Songklanakarin J. Sci. Technol.
47 (2), 102–112, Mar. – Apr. 2025

http://www.sjst.psu.ac.th

Original Article

In silico assessment of self-assembly and bioactivity profiles of peptides from lentil protein

Matthew D. Ezema[1, 2], Toluwase H. Fatoki[2, 3*], Seyi Ogunrinola[1], and Chibuike C. Udenigwe[1*]

[1] *School of Nutrition Sciences, Faculty of Health sciences, University of Ottawa, Ottawa, Ontario, Canada*

[2] *Department of Biochemistry, Federal University Oye-Ekiti, Oye-Ekiti, Ekiti State, Nigeria*

[3] *Department of Chemistry and Biochemistry, Concordia University, Montreal, Quebec, Canada*

Received: 17 August 2024; Revised: 10 September 2024; Accepted: 11 March 2025

Abstract

Peptides are versatile biomolecules with diverse biological functions, including self-assembly into nanostructures such as nanoparticles, nanosheets, and nanotubes, which have applications in bioadhesives, hydrogels, and drug delivery systems. Lentil proteins, known for their antioxidant, antifungal, antihypertensive, and antidiabetic properties, are promising candidates for such biomaterial applications. This study aimed to computationally explore the self-assembly properties of lentil peptides derived from lentil protein-tannic acid (LPTA) interactions, previously identified through mass spectrometry. A database of 1,146 peptides was analyzed using peptide self-aggregation prediction, bioactivity profiling, physicochemical characterization, 3D structures modeling, and molecular dynamics simulations. The results identified 44 amyloid-like peptides, of which seven were non-toxic and poorly soluble, including peptide 1137 (YVIVNDSCYHQLVSHWLNT), peptide 494 (IVHIAKQL) and peptide 711 (LVVVPQNF). These peptides demonstrated a range of bioactivities, including antihypertensive, antidiabetic, antimicrobial, antimalarial, quorum sensing, anticancer, and antioxidant activities. Peptide 1137 exhibited moderate self-assembly, balanced bioactivity, and high structural stability, making it the most promising candidate. Poor solubility, attributed to the presence of tryptophan (W) and histidine (H), was observed. In conclusion, this study underscores the potential of lentil-derived peptides as biomaterials, emphasizing the benefits of poor solubility for controlled bioactive release. Future work should validate these computational predictions experimentally, optimize safety, solubility and stability, and explore advanced biomaterial applications in tissue engineering and drug delivery systems.

Keywords: lentil peptides, self-assembly, amyloid formation, bioactivity, 3D structure, molecular dynamics

1. Introduction

Proteins are integral food components that provide nutritional, technological, and functional properties. Peptides of less than 40 amino acids are ubiquitously generated during proteolytic cleavage or maturation of functional proteins and exhibit similar or unrelated biological properties in comparison to the parent protein (Fatoki, Aluko, & Udenigwe, 2022; Udenigwe, Abioye, Okagu, & Obeme-Nmom, 2021). Peptides serve diverse biological functions and the knowledge of the conformational ensemble of polypeptides in various experimental conditions is important for biological applications. Peptides that can assemble spontaneously into a variety of nanostructures (nanoparticles, nanosheets, nanofibers, and nanotubes) find applications in a wide variety

*Corresponding author
Email address: toluwase.fatoki@fuoye.edu.ng; cudenigw@uottawa.ca

of fields, such as bioadhesives, hydrogels, and drug delivery systems (Miserez, Yu, & Mohammadi, 2023; Tufféry & Derreumaux, 2023).

Amyloid-like protein/peptide nanofibrils (PNFs) form through self-assembly of partly or fully denatured proteins into highly ordered, β-sheet rich fibrous structures, and give solutions with low ionic strength at a low pH (typically around 2) when heated to 70−90 °C, which are generally favorable conditions for PNF formation (Josefsson *et al*., 2020; Ye, Lendel, Langton, Olsson, & Hedenqvist, 2019). It has been reported that proteins from several sources such as proteins from hen egg, soy, bovine whey, and various legumes, could form amyloid-like nanofibrils (Josefsson *et al*., 2020). Peptides have gained considerable interest as therapeutic agents, particularly to target protein-protein interactions (Cabri *et al*., 2021; Fatoki, Chukwuejim, Udenigwe, & Aluko, 2023). They are currently used in the development of new functional biomimetic materials (Levin *et al*., 2020). Natural biomaterials derived from renewable resources, such as plants, animals and microorganisms, exhibit a large diversity of unique yet complex constituents, microstructures, and physiological properties, and they usually offer a biological support suitable for cell attachment and growth with a diverse set of functions in their native setting (Huang *et al*., 2017; Ullah & Chen, 2020).

Lentil (*Lens culinaris* L.) is a small legume seed belonging to the family Leguminosae (Fabaceae or Papilionaceae). The name **"**lentil**"** derives from its typical lens-shaped seeds (Samaranayaka, 2017). Lentil is classified as a pulse and ranked the third-most important cool-season grain legume in the world after chickpea and pea (Sehgal, *et al*, 2021). Lentils are a rich source of protein making them a great alternative to meat or fish. Lentil protein, like other pulse protein, is a good source of the essential amino acids, particularly leucine, lysine, threonine, and phenylalanine, but lentil is deficient in methionine, tryptophan and cysteine (Samaranayaka, 2017; Zeece, 2020). The composition of lentils is approximately 28% protein, 63% total carbohydrate (47% of which is starch and 12% dietary fiber), and only about 1% fat (Maqbool, Bakhsh, & Aksoy, 2021; Thavarajah, Ruszkowski, & Vandenberg, 2008; Zeece, 2020;). Principal proteins in lentils are characterized as albumins, globulins, 7S vicilin, and 11S legumins, which make up most of its extractable protein (Shevkani *et al*., 2019; Zeece, 2020).

Lentil protein has been known as a strong antioxidant, having also antifungal, antihypertensive, and antidiabetic properties (Boachie *et al*., 2022). Lentil peptides have shown promising bioactive properties regarding the antioxidant activity and also inhibitory activity of angiotensin-I-converting enzyme (ACE). Sequential hydrolysis of proteins has shown a higher degree of hydrolysis with enhanced antioxidant and ACE-inhibitory activities (Rezvankhah, Yarmand, Ghanbarzadeh, & Mirzaee, 2023).

In this study, an *in silico* approach was utilized, as it has been applied to elucidate peptides biological activities and predict their 3D structures, in isolation, or in interaction with a receptor (Fatoki, Aluko, & Udenigwe, 2022; Langyan *et al*., 2021). Accurate and fast structure prediction of peptides of less than 40 amino acids in aqueous solution has many biological applications, but their conformations are pH- and salt concentration-dependent (Rey, Murail, de Vries, Derreumaux, & Tuffery, 2023). A prior study has used a nearly similar *in silico* approach to study bioactive peptides from sequences of fatty acid desaturase 3 (FAD3) of flaxseed protein, where aggregation, allergenicity, toxicity, water solubility and physicochemical properties were predicted (Langyan *et al*., 2021).

