L.Z.; Balogh, E.; Conceição, K.; Queniaux, V.F.; Ryffel, B.; Lopes-Ferreira, M. Natterin an Aerolysin-like Fish Toxin Drives IL-1 β -Dependent Neutrophilic Inflammation Mediated by Caspase-1 and Caspase-11 Activated by the Inflammasome Sensor NLRP6. *Int. Immunopharmacol.* **2021**, *91*, 107287. [[CrossRef](#)]
21. Lewis, R.J.; Garcia, M.L. Therapeutic Potential of Venom Peptides. *Nat. Rev. Drug Discov.* **2003**, *2*, 790–802. [[CrossRef](#)] [[PubMed](#)]
22. Heitz, F.; Morris, M.C.; Divita, G. Themed Section: Vector Design and Drug Delivery Review Twenty Years of Cell-Penetrating Peptides: From Molecular Mechanisms to Therapeutics. *Br. J. Pharmacol.* **2009**, *157*, 195–206. [[CrossRef](#)] [[PubMed](#)]
23. Säälk, P.; Elmquist, A.; Hansen, M.; Padari, K.; Saar, K.; Viht, K.; Langel, Ü.; Pooga, M. Protein Cargo Delivery Properties of Cell-Penetrating Peptides. A Comparative Study. *Bioconjugate Chem.* **2004**, *15*, 1246–1253. [[CrossRef](#)]
24. Pooga, M.; Soomets, U.; Hällbrink, M.; Valkna, A.; Saar, K.; Rezaei, K.; Kahl, U.; Hao, J.X.; Xu, X.J.; Wiesenfeld-Hallin, Z.; et al. Cell penetrating PNA constructs regulate galanin receptor levels and modify pain transmission in vivo. *Nat. Biotechnol.* **1998**, *16*, 857–861. [[CrossRef](#)] [[PubMed](#)]
25. Vives, E. Present and Future of Cell-Penetrating Peptide Mediated Delivery Systems: “Is the Trojan Horse Too Wild to Go Only to Troy?”. *J. Control. Release* **2005**, *109*, 77–85. [[CrossRef](#)]
26. Reddy, K.V.R.; Yedery, R.D.; Aranha, C. Antimicrobial Peptides: Premises and Promises. *Int. J. Antimicrob. Agents* **2004**, *24*, 536–547. [[CrossRef](#)]

27. Zasloff, M. Antimicrobial Peptides of Multicellular Organisms: My Perspective. In *Advances in Experimental Medicine and Biology*; Springer: New York, NY, USA, 2019; Volume 1117, pp. 3–6.
28. Ganz, T. Defensins: Antimicrobial Peptides of Innate Immunity. *Nat. Rev. Immunol.* **2003**, *3*, 710–720. [[CrossRef](#)]
29. Ellerby, H.M.; Bredesen, D.E.; Fujimura, S.; John, V. Hunter-Killer Peptide (HKP) for Targeted Therapy. *J. Med. Chem.* **2008**, *51*, 5887–5892. [[CrossRef](#)]
30. Kolonin, M.G.; Saha, P.K.; Chan, L.; Pasqualini, R.; Arap, W. Reversal of Obesity by Targeted Ablation of Adipose Tissue. *Nat. Med.* **2004**, *10*, 625–632. [[CrossRef](#)]
31. Barnhart, K.F.; Christianson, D.R.; Hanley, P.W.; Driessen, W.H.P.; Bernacky, B.J.; Baze, W.B.; Wen, S.; Tian, M.; Ma, J.; Kolonin, M.G.; et al. A Peptidomimetic Targeting White Fat Causes Weight Loss and Improved Insulin Resistance in Obese Monkeys. *Sci. Transl. Med.* **2011**, *3*, 108ra112. [[CrossRef](#)]
32. Ellerby, H.; Arap, W.; Ellerby, L.M.; Kain, R.; Andrusiak, R.; Rio, G.D.; Krajewski, S.; Lombardo, C.R.; Rao, R.; Ruoslahti, E.; et al. Anti-cancer activity of targeted pro-apoptotic peptides. *Nat. Med.* **1999**, *5*, 1032–1038. [[CrossRef](#)] [[PubMed](#)]