In silico approaches can aid to predict the aggregation propensity and amyloidogenicity as well as identify aggregation-prone regions (Prabakaran, Rawat, Kumar, & Gromiha, 2021). Computational tools enable the rapid prediction of peptide sequences with desirable bioactivity, such as antioxidant, antihypertensive, or antimicrobial properties, by analyzing protein databases and simulating enzymatic hydrolysis. This accelerates the identification of bioactive peptides without the need for extensive wet-lab experiments. Furthermore, bioactivity prediction, molecular docking, molecular structure modeling and molecular dynamic simulation studies can help to predict peptide self-assembly behaviors and stability in physiological conditions, providing insights into safety and biological applications (Peredo-Lovillo, Hernández-Mendoza, Vallejo-Cordoba, & Romero-Luna, 2021; Prabakaran, Rawat, Kumar, & Gromiha, 2021). This approach is cost-effective, time-efficient, and environmentally sustainable, aiding in the development of functional biomaterials.

Therefore, the aim of this study was to computationally assess the self-assembly and bioactivity profile of lentil peptides obtained by mass spectrometry analyses in our previous study, based on lentil protein - tannic acid (LPTA) interactions. In our previous study (Boachie *et al*., 2022), a total of 1,146 peptides were identified from the high molecular weight hydrolysates of lyophilized lentil protein (LPI) and LPTA 0.5% w/v, with 589 overlapping peptides among the two samples, and the molecular masses of these peptides were between 698.4 and 3,095.5 Da. This suggested that a 0.5% w/v concentration of tannic acid promoted protein aggregation in lentil proteins, thereby impairing their digestibility. This effect was attributed to tannic acid's interference with pepsin accessibility to the protein matrix, ultimately altering the composition and profile of the peptides released during digestion (Boachie *et al*., 2022).

2. Materials and Methods

2.1 Materials

In this study, a laptop PC with Core i5 processor and 16 GB RAM was used to access standard bioinformatics tools such as software, databases, and webservers.

2.2 Lentil Peptide Sequence Preparation

The lentil peptides database used in this study was generated in our previous work on lentil protein - tannic acid (LPTA) interaction (Boachie *et al*., 2022). Briefly, 250 g of dry green lentil seeds were soaked in water at 20°C for 10 hours, then partly ground and freeze-dried before finally pulverizing into a fine powder. The flour (10% w/v) was prepared in NaOH (0.05 M) and stirred constantly for 4 h with pH maintained at 10.0. The suspension was then centrifuged at 7,000g and 4 °C for 30 min. The supernatant was recovered,

pH adjusted to 4.0 using 3 M HCl, and stirred for 2 h to precipitate the lentil proteins; then recentrifuged. The pellet fraction was then resuspended in deionized water and pH adjusted to 7.0 with 3 M NaOH, after which it was lyophilized. The lyophilized lentil protein (LLP) was stored at −20 °C. The LPTA complexation solutions were prepared from the stock solutions of LLP (2% w/v) and tannic acid (TA, 1% w/v) in deionized water, to obtain mixtures containing 1% w/v LP and 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, or 0.5% TA (ref 18). In vitro peptic hydrolysis of LPTA was carried out by mixing 30 µL pepsin stock solution (25,000 U/mL) prepared in Milli Q water and 5 mL solutions of the LPTA complexes, with pH adjusted to pH 3.0 using 1 M HCl. The mixture samples were incubated at 37 °C for 2 h with constant shaking at 90 rpm. Afterward, the pH of digesta was adjusted to 7.0, and the mixture was vortex-mixed and frozen to halt enzymatic activities. The samples (LLP and LPTA 0.5% digesta) were adjusted to pH 4.0, and centrifuged for 30 min at 7,000g and 20 °C. The precipitates were then freeze-dried for characterization of the peptides by Liquid Chromatography with Tandem Mass Spectrometry (LC−MS/MS) approach. Elaborate steps and analyses were provided in our previous work (Boachie *et al*., 2022). The peptide sequences obtained were used in this present computational study.

2.3 Peptide self-aggregation prediction

The LPTA peptide database consisting of 1,146 peptides was obtained as described above and curated from our previous study (Boachie *et al*., 2022). The 1,146 peptides were converted to FASTA format, and used to predict potential self-aggregation properties using the Prediction of Amyloid STructure Aggregation 2.0 (PASTA 2.0) web server: available at http://old.protein.bio.unipd.it/pasta2/index.html (Walsh, Seno, Tosatto, & Trovato, 2014), and Generalized Aggregation Proneness (GAP) webserver, available at https://www.iitm.ac.in/bioinfo/GAP (Thangakani, Kumar, Nagarajan, Velmurugan, & Gromiha, 2014). The PASTA webserver ranked these peptides based on the possibility of forming aggregation, and the positive cases (44 peptides) obtained from PASTA were used on GAP webserver to determine the probability of forming aggregation.

2.4 Bioactivity prediction of self-aggregating peptides

The bioactivities such as anticancer, antioxidant, antimicrobial, cell penetration and neuroactivity, of the identified 44 self-aggregating peptides, were predicted on the Universal Deep Learning Architecture for Bioactive Peptide Prediction (UniDL4BioPep) webserver, available at https://nepc2pvmzy.us-east-1.awsapprunner.com (Du, Ding., Xu, & Li, 2023).

2.5 Physicochemical properties prediction

The physicochemical properties, such as isoelectric point (pI) and solubility, of the identified 44 peptides with good bioactivity, were predicted on the peptide solubility calculator webserver, available at https://pepcalc.com/peptide-solubility-calculation.php.

2.6 Peptide 3D structure prediction

The three-dimensional (3D) structures of the identified 44 peptides were predicted on PEP-FOLD4 webserver, available at https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD4/, which was built based on the coarse grained OPEP force field and Debye-Hueckel for instantiated pH and salt concentration conditions (Maupetit, Tuffery, & Derreumaux, 2007; Rey *et al*., 2023; Tuffery and Derreumaux, 2023). The predictions were carried out using Monte Carlo steps 0 and temperature 370 K, solvent pH 7, and 1 mM ionic strength, while other parameters were at default settings. The sOPEP (optimized potential energy for proteins) score was used to rank the model, and it is a simplified potential used to evaluate the energetic quality of protein models. The goal is to identify conformations with lower energies that indicate better structural candidates for further analysis or validation.