33. Morris, M.C.; Depollier, J.; Mery, J.; Heitz, F.; Divita, G. A Peptide Carrier for the Delivery of Biologically Active Proteins into Mammalian Cells. *Nat. Biotechnol.* **2001**, *19*, 1173–1176. [[CrossRef](#)] [[PubMed](#)]
34. Wang, F.; Wang, Y.; Zhang, X.; Zhang, W.; Guo, S.; Jin, F. Recent Progress of Cell-Penetrating Peptides as New Carriers for Intracellular Cargo Delivery. *J. Control. Release* **2014**, *174*, 126–136. [[CrossRef](#)] [[PubMed](#)]
35. Gautam, A.; Singh, H.; Tyagi, A.; Chaudhary, K.; Kumar, R.; Kapoor, P.; Raghava, G.P.S. CPPsite: A Curated Database of Cell Penetrating Peptides. *Database* **2012**, *2012*, bas015. [[CrossRef](#)] [[PubMed](#)]
36. Farnaud, S.; Evans, R.W. Lactoferrin—A Multifunctional Protein with Antimicrobial Properties. *Mol. Immunol.* **2003**, *40*, 395–405. [[CrossRef](#)]
37. Loose, C.; Jensen, K.; Rigoutsos, I.; Stephanopoulos, G. A Linguistic Model for the Rational Design of Antimicrobial Peptides. *Nature* **2006**, *443*, 867–869. [[CrossRef](#)]
38. Brand, G.D.; Ramada, M.H.S.; Genaro-Mattos, T.C.; Bloch, C. Towards an Experimental Classification System for Membrane Active Peptides. *Sci. Rep.* **2018**, *8*, 1194. [[CrossRef](#)]
39. Huang, Y.W.; Lee, H.J. Cell-Penetrating Peptides for Medical Theranostics and Targeted Drug Delivery. In *Peptide Applications in Biomedicine, Biotechnology and Bioengineering*; Elsevier Inc.: Amsterdam, The Netherlands, 2018; pp. 359–370. ISBN 9780081007426.
40. Porosk, L.; Gaidutšik, I.; Langel, Ü. Approaches for the Discovery of New Cell-Penetrating Peptides. *Expert Opin. Drug Discov.* **2021**, *16*, 553–565. [[CrossRef](#)]
41. Cardoso, M.H.; Orozco, R.Q.; Rezende, S.B.; Rodrigues, G.; Oshiro, K.G.N.; Cândido, E.S.; Franco, O.L. Computer-Aided Design of Antimicrobial Peptides: Are We Generating Effective Drug Candidates? *Front. Microbiol.* **2020**, *10*, 3097. [[CrossRef](#)]
42. Lertampaiporn, S.; Vorapreeda, T.; Hongsthong, A.; Thammarongtham, C. Ensemble-AmpPred: Robust Amp Prediction and Recognition Using the Ensemble Learning Method with a New Hybrid Feature for Differentiating Amps. *Genes* **2021**, *12*, 137. [[CrossRef](#)]
43. Tyagi, A.; Kapoor, P.; Kumar, R.; Chaudhary, K.; Gautam, A.; Raghava, G.P.S. In Silico Models for Designing and Discovering Novel Anticancer Peptides. *Sci. Rep.* **2013**, *3*, 2984. [[CrossRef](#)] [[PubMed](#)]
44. de Oliveira, E.C.L.; Santana, K.; Josino, L.; Lima e Lima, A.H.; de Souza de Sales, C., Jr. Predicting Cell-Penetrating Peptides Using Machine Learning Algorithms and Navigating in Their Chemical Space. *Sci. Rep.* **2021**, *11*, 7628. [[CrossRef](#)]
45. Robles-Loaiza, A.A.; Pinos-Tamayo, E.A.; Mendes, B.; Ortega-Pila, J.A.; Proaño-Bolaños, C.; Plisson, F.; Teixeira, C.; Gomes, P.; Almeida, J.R. Traditional and Computational Screening of Non-Toxic Peptides and Approaches to Improving Selectivity. *Pharmaceuticals* **2022**, *15*, 323. [[CrossRef](#)] [[PubMed](#)]