2.7 Molecular dynamics simulation

The predicted 3D structures of the 44 peptides were subjected to short molecular dynamics (MD) simulation on BioExcel Building Block (BioBB) webserver, https://mmb.irbbarcelona.org/biobb-wfs (Bayarri, Andrio, Hospital, Orozco, & Gelpí, 2022), using protein MD gromacs workflow that consist of AMBER99SB forcefield (Hornak *et al*., 2006), TIP3P water model, octahedron box extending 0.8 nm beyond the peptide; and 0.05 M NaCl was included to achieve neutrality. The energy minimization was conducted and the system was equilibrated. Finally, the production runs of 0.5 ns (500 ps) were implemented based on 250,000 total steps of 2fs timestep with all-atom output frames saved every 1000 steps. The potential and total energies, radius of gyration (Rg) and root-mean square deviation (RMSD) were obtained.

3. Results and Discussion

3.1 Effect of self-aggregation profiles prediction of lentil peptide database

The need for peptide self-assembly study has become paramount due to the emerging applications in biomedicine and material science. Therefore, identifying peptides with self-aggregating potential from lentil protein has become an interesting topic. Amyloid fibrils with the common cross-β pattern are often formed by various proteins and peptides, and they usually consist of several protofibrils and have up to several micrometers in length, with 6-12 nanometers width (Ke *et al*., 2020). Differences in the hydrolysis protocol could lead to different peptide fragments obtained. The results of self-aggregation potential of the lentil peptides from our database that consisted of 1,146 peptides showed that only 44 peptides were amyloid-like at the energy threshold of -5.00 PASTA energy unit (PEU), which is equivalent to 5.960 kcal.mol^{-1} (Table 1). The peptides 495-499 showed the best energy score of 10.038397 PEU as well as the highest number (20) of amyloids that could be formed. Peptide 344 has the least energy of -5.10079 PEU but has 3 potential amyloids, while 11 peptides have only 1 possible amyloid. The results showed that peptide 499 (IVIVTVNEGKGDFELVGQRNENQQE) and peptide 289

Table 1. Selected lentil self-aggregating peptides obtained at a threshold of -5.00 PASTA energy units with GAP amyloid probability

SN	Peptide custom database ID	Sequence	Self-aggregation Profile							
			Length	Amyloids (#)	PASTA energy unit	Disorder (%)	α-Helix (%)	β-Strand (%)	% Coil	GAP (YES) amyloids probability
1	497	IVIVTVNEGKGDFE	14	20	-10.038397	100.0	0.0	35.71	64.29	0.863
2	498	IVIVTVNEGKGDFEL	15	20	-10.038397	66.66	0.0	33.33	66.67	0.871
3	499	IVIVTVNEGKGDFELVGQRNENQQE	25	20	-10.038397	68.0	0.0	36.0	64.0	0.799
4	496	IVIVTVNEGKGDF	13	20	-10.038397	100.0	0.0	38.46	61.54	0.969
5	495	IVIVTVNEGKGD	12	20	-10.038397	100.0	0.0	41.67	58.33	0.964
6	994	VIVTVNEGKGDF	12	6	-8.065847	100.0	0.0	33.33	66.67	0.964
7	995	VIVTVNEGKGDFEL	14	6	-8.065847	100.0	0.0	28.57	71.43	0.857
8	872	SPGDVVIIPAGHPVA	15	5	-7.953472	100.0	0.0	26.67	73.33	0.857
9	455	ILVVLSGKAIL	11	11	-7.49436	100.0	0.0	36.36	63.64	0.927
10	1137	YVIVNDSCYHQLVSHWLNT	19	15	-6.956513	21.05	63.16	5.26	31.58	0.972
11	500	IVIWTASAL	9	6	-6.931236	100.0	0.0	44.44	55.56	0.921
12	417	IIIWTASAL	9	6	-6.8996	100.0	0.0	44.44	55.56	0.921
13	400	IFIIQGKGVLG	11	2	-6.806883	100.0	0.0	27.27	72.73	0.695
14	399	IFIIQGKGVL	10	2	-6.806883	100.0	0.0	30.0	70.0	0.634
15	711	LVVVPQNF	8	1	-6.769277	100.0	0.0	12.5	87.5	0.992
16	531	KLVIVGDGGTGKTTF	15	3	-6.737641	100.0	0.0	26.67	73.33	0.910
17	775	NVIVKVSKKQIEE	13	5	-6.645982	100.0	15.38	30.77	53.85	0.828
18	774	NVIVKISRKQIEE	13	5	-6.614346	100.0	23.08	30.77	46.15	0.709
19	993	VIVKISRKQIEE	12	2	-6.614346	100.0	25.0	33.33	41.67	0.710
20	1046	VVIGHVDSGKSTTTGHL	17	2	-6.567368	100.0	0.0	23.53	76.47	0.865
21	1143	YVVVNDSCYHQL	12	3	-6.446522	100.0	0.0	25.0	75.0	0.991
22	493	IVFWMYNDQDTPVIA	15	4	-6.379771	73.33	0.0	33.33	66.67	0.835
23	806	QDLDIFVNSVE	11	2	-6.182266	100.0	0.0	27.27	72.73	0.736
24	571	LDIFVNSVE	9	2	-6.182266	100.0	0.0	33.33	66.67	0.747
25	231	FEITPEKNPQLQDLDIFVNSVE	22	2	-6.182266	100.0	22.73	0.0	77.27	0.790
26	922	TSNIANQLDSTPRVFYLV	18	1	-6.088288	100.0	22.22	22.22	55.56	0.740
27	507	IVTVNEGKGDFE	12	2	-6.061661	100.0	0.0	25.0	75.0	0.824
28	508	IVTVNEGKGDFEL	13	2	-6.061661	100.0	0.0	23.08	76.92	0.839
29	509	IVTVNEGKGDFELVGQRNENQQE	23	2	-6.061661	100.0	0.0	30.43	69.57	0.777
30	506	IVTVNEGKGDF	11	2	-6.061661	100.0	0.0	27.27	72.73	0.958
31	494	IVHIAKQL	8	2	-5.866514	100.0	50.0	0.0	50.0	0.989
32	416	IIAVPTGIVFW	11	3	-5.767179	100.0	0.0	18.18	81.82	0.625
33	823	QVVNCNGNTVFDGELEAGRA	20	1	-5.652111	100.0	0.0	5.0	95.0	0.797
34	1001	VLLVNEGKGNLEL	13	2	-5.52181	100.0	0.0	23.08	76.92	0.756
35	1062	WMYNDQDTPVIAV	13	1	-5.480359	100.0	0.0	15.38	84.62	0.808
36	824	REGDIIAVPTGIVF	14	1	-5.448723	57.14	0.0	21.43	78.57	0.696
37	289	FRKGDIIAIPSGIPYWTYNHGDEPL	25	1	-5.417087	52.0	0.0	24.0	76.0	0.718
38	384	IDVMRVIDE	9	1	-5.395658	100.0	0.0	55.56	44.44	0.848
39	1141	YVVFKTNDRAAVSHVQQ	17	1	-5.331569	100.0	35.29	23.53	41.18	0.864
40	1140	YVVFKTNDRAAVSHV	15	1	-5.331569	100.0	13.33	26.67	60.0	0.839
41	1142	YVVFKTNDRAAVSHVQQV	18	1	-5.331569	100.0	44.44	22.22	33.33	0.873
42	886	SVIIPATTGEMGVLPGHVAT	20	1	-5.283375	100.0	0.0	10.0	90.0	0.977
43	295	FTTITVNEPAPGVKAILS	18	2	-5.108186	100.0	0.0	38.89	61.11	0.834
44	344	GLFGGAGVGKTVLIMELIN	19	3	-5.100790	47.36	52.63	0.0	47.37	0.792