46. Gautam, A.; Chaudhary, K.; Kumar, R.; Sharma, A.; Kapoor, P.; Tyagi, A.; Raghava, G.P.S. In Silico Approaches for Designing Highly Effective Cell Penetrating Peptides. *J. Transl. Med.* **2013**, *11*, 74. [[CrossRef](#)] [[PubMed](#)]
47. Gaspar, D.; Salomé Veiga, A.; Castanho, M.A.R.B. From Antimicrobial to Anticancer Peptides. A Review. *Front. Microbiol.* **2013**, *4*, 294. [[CrossRef](#)] [[PubMed](#)]
48. Maccari, G.; di Luca, M.; Nifosi, R. In Silico Design of Antimicrobial Peptides. *Methods Mol. Biol.* **2015**, *1268*, 195–219. [[CrossRef](#)]
49. Wishart, D.S.; Feunang, Y.D.; Guo, A.C.; Lo, E.J.; Marcu, A.; Grant, J.R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; et al. DrugBank 5.0: A Major Update to the DrugBank Database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074–D1082. [[CrossRef](#)]
50. Zhao, Y.; Zhang, M.; Qiu, S.; Wang, J.; Peng, J.; Zhao, P.; Zhu, R.; Wang, H.; Li, Y.; Wang, K.; et al. Antimicrobial Activity and Stability of the D-Amino Acid Substituted Derivatives of Antimicrobial Peptide Polybia-MPI. *AMB Express* **2016**, *6*, 122. [[CrossRef](#)]
51. Pachón-Ibáñez, M.E.; Smani, Y.; Pachón, J.; Sánchez-Céspedes, J. Perspectives for Clinical Use of Engineered Human Host Defense Antimicrobial Peptides. *FEMS Microbiol. Rev.* **2017**, *41*, 323–342. [[CrossRef](#)]
52. Lombardi, L.; Maisetta, G.; Batoni, G.; Tavanti, A. Insights into the Antimicrobial Properties of Hepcidins: Advantages and Drawbacks as Potential Therapeutic Agents. *Molecules* **2015**, *20*, 6319–6341. [[CrossRef](#)]
53. Oshiro, K.G.N.; Cândido, E.S.; Chan, L.Y.; Torres, M.D.T.; Monges, B.E.D.; Rodrigues, S.G.; Porto, W.F.; Ribeiro, S.M.; Henriques, S.T.; Lu, T.K.; et al. Computer-Aided Design of Mastoparan-like Peptides Enables the Generation of Nontoxic Variants with Extended Antibacterial Properties. *J. Med. Chem.* **2019**, *62*, 8140–8151. [[CrossRef](#)] [[PubMed](#)]
54. Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatsky, M. Synthetic Therapeutic Peptides: Science and Market. *Drug Discov. Today* **2010**, *15*, 40–56. [[CrossRef](#)] [[PubMed](#)]

55. Li, Z.J.; Cho, C.H. Peptides as Targeting Probes against Tumor Vasculature for Diagnosis and Drug Delivery. *J. Transl. Med.* **2012**, *10* (Suppl S1), S1. [[CrossRef](#)] [[PubMed](#)]
56. Milletti, F. Cell-Penetrating Peptides: Classes, Origin, and Current Landscape. *Drug Discov. Today* **2012**, *17*, 850–860. [[CrossRef](#)]
57. Calis, J.J.A.; Maybeno, M.; Greenbaum, J.A.; Weiskopf, D.; de Silva, A.D.; Sette, A.; Keşmir, C.; Peters, B. Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. *PLoS Comput. Biol.* **2013**, *9*, e1003266. [[CrossRef](#)] [[PubMed](#)]
58. Thundimadathil, J. Cancer Treatment Using Peptides: Current Therapies and Future Prospects. *J. Amino Acids* **2012**, *2012*, 967347. [[CrossRef](#)]
59. Smolarczyk, R.; Cichoń, T.; Szala, S. Peptydy-Nowa Klasa Leków Przeciwnowotworowych* Peptides: A New Class of Anticancer Drugs. *Postepy Hig. Med. Dosw.* **2009**, *63*, 360–368.
60. Li, Z.J.; Cho, C.H. Development of Peptides as Potential Drugs for Cancer Therapy. *Curr. Pharm. Des.* **2010**, *16*, 1180–1189. [[CrossRef](#)]
61. Thayer, A.M. Making peptides at large scale. *Chem. Eng. News* **2011**, *89*, 81–85. [[CrossRef](#)]
62. Lau, J.L.; Dunn, M.K. Therapeutic Peptides: Historical Perspectives, Current Development Trends, and Future Directions. *Bioorg. Med. Chem.* **2018**, *26*, 2700–2707. [[CrossRef](#)]
63. Pirtskhalava, M.; Vishnepolsky, B.; Grigolava, M.; Managadze, G. Physicochemical Features and Peculiarities of Interaction of Amp with the Membrane. *Pharmaceuticals* **2021**, *14*, 471. [[CrossRef](#)] [[PubMed](#)]
64. Tan, T.; Wu, D.; Li, W.; Zheng, X.; Li, W.; Shan, A. High Specific Selectivity and Membrane-Active Mechanism of Synthetic Cationic Hybrid Antimicrobial Peptides Based on the Peptide FV7. *Int. J. Mol. Sci.* **2017**, *18*, 339. [[CrossRef](#)] [[PubMed](#)]
65. Matsuzaki, K. Control of Cell Selectivity of Antimicrobial Peptides. *Biochim. Et Biophys. Acta Biomembr.* **2009**, *1788*, 1687–1692. [[CrossRef](#)] [[PubMed](#)]
66. Matsuzaki, K.; Sugishita, K.-I.; Fujii, N.; Miyajima, K. Molecular Basis for Membrane Selectivity of an Antimicrobial Peptide, Magainin 2. *Biochemistry* **1995**, *34*, 3423–3429. [[CrossRef](#)]
67. Yeaman, M.R.; Yount, N.Y. Mechanisms of Antimicrobial Peptide Action and Resistance. *Pharmacol. Rev.* **2003**, *55*, 27–55. [[CrossRef](#)]
68. Jiménez-Rodríguez, A.; Guardado-Félix, D.; Antunes-Ricardo, M. Challenges and Strategies for Topical and Transdermal Delivery of Bioactive Peptides. *Crit. Rev. Ther. Drug Carr. Syst.* **2022**, *39*, 1–31. [[CrossRef](#)]
69. Kumar, P.; Kizhakkedathu, J.N.; Straus, S.K. Antimicrobial Peptides: Diversity, Mechanism of Action and Strategies to Improve the Activity and Biocompatibility in Vivo. *Biomolecules* **2018**, *8*, 4. [[CrossRef](#)]
70. Klepach, A.; Tran, H.; Ahmad Mohammed, F.; ElSayed, M.E.H. Characterization and Impact of Peptide Physicochemical Properties on Oral and Subcutaneous Delivery. *Adv. Drug Deliv. Rev.* **2022**, *186*, 114322. [[CrossRef](#)]
71. Jobin, M.L.; Blanchet, M.; Henry, S.; Chaignepain, S.; Manigand, C.; Castano, S.; Lecomte, S.; Burlina, F.; Sagan, S.; Alves, I.D. The Role of Tryptophans on the Cellular Uptake and Membrane Interaction of Arginine-Rich Cell Penetrating Peptides. *Biochim. Et Biophys. Acta Biomembr.* **2015**, *1848*, 593–602. [[CrossRef](#)]
72. Boman, H.G. Antibacterial Peptides: Basic Facts and Emerging Concepts. *J. Intern. Med.* **2003**, *254*, 197–215. [[CrossRef](#)]
73. Nasiri, F.; Atanaki, F.F.; Behrouzi, S.; Kavousi, K.; Bagheri, M. CpACpP: In Silico Cell-Penetrating Anticancer Peptide Prediction Using a Novel Bioinformatics Framework. *ACS Omega* **2021**, *6*, 19846–19859. [[CrossRef](#)] [[PubMed](#)]
74. Behzadi, M.; Arasteh, S.; Bagheri, M. Palmitoylation of Membrane-Penetrating Magainin Derivatives Reinforces Necroptosis in A549 Cells Dependent on Peptide Conformational Propensities. *ACS Appl. Mater. Interfaces* **2020**, *12*, 56815–56829. [[CrossRef](#)]
75. Doneva, N.; Doytchinova, I.; Dimitrov, I. Predicting Immunogenicity Risk in Biopharmaceuticals. *Symmetry* **2021**, *13*, 388. [[CrossRef](#)]