Note: 1 PASTA energy unit is equivalent to 1.192 kcal.mol^{-1}

(FRKGDIIAIPSGIPYWTYNHGDEPL) have amino acid residue length of 25, but the differences in composition resulted in aggregation profile.

The amino acid residue IIPA of the peptide 872 and 886 was found to be present in the peptide fragments reported by Josefsson *et al*. (2020). It has been suggested that the hydrophobic side chains in peptides are usually the driving force to aggregation (Roberts, 2014). An increase in peptide concentration increases soluble aggregation (Shire, Shahrokh, & Liu, 2004). Manufacturing processes to fabricate a versatile range of materials are usually achieved from aqueous solutions by altering conditions such as pH, temperature, protein concentration, ionic strength, and composition, which are the factors that generally increase peptide aggregation or affect stability (Høgstedt, Østergaard, Weiss, Sjögren, & van de Weert, 2018; Miserez *et al*., 2023; Payne & Manning, 2009).

A study has reported that highly rigid peptides (peptides with high β-sheet propensity) could assemble predominantly into fibrillar structures, and that increasing the flexibility of the peptide leads to a variety of structures, including fibrils, β-barrel structures, and amorphous aggregates (Bellesia, & Shea, 2009). Peptide/protein aggregation could form either structured amyloid aggregates or non-structured precipitated/soluble aggregates (Dobson, 2003; Høgstedt *et al*., 2018). The mechanism for soluble peptide aggregation and its relation to phase-separated aggregates and/or amyloid fibrils has not been extensively studied. Moreover, it has been well-established that arginine modulates protein aggregation, acting as a molecular

cosolvent and "chemical chaperone" in solution (Sharma, Sarkar, Paul, Roy, & Chattopadhyay, 2013). A study has revealed that amyloid-beta peptide exhibits increased solubility and decreased fibril formation in the presence of arginine, as a result of hydrophobic surfaces of arginine clusters masking the pro-aggregatory hydrophobic residues of amyloid-beta monomers, thereby preventing their self-association (Das *et al*., 2007).

3.2. Effect of bioactivity profile screening of selected lentil self-aggregating peptides

Bioactive peptides are known to be normally buried in the structure of parent protein and become active after the cleavage of the protein, hence the need to explore their bioactive applications. The bioactivity prediction results showed that all the 44 peptides obtained were found to be active anti-microbial peptides (AMP), whereas more than half of these peptides were active for quorum sensing (QS) activity. Altogether, 11 peptides had antioxidant activity (AO). Moreover, only seven peptides: peptide 1,137 (YVIVNDSCYHQLVSHWLNT), Peptide 711 (LVVVPQ NF), peptide 1,143 (YVVVNDSCYHQL), peptide 493 (IVFWMYNDQDTPVIA), peptide 571 (LDIFVNSVE), peptide 494 (IVHIAKQL), and peptide 1,062 (WMYN DQDTPVIAV) were predicted to be non-toxic (Table 2), and they possess variably suitable bioactivities, which include antihypertensive (peptides 494 and 711), antidiabetics (peptides 571 and 711), antimicrobial (peptides 494, 571, 711, 1,062, 1,137, and 1,143). antimalarial (peptides 494, 711, and 1,062), quorum sensing (peptides 494, 571, 711, 1,137, and 1,143), anticancer (peptides 493, 571, and 711) anti-methicillin resistance *S. aureus* (peptide 494), tumor T-cell antigens (peptides 571 and 1,143), antioxidant (peptides 493, 1,062, 1,137, and 1,143), antibacterial (peptide 494), blood-brain barrier permeant (peptides 493, 494, 711, 1,062, and

Table 2. Bioactivity profiles of selected lentil self-aggregating peptides obtained from UniDL4BioPep webserver

SN	Peptide custom database ID	Sequence	Bioactivity profile								
			AHT	DPPIV	AMP	AMAP main	AMAP alt	QS	ACP main	MRSA main	TTCA
1	497	IVIVTVNEGKGDFE	NA	NA	A	NA	NA	A	NA	NA	NA
2	498	IVIVTVNEGKGDFEL	NA	NA	A	NA	NA	A	NA	NA	NA
3	499	IVIVTVNEGKGDFELVGQRNENQQE	NA	NA	A	NA	NA	NA	NA	NA	NA
4	496	IVIVTVNEGKGDF	NA	NA	A	NA	NA	A	NA	NA	NA
5	495	IVIVTVNEGKGD	NA	NA	A	NA	NA	A	NA	NA	NA
6	994	VIVTVNEGKGDF	NA	NA	A	NA	NA	A	NA	NA	NA
7	995	VIVTVNEGKGDFEL	NA	NA	A	NA	NA	A	NA	NA	NA
8	872	SPGDVVIIPAGHPVA	NA	NA	A	NA	NA	A	NA	NA	NA
9	455	ILVVLSGKAIL	NA	NA	A	NA	NA	A	A	A	NA
10	1137	YVIVNDSCYHQLVSHWLNT	NA	NA	A	NA	NA	A	NA	NA	NA
11	500	IVIWTASAL	NA	NA	A	NA	NA	A	NA	NA	NA
12	417	IIIWTASAL	NA	NA	A	NA	A	A	A	NA	A
13	400	IFIIQGKGVLG	NA	NA	A	NA	NA	A	A	NA	A
14	399	IFIIQGKGVL	NA	NA	A	NA	NA	A	A	NA	A
15	711	LVVVPQNF	A	A	A	NA	A	A	A	NA	NA
16	531	KLVIVGDGGTGKTTF	NA	NA	A	NA	NA	NA	NA	NA	NA
17	775	NVIVKVSKKQIEE	NA	NA	A	NA	NA	NA	NA	NA	NA
18	774	NVIVKISRKQIEE	NA	NA	A	NA	NA	NA	NA	NA	NA
19	993	VIVKISRKQIEE	A	NA	A	NA	NA	NA	NA	NA	NA
20	1046	VVIGHVDSGKSTTTGHL	NA	NA	A	NA	NA	NA	NA	NA	NA
21	1143	YVVVNDSCYHQL	NA	NA	A	NA	NA	A	NA	NA	A
22	493	IVFWMYNDQDTPVIA	NA	NA	A	NA	NA	A	A	NA	NA
23	806	QDLDIFVNSVE	NA	A	A	NA	NA	A	NA	NA	NA
24	571	LDIFVNSVE	NA	A	A	NA	NA	A	A	NA	A
25	231	FEITPEKNPQLQDLDIFVNSVE	A	NA	A	NA	NA	NA	NA	NA	NA
26	922	TSNIANQLDSTPRVFYLV	NA	NA	A	NA	NA	A	A	NA	NA
27	507	IVTVNEGKGDFE	NA	NA	A	NA	NA	A	NA	NA	NA
28	508	IVTVNEGKGDFEL	NA	NA	A	NA	NA	A	NA	NA	NA
29	509	IVTVNEGKGDFELVGQRNENQQE	NA	NA	A	NA	NA	NA	NA	NA	NA
30	506	IVTVNEGKGDF	NA	NA	A	NA	NA	A	NA	NA	NA
31	494	IVHIAKQL	A	NA	A	NA	A	A	A	NA	A
32	416	IIAVPTGIVFW	NA	NA	A	NA	NA	A	A	NA	A
33	823	QVVNCNGNTVFDGELEAGRA	NA	NA	A	NA	NA	A	NA	NA	NA
34	1001	VLLVNEGKGNLEL	NA	NA	A	NA	NA	NA	NA	NA	NA
35	1062	WMYNDQDTPVIAV	NA	NA	A	NA	NA	A	NA	NA	NA
36	824	REGDIIAVPTGIVF	NA	NA	A	NA	NA	A	NA	NA	NA
37	289	FRKGDIIAIPSGIPYWTYNHGDEPL	NA	NA	A	NA	NA	NA	NA	NA	NA
38	384	IDVMRVIDE	NA	A	A	NA	NA	A	NA	NA	A
39	1141	YVVFKTNDRAAVSHVQQ	NA	NA	A	NA	NA	A	NA	NA	NA
40	1140	YVVFKTNDRAAVSHV	NA	NA	A	NA	NA	A	NA	NA	NA
41	1142	YVVFKTNDRAAVSHVQQV	NA	NA	A	NA	NA	A	NA	NA	NA
42	886	SVIIPATTGEMGVLPGHVAT	NA	NA	A	NA	NA	NA	NA	NA	NA
43	295	FTTITVNEPAPGVKAILS	NA	NA	A	NA	NA	A	A	NA	NA
44	344	GLFGGAGVGKTVLIMELIN	NA	NA	A	NA	NA	NA	NA	NA	A

Table 2. Continued.

SN	Peptide custom database ID	Sequence	Bioactivity profile							
			AF	AO	AV	AB	BBP	Tox	Neuro	APP
1	497	IVIVTVNEGKGDFE	NA	NA	NA	NA	NA	A	NA	NA
2	498	IVIVTVNEGKGDFEL	NA	NA	NA	NA	NA	A	NA	NA
3	499	IVIVTVNEGKGDFELVGQRNENQQE	NA	NA	NA	NA	NA	A	NA	NA
4	496	IVIVTVNEGKGDF	NA	NA	NA	NA	NA	A	NA	NA
5	495	IVIVTVNEGKGD	NA	NA	NA	NA	NA	A	NA	NA
6	994	VIVTVNEGKGDF	NA	NA	NA	NA	NA	A	NA	NA
7	995	VIVTVNEGKGDFEL	NA	NA	NA	NA	NA	A	NA	NA
8	872	SPGDVVIIPAGHPVA	A	A	NA	NA	NA	A	NA	A
9	455	ILVVLSGKAIL	NA	NA	A	A	NA	A	NA	A
10	1137	YVIVNDSCYHQLVSHWLNT	NA	A	NA	NA	A	NA	NA	NA
11	500	IVIWTASAL	NA	NA	A	A	NA	A	NA	A
12	417	IIIWTASAL	NA	NA	A	A	NA	A	NA	A
13	400	IFIIQGKGVLG	NA	A	A	A	NA	A	NA	A
14	399	IFIIQGKGVL	NA	A	A	A	NA	A	NA	A
15	711	LVVVPQNF	NA	NA	NA	NA	A	NA	A	A
16	531	KLVIVGDGGTGKTTF	NA	NA	NA	NA	NA	A	NA	NA
17	775	NVIVKVSKKQIEE	NA	NA	NA	NA	A	A	NA	NA
18	774	NVIVKISRKQIEE	NA	NA	NA	NA	A	A	NA	NA
19	993	VIVKISRKQIEE	NA	NA	NA	NA	A	A	NA	NA
20	1046	VVIGHVDSGKSTTTGHL	NA	NA	A	NA	NA	A	NA	NA
21	1143	YVVVNDSCYHQL	NA	A	NA	NA	A	NA	NA	NA
22	493	IVFWMYNDQDTPVIA	NA	A	NA	NA	A	NA	NA	NA
23	806	QDLDIFVNSVE	NA	NA	NA	NA	NA	A	NA	NA
24	571	LDIFVNSVE	NA	NA	NA	NA	NA	NA	NA	NA
25	231	FEITPEKNPQLQDLDIFVNSVE	NA	NA	NA	NA	NA	A	NA	NA
26	922	TSNIANQLDSTPRVFYLV	NA	NA	NA	NA	NA	A	NA	NA
27	507	IVTVNEGKGDFE	NA	NA	NA	NA	NA	A	NA	NA
28	508	IVTVNEGKGDFEL	NA	NA	NA	NA	NA	A	NA	NA
29	509	IVTVNEGKGDFELVGQRNENQQE	NA	NA	NA	NA	NA	A	NA	NA
30	506	IVTVNEGKGDF	NA	NA	NA	NA	NA	A	NA	NA
31	494	IVHIAKQL	NA	NA	NA	A	A	NA	NA	A
32	416	IIAVPTGIVFW	A	A	A	A	NA	A	NA	A
33	823	QVVNCNGNTVFDGELEAGRA	NA	A	NA	NA	NA	A	NA	NA
34	1001	VLLVNEGKGNLEL	NA	NA	NA	NA	NA	A	NA	A
35	1062	WMYNDQDTPVIAV	NA	A	NA	NA	A	NA	NA	NA
36	824	REGDIIAVPTGIVF	NA	NA	NA	NA	NA	A	NA	NA
37	289	FRKGDIIAIPSGIPYWTYNHGDEPL	NA	A	NA	NA	A	A	NA	NA
38	384	IDVMRVIDE	NA	NA	NA	NA	A	A	NA	NA
39	1141	YVVFKTNDRAAVSHVQQ	NA	NA	NA	NA	A	A	NA	A
40	1140	YVVFKTNDRAAVSHV	NA	NA	NA	NA	A	A	NA	A
41	1142	YVVFKTNDRAAVSHVQQV	NA	NA	NA	NA	A	A	NA	A
42	886	SVIIPATTGEMGVLPGHVAT	NA	A	NA	NA	NA	A	NA	A
43	295	FTTITVNEPAPGVKAILS	A	NA	NA	NA	NA	A	NA	NA
44	344	GLFGGAGVGKTVLIMELIN	NA	NA	NA	NA	NA	A	NA	A