76. Dhanda, S.K.; Mahajan, S.; Paul, S.; Yan, Z.; Kim, H.; Jespersen, M.C.; Jurtz, V.; Andreatta, M.; Greenbaum, J.A.; Marcatili, P.; et al. IEDB-AR: Immune Epitope Database—Analysis Resource in 2019. *Nucleic Acids Res.* **2019**, *47*, W502–W506. [[CrossRef](#)] [[PubMed](#)]
77. Kuriakose, A.; Chirmule, N.; Nair, P. Immunogenicity of Biotherapeutics: Causes and Association with Posttranslational Modifications. *J. Immunol. Res.* **2016**, *2016*, 1298473. [[CrossRef](#)] [[PubMed](#)]
78. Shankar, G.; Arkin, S.; Cocea, L.; Devanarayan, V.; Kirshner, S.; Kromminga, A.; Quarmby, V.; Richards, S.; Schneider, C.K.; Subramanyam, M.; et al. Assessment and Reporting of the Clinical Immunogenicity of Therapeutic Proteins and Peptides—Harmonized Terminology and Tactical Recommendations. *AAPS J.* **2014**, *16*, 658–673. [[CrossRef](#)]
79. Dimitrov, I.; Bangov, I.; Flower, D.R.; Doytchinova, I. AllerTOP v.2—A Server for in Silico Prediction of Allergens. *J. Mol. Modeling* **2014**, *20*, 2278. [[CrossRef](#)]
80. Gupta, S.; Kapoor, P.; Chaudhary, K.; Gautam, A.; Kumar, R.; Raghava, G.P.S. In Silico Approach for Predicting Toxicity of Peptides and Proteins. *PLoS ONE* **2013**, *8*, e73957. [[CrossRef](#)]
81. Gupta, S.; Kapoor, P.; Chaudhary, K.; Gautam, A.; Kumar, R.; Raghava, G.P.S. Peptide Toxicity Prediction. *Methods Mol. Biol.* **2015**, *1268*, 143–157. [[CrossRef](#)]
82. Chowdhury, A.S.; Reehl, S.M.; Kehn-Hall, K.; Bishop, B.; Webb-Robertson, B.J.M. Better Understanding and Prediction of Antiviral Peptides through Primary and Secondary Structure Feature Importance. *Sci. Rep.* **2020**, *10*, 19260. [[CrossRef](#)]
83. Gleenberg, I.O.; Herschhorn, A.; Hizi, A. Inhibition of the Activities of Reverse Transcriptase and Integrase of Human Immunodeficiency Virus Type-1 by Peptides Derived from the Homologous Viral Protein R (Vpr). *J. Mol. Biol.* **2007**, *369*, 1230–1243. [[CrossRef](#)]

84. Qureshi, A.; Kaur, G.; Kumar, M. AVCPred: An Integrated Web Server for Prediction and Design of Antiviral Compounds. *Chem. Biol. Drug Des.* **2017**, *89*, 74–83. [[CrossRef](#)]
85. Littler, E.; Oberg, B. Achievements and challenges in antiviral drug discovery. *Antiviral Chem. Chemother.* **2005**, *16*, 155–168. [[CrossRef](#)]
86. Chiangjong, W.; Chutipongtanate, S.; Hongeng, S. Anticancer Peptide: Physicochemical Property, Functional Aspect and Trend in Clinical Application (Review). *Int. J. Oncol.* **2020**, *57*, 678–696. [[CrossRef](#)]
87. Sok, M.; Šentjurs, M.; Schara, M. Membrane fluidity Characteristics of Human Lung Cancer. *Cancer Lett.* **1999**, *139*, 215–220. [[CrossRef](#)]
88. Gaspar, D.; Freire, J.M.; Pacheco, T.R.; Barata, J.T.; Castanho, M.A.R.B. Apoptotic Human Neutrophil Peptide-1 Anti-Tumor Activity Revealed by Cellular Biomechanics. *Biochim. Et Biophys. Acta Mol. Cell Res.* **2015**, *1853*, 308–316. [[CrossRef](#)]
89. Deslouches, B.; Di, Y.P. Antimicrobial Peptides with Selective Antitumor Mechanisms: Prospect for Anticancer Applications. *Oncotarget* **2017**, *8*, 46635. [[CrossRef](#)]
90. Dobrzyńska, I.; Szachowicz-Petelska, B.; Law Sulkowski, S.; Figaszewski, Z. *Changes in Electric Charge and Phospholipids Composition in Human Colorectal Cancer Cells*; Springer: Berlin/Heidelberg, Germany, 2005; Volume 276.