Legend: AHT - antihypertensive (ACE inhibitory activity). DPPIV – dipeptidyl peptidase IV. AMP – antimicrobial peptide. AMAP – antimalarial activity (main and alternative). QS – quorum sensing activity. ACP – anticancer activity (main). MRSA – anti-methicillin resistance *S. aureus* activity (main). TTCA – tumor T-cell antigens. AF – antifungal activity. AO – antioxidant. AV – antiviral activity. AB – antibacterial activity. BBP – blood-brain barrier permeant. Tox – toxicity. Neuro – neuropeptide. APP – anti-parasitic activity. A - Active, NA - Non-active

1,143) and anti-parasitic (peptides 494 and 711) activities.

It is interesting to note that lentil peptides have such significant biological positive effects as enumerated in Table 2 on body function or condition, hence the range of attention on lentil due to the potential application in increasing the quality of health. Specifically, having the above 7 sequences as non-toxic is a welcome development. Bioactive peptides from plant sources have recently received more attention due to their cost effectiveness and lower immunogenetic effects (Akbarian, Khani, Eghbalpour, & Uversky, 2022).

Oxidative stress changes the balance between pro-oxidants/antioxidants in favour of pro-oxidant, potentially leading to biological damage, and this challenging health anomaly is rampant in people, both in developing and developed societies (Akbarian *et al*, 2022). As earlier mentioned, 11 peptides had antioxidant activity, which is associated with the presence of hydrophilic amino acids such as proline, alanine, valine, and leucine in the N- terminal position; and the amino acids tyrosine, valine, methionine, leucine, isoleucine, glutamine, and tryptophan in the C-terminal position as shown in the 11 sequences indicated in Table 2. This agrees with a study by Rezvankhah *et al*. (2023), showing that lentil peptides have promising bioactive properties regarding the antioxidant activity and also

inhibitory activity of angiotensin-I-converting enzyme (ACE). Sequential hydrolysis of proteins has shown a higher degree of hydrolysis with enhanced antioxidant and ACE-inhibitory activities.

3.3. Physicochemical properties of the selected lentil self-aggregating peptides

Physicochemical characteristics of lentil peptides were determined to establish their viability and best use in the food industry and other applications. The results on physicochemical properties of the 44 potential self-aggregating peptides obtained in the previous step of this study showed that peptides 495-499, 506-509, 571 and few others have good solubility, while peptide 1,137 (YVIVNDSCYHQLVSHWLNT), peptide 711 (LVVVP QNF), peptide 1,143 (YVVVNDSCYHQL), peptide 493 (IVFWMYNDQDTPVIA), peptide 494 (IVHIAKQL), peptide 1,062 (WMYNDQDTPVIAV) and a few others have poor solubility (Table 3). Overall, these results suggest that the presence of tryptophan (W), histidine (H) or proline (P) in a potential self-assembly peptide decreases its solubility.

Advantageously, poor solubility of biomaterials can: (i) enhance therapeutic efficacy by maintaining consistent bioavailability through slower dissolution rates, enabling sustained or controlled release; (ii) limit the systemic distribution of the bioactive material, concentrating its effects at the target site. This property is particularly beneficial in localized drug delivery systems, such as implants or topical applications; (iii) protect the bioactive material from premature degradation in aqueous environments, such as bodily fluids. This can extend the shelf life and activity of the biomaterial; (iv) reduce the risk of off-target effects and systemic toxicity by minimizing systemic absorption; (v) be useful in the fabrication of biomaterial scaffolds for tissue engineering, where gradual biodegradation and release of bioactive molecules can stimulate cellular growth and regeneration.

Table 3. Physicochemical properties of the selected lentil self-aggregating peptides