91. Figueiredo, C.R.; Matsuo, A.L.; Massaoka, M.H.; Polonelli, L.; Travassos, L.R. Anti-Tumor Activities of Peptides Corresponding to Conserved Complementary Determining Regions from Different Immunoglobulins. *Peptides* **2014**, *59*, 14–19. [[CrossRef](#)]
92. Cooper, B.M.; Iegre, J.; O'Donovan, D.H.; Öwegård Halvarsson, M.; Spring, D.R. Peptides as a Platform for Targeted Therapeutics for Cancer: Peptide-Drug Conjugates (PDCs). *Chem. Soc. Rev.* **2021**, *50*, 1480–1494. [[CrossRef](#)] [[PubMed](#)]
93. Xiong, G.; Wu, Z.; Yi, J.; Fu, L.; Yang, Z.; Hsieh, C.; Yin, M.; Zeng, X.; Wu, C.; Lu, A.; et al. ADMETlab 2.0: An Integrated Online Platform for Accurate and Comprehensive Predictions of ADMET Properties. *Nucleic Acids Res.* **2021**, *49*, W5–W14. [[CrossRef](#)] [[PubMed](#)]
94. Bockus, A.T.; McEwen, C.M.; Lokey, R.S. Form and Function in Cyclic Peptide Natural Products: A Pharmacokinetic Perspective. *Curr. Top. Med. Chem.* **2013**, *13*, 821–836. [[CrossRef](#)] [[PubMed](#)]
95. Diaz-Eufracio, B.I.; Palomino-Hernández, O.; Houghten, R.A.; Medina-Franco, J.L. Exploring the Chemical Space of Peptides for Drug Discovery: A Focus on Linear and Cyclic Penta-Peptides. *Mol. Divers.* **2018**, *22*, 259–267. [[CrossRef](#)] [[PubMed](#)]
96. Santos, G.B.; Ganesan, A.; Emery, F.S. Oral Administration of Peptide-Based Drugs: Beyond Lipinski's Rule. *ChemMedChem* **2016**, *11*, 2245–2251. [[CrossRef](#)] [[PubMed](#)]
97. Daina, A.; Zoete, V. A Boiled-Egg to Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. *ChemMedChem* **2016**, *11*, 1117–1121. [[CrossRef](#)] [[PubMed](#)]
98. Lipinski, C.A.; Dominy, B.W.; Feeney, P.J. Drug Delivery Reviews Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3–25. [[CrossRef](#)]
99. Veber, D.F.; Johnson, S.R.; Cheng, H.Y.; Smith, B.R.; Ward, K.W.; Kopple, K.D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45*, 2615–2623. [[CrossRef](#)]
100. Doak, B.C.; Over, B.; Giordanetto, F.; Kihlberg, J. Oral Druggable Space beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chem. Biol.* **2014**, *21*, 1115–1142. [[CrossRef](#)]
101. Lovering, F.; Bikker, J.; Humblet, C. Escape from Flatland: Increasing Saturation as an Approach to Improving Clinical Success. *J. Med. Chem.* **2009**, *52*, 6752–6756. [[CrossRef](#)]
102. Ritchie, T.J.; MacDonald, S.J.F.; Young, R.J.; Pickett, S.D. The Impact of Aromatic Ring Count on Compound Developability: Further Insights by Examining Carbo- and Hetero-Aromatic and -Aliphatic Ring Types. *Drug Discov. Today* **2011**, *16*, 164–171. [[CrossRef](#)]
103. Gautier, R.; Douguet, D.; Antony, B.; Drin, G. HELIQUEST: A Web Server to Screen Sequences with Specific α -Helical Properties. *Bioinformatics* **2008**, *24*, 2101–2102. [[CrossRef](#)]
104. Santos, J.; Iglesias, V.; Pintado, C.; Santos-Suárez, J.; Ventura, S. Disphred: A Server to Predict Ph-Dependent Order–Disorder Transitions in Intrinsically Disordered Proteins. *Int. J. Mol. Sci.* **2020**, *21*, 5814. [[CrossRef](#)] [[PubMed](#)]
105. Khandogin, J.; Chen, J.; Brooks III, C.L. Exploring Atomistic Details of PH-Dependent Peptide Folding. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 18546–18550. [[CrossRef](#)] [[PubMed](#)]
106. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)] [[PubMed](#)]
107. Burdukiewicz, M.; Sidorczuk, K.; Rafacz, D.; Pietluch, F.; Chilimoniuk, J.; Rödiger, S.; Gagat, P. Proteomic Screening for Prediction and Design of Antimicrobial Peptides with Ampgram. *Int. J. Mol. Sci.* **2020**, *21*, 4310. [[CrossRef](#)] [[PubMed](#)]
108. Thomas, S.; Karnik, S.; Barai, R.S.; Jayaraman, V.K.; Idicula-Thomas, S. CAMP: A Useful Resource for Research on Antimicrobial Peptides. *Nucleic Acids Res.* **2009**, *38*, D774–D780. [[CrossRef](#)]
109. Torrent, M.; Pulido, D.; de La Torre, B.G.; García-Mayoral, M.F.; Nogués, M.V.; Bruix, M.; Andreu, D.; Boix, E. Refining the Eosinophil Cationic Protein Antibacterial Pharmacophore by Rational Structure Minimization. *J. Med. Chem.* **2011**, *54*, 5237–5244. [[CrossRef](#)]
110. Tang, H.; Su, Z.D.; Wei, H.H.; Chen, W.; Lin, H. Prediction of Cell-Penetrating Peptides with Feature Selection Techniques. *Biochem. Biophys. Res. Commun.* **2016**, *477*, 150–154. [[CrossRef](#)]
111. Schaduangrat, N.; Nantasenamat, C.; Prachayasittikul, V.; Shoombuatong, W. ACPred: A Computational Tool for the Prediction and Analysis of Anticancer Peptides. *Molecules* **2019**, *24*, 1973. [[CrossRef](#)]

112. Schaduangrat, N.; Nantasenamat, C.; Prachayasittikul, V.; Shoombuatong, W. Meta-Iavp: A Sequence-Based Meta-Predictor for Improving the Prediction of Antiviral Peptides Using Effective Feature Representation. *Int. J. Mol. Sci.* **2019**, *20*, 5743. [[CrossRef](#)]
113. Baell, J.B.; Holloway, G.A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53*, 2719–2740. [[CrossRef](#)]
114. Lamiable, A.; Thevenet, P.; Rey, J.; Vavrusa, M.; Derreumaux, P.; Tuffery, P. PEP-FOLD3: Faster Denovo Structure Prediction for Linear Peptides in Solution and in Complex. *Nucleic Acids Res.* **2016**, *44*, W449–W454. [[CrossRef](#)] [[PubMed](#)]
115. Roman Laskowski, B.A.; Macarthur, M.W.; Thornton, J.M. Computer Programs PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* **1983**, *13*, 283–291.
116. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [[CrossRef](#)] [[PubMed](#)]