SN	Peptide custom database ID	Sequence	Physiochemical property					
			Length	Molecular weight (g/mol)	Extinction coefficient ($M^{-1}cm^{-1}$)	Iso-electric point (pH)	Net charge at pH 7	Estimated solubility
1	497	IVIVTVNEGKGDFE	14	1519.69	0	3.69	-2	Good
2	498	IVIVTVNEGKGDFEL	15	1632.85	0	3.69	-2	Good
3	499	IVIVTVNEGKGDFELVGQRNENQQE	25	2816.04	0	3.88	-3	Good
4	496	IVIVTVNEGKGDF	13	1390.58	0	3.93	-1	Good
5	495	IVIVTVNEGKGD	12	1243.41	0	3.93	-1	Good
6	994	VIVTVNEGKGDF	12	1277.42	0	3.93	-1	Good
7	995	VIVTVNEGKGDFEL	14	1519.69	0	3.69	-2	Good
8	872	SPGDVVIIPAGHPVA	15	1428.63	0	4.87	-0.9	Poor
9	455	ILVVLSGKAIL	11	1125.44	0	10.14	1	Poor
10	1137	YVIVNDSCYHQLVSHWLNT	19	2291.54	8250	6.04	-0.9	Poor
11	500	IVIWTASAL	9	973.17	5690	3.64	0	Poor
12	417	IIIWTASAL	9	987.19	5690	3.64	0	Poor
13	400	IFIIQGKGVLG	11	1144.41	0	10.14	1	Poor
14	399	IFIIQGKGVL	10	1087.35	0	10.14	1	Poor
15	711	LVVVPQNF	8	915.09	0	3.7	0	Poor
16	531	KLVIVGDGGTGKTTF	15	1492.72	0	9.74	1	Good
17	775	NVIVKVSKKQIEE	13	1513.78	0	9.43	1	Good
18	774	NVIVKISRKQIEE	13	1555.82	0	9.53	1	Good
19	993	VIVKISRKQIEE	12	1441.72	0	9.9	1	Good
20	1046	VVIGHVDSGKSTTTGHL	17	1707.88	0	7.92	0.2	Poor
21	1143	YVVVNDSCYHQL	12	1439.59	2560	4.87	-1	Poor
22	493	IVFWMYNDQDTPVIA	15	1812.05	6970	0.54	-2	Poor
23	806	QDLDIFVNSVE	11	1278.36	0	0.59	-3	Good
24	571	LDIFVNSVE	9	1035.15	0	0.71	-2	Good
25	231	FEITPEKNPQLQDLDIFVNSVE	22	2575.82	0	3.34	-4	Good
26	922	TSNIANQLDSTPRVFYLV	18	2038.26	1280	6.32	0	Poor
27	507	IVTVNEGKGDFE	12	1307.41	0	3.69	-2	Good
28	508	IVTVNEGKGDFEL	13	1420.56	0	3.69	-2	Good
29	509	IVTVNEGKGDFELVGQRNENQQE	23	2603.75	0	3.88	-3	Good
30	506	IVTVNEGKGDF	11	1178.29	0	3.93	-1	Good
31	494	IVHIAKQL	8	921.14	0	10.14	1.1	Poor
32	416	IIAVPTGIVFW	11	1215.48	5690	3.61	0	Poor
33	823	QVVNCNGNTVFDGELEAGRA	20	2093.24	0	3.69	-2.1	Good
34	1001	VLLVNEGKGNLEL	13	1397.62	0	4.15	-1	Good
35	1062	WMYNDQDTPVIAV	13	1551.72	6970	0.57	-2	Poor
36	824	REGDIIAVPTGIVF	14	1486.71	0	3.93	-1	Poor
37	289	FRKGDIIAIPSGIPYWTYNHGDEPL	25	2860.18	8250	5.23	-0.9	Poor
38	384	IDVMRVIDE	9	1089.27	0	3.54	-2	Good
39	1141	YVVFKTNDRAAVSHVQQ	17	1962.17	1280	9.49	1.1	Poor
40	1140	YVVFKTNDRAAVSHV	15	1705.91	1280	9.49	1.1	Poor
41	1142	YVVFKTNDRAAVSHVQQV	18	2061.3	1280	9.49	1.1	Poor
42	886	SVIIPATTGEMGVLPGHVAT	20	1950.26	0	5.1	-0.9	Poor
43	295	FTTITVNEPAPGVKAILS	18	1858.14	0	6.61	0	Poor
44	344	GLFGGAGVGKTVLIMELIN	19	1889.27	0	6.85	0	Poor

It has been found that glycine (G) and proline (P) could control protein self-assembly to amyloid fibrils (Rauscher, Baud, Miao, Keeley, & Pomès, 2006). Histidine (H) has been shown by MD simulations to regulate the pH-dependent fibrillation of M1 peptides (Hiew *et al*., 2016). Moreover, various hydrophobic residues, including tryptophan (W), tyrosine (Y), and phenylalanine (F), could contribute to the adhesion of peptides on hydrophobic surfaces (Yu *et al*., 2013).

It has been proposed that peptides with broad antimicrobial activities (anti-bacterial, antiviral and antifungal) usually have a common amino acid residue Gly or Pro at position 1 or 2 of the N-terminal; peptides with amino acid residue Gly at position 1 of N-terminal showed no anticancer activity, while all the peptides with anti-inflammatory activity have amino acid residue Ile, Gly, Ala, Ser, or Val at position 3 of the N-terminal, and the additional presence of the same amino acid residues at position 2 or 4 improved the anti-inflammatory specificity and sensitivity (Fatoki *et al*., 2022). A study has shown that TNKPVI peptide, a product of *in silico* simulated digestion of potato patatin using pepsin and trypsin, greatly lowered the level of proinflammatory IL-6, IL-8, and TNF in lipopolysaccharide-activated human monocytic (THP-1) cells (Okeke, Abioye, Ventura-Santana, Sun, & Udenigwe, 2022).Neuropeptides play vital roles in neurodegenerative diseases (Ben-Shushan & Miller, 2021), antimicrobial peptides contribute to host defense (Mookherjee, Anderson, Haagsman, & Davidson, 2020), and immunomodulatory peptides support vaccine design (Pavlicevic, Marmiroli, & Maestri, 2022). It has been suggested that the transient oligomers can be more toxic than the mature fibrils (Ke *et al*., 2020).

3.4 3D structures' predicted energy and molecular dynamics profiles of the selected lentil self-aggregating peptides

The results of 3D structure prediction for the 44 peptides are presented in Table 4, which showed that peptide 1,137 (YVIVNDSCYHQLVSHWLNT) has the highest oSPEP energy while peptide 806 (QDLDIFVNSVE) has the lowest oSPEP energy followed by peptide 711 (LVVVPQNF), and selected structures of 27 peptides have low sOPEP energy (≤ -15.0) as presented in Figure 1. In the context of structure modeling, a high sOPEP energy typically indicates a less stable or less favorable conformation of a structure. Lower sOPEP energies generally correspond to structures that are more energetically favorable and closer to native-like states, which are desirable in structural modeling and molecular dynamics studies. In this study, the structure of peptide 1,137 (YVIVNDSCYHQLVSHWLNT) with sOPEP score of -33.4285 (lowest) is the most stable compared to the structure of peptide 806 (QDLDIFVNSVE) with sOPEP score of -4.82228 (highest). Therefore, high sOPEP energy is not ideal for structure modeling, as it suggests higher potential energy and reduced stability or biological plausibility of the structure.

The molecular dynamic (MD) simulation analyses indicated that peptides 1,142, 289, 922, 509 and 499 have potential energy below -100 kJ.mol^{-1} (Table 4). MD simulation studies can provide an efficient tool to characterize the aggregation and conformational changes of biomolecules in water at a molecular level (Fatoki, 2023). MD studies are commonly used to investigate such processes as drug targeting and docking, neuronal signaling, protein folding, conformational dynamics of biomacromolecules, and proteins' and peptides' self-assembly (Hospital, Goñi, Orozco, & Gelpí, 2015; Hollingsworth & Dror, 2018). MD simulation has been used to study the effect of a finite system size on the aggregation of amyloidogenic peptides of the islet amyloid polypeptide IAPP, and it was found that peptide aggregation, as well as various properties of the peptide-water system, are highly sensitive to the system size and concentration, as expected for any phase transition (Singh, Brovchenko, Oleinikova, & Winter, 2008).

Furthermore, all-atom molecular dynamics (AA-MD) simulations has been used to study a series of structurally similar therapeutic peptides, which provided molecular-level insights into aggregation dynamics, aggregation pathways, and the influence of different structural elements on peptides' aggregation propensity and on intermolecular interactions within the aggregates (Hjalte *et al*., 2022).

4. Conclusions

In this study, we computationally explored the self-assembly properties of lentil hydrolysate peptide derived from LPTA interaction and the peptides with proneness to

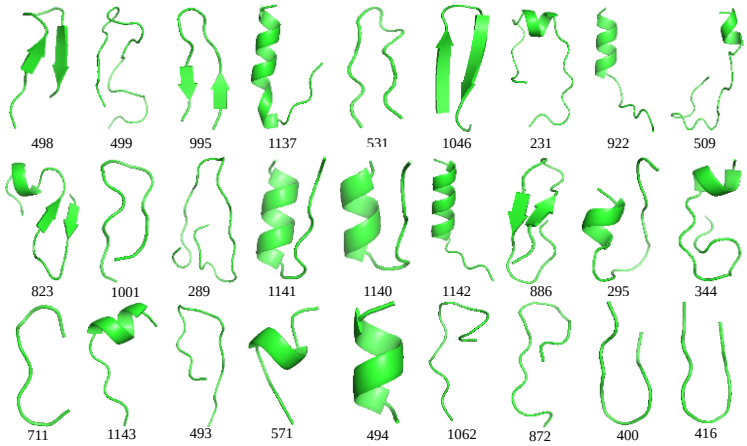

Figure 1. 3D structures of the selected 27 peptides

Table 4. 3D structures' predicted energy and molecular dynamics profile at of the selected lentil self-aggregating peptides

SN	Peptide custom database ID	Sequence	3D structure	Molecular dynamic simulation profile			
			sOPEP energy	Potential energy ($kJ.mol^{-1}$)	Total energy ($kJ.mol^{-1}$)	Total Rg (nm)	RMSD (nm)
1	497	IVIVTVNEGKGDFE	-8.46027	-79	-65	0.80	0.40
2	498	IVIVTVNEGKGDFEL	-17.9051	-94	-77	0.83	0.35
3	499	IVIVTVNEGKGDFELVGQRNENQQE	-15.3421	-103	-84	1.10	0.50
4	496	IVIVTVNEGKGDF	-8.09250	-89	-73	0.80	0.40
5	495	IVIVTVNEGKGD	-6.40239	-84	-78	0.75	0.30
6	994	VIVTVNEGKGDF	-9.58308	-94	-76	0.75	0.35
7	995	VIVTVNEGKGDFEL	-22.9596	-74	-61	0.76	0.30
8	872	SPGDVVIIPAGHPVA	-11.4229	-78	-63	0.85	0.50
9	455	ILVVLSGKAIL	-13.6015	-68	-55	0.65	0.30
10	1137	YVIVNDSCYHQLVSHWLNT	-33.4285	-70	-57	0.95	0.42
11	500	IVIWTASAL	-8.64756	-49	-39	0.75	0.40
12	417	IIIWTASAL	-9.60032	-49	-39	0.80	0.35
13	400	IFIIQGKGVLG	-10.1915	-81	-65	0.65	0.25
14	399	IFIIQGKGVL	-8.52894	-65	-53	0.80	0.55
15	711	LVVVPQNF	-4.86906	-59	-40	0.65	0.40
16	531	KLVIVGDGGTGKTTF	-20.8714	-80	-65	0.80	0.35
17	775	NVIVKVSKKQIEE	-7.43332	-115	-93	0.95	0.30
18	774	NVIVKISRKQIEE	-8.79041	-109	-89	1.00	0.50
19	993	VIVKISRKQIEE	-8.76144	-83	-67	0.90	0.35
20	1046	VVIGHVDSGKSTTTGHL	-16.2685	-86	-70	0.95	0.25
21	1143	YVVVNDSCYHQL	-11.8063	-87	-71	0.85	0.40
22	493	IVFWMYNDQDTPVIA	-10.5341	-89	-73	0.80	0.40
23	806	QDLDIFVNSVE	-4.82228	-87	-71	0.85	0.50
24	571	LDIFVNSVE	-7.23960	-49	-41	0.65	0.25
25	231	FEITPEKNPQLQDLDIFVNSVE	-18.5492	-95	-78	1.10	0.60
26	922	TSNIANQLDSTPRVFYLV	-19.0342	-140	-115	1.20	0.40
27	507	IVTVNEGKGDFE	-6.13114	-112	-92	0.75	0.30
28	508	IVTVNEGKGDFEL	-11.4718	-55	-45	0.75	0.35
29	509	IVTVNEGKGDFELVGQRNENQQE	-15.8872	-123	-104	1.10	0.50
30	506	IVTVNEGKGDF	-6.35078	-48	-39	0.65	0.25
31	494	IVHIAKQL	-10.9227	-34	-28	0.65	0.45
32	416	IIAVPTGIVFW	-14.6770	-78	-63	0.75	0.35
33	823	QVVNCNGNTVFDGELEAGRA	-23.8820	-70	-56	0.80	0.40
34	1001	VLLVNEGKGNLEL	-18.4874	-53	-43	0.70	0.25
35	1062	WMYNDQDTPVIAV	-6.66513	-82	-71	0.90	0.50
36	824	REGDIIAVPTGIVF	-7.21185	-92	-73	0.95	0.40
37	289	FRKGDIIAIPSGIPYWTYNHGDEPL	-22.0269	-114	-92	0.93	0.40
38	384	IDVMRVIDE	-13.8836	-43	-35	0.60	0.20
39	1141	YVVFKTNDRAAVSHVQQ	-20.9055	-92	-74	0.90	0.50
40	1140	YVVFKTNDRAAVSHV	-16.7311	-92	-75	0.85	0.45
41	1142	YVVFKTNDRAAVSHVQQV	-25.0906	-147	-120	1.00	0.35
42	886	SVIIPATTGEMGVLPGHVAT	-25.7989	-68	-54	0.80	0.20
43	295	FTTITVNEPAPGVKAILS	-18.3091	-73	-58	0.85	0.35
44	344	GLFGGAGVGKTVLIMELIN	-24.1323	-58	-47	0.80	0.60

formation of amyloid were obtained, of which only seven were predicted to be non-toxic and consecutively poorly soluble (peptide 494, 494, 571, 711, 1,062, 1,137, and 1,143). They possess variably suitable bioactivities, which include antihypertensive, antidiabetic, antimicrobial, antimalarial, quorum sensing, anticancer, and antioxidant activities. Peptide 711 (LVVVPQNF) showed the highest non-toxic bioactivity profile but high sOPEP energy which indicates an unstable structure. Overall, this study has identified peptide 1,137 (YVIVNDSCYHQLVSHWLNT) as the best peptide based on its moderate self-assembly profile, moderate non-toxic bioactivity, physicochemical properties, the lowest 3D structure sOPEP energy, and moderately stable structure in simulated physiological conditions. Moreover, the structural optimization to eliminate potential toxicity will be beneficial for the peptides 995, 531, 1,046, 231, 823, 289, 1,141, 1,140, 886, 295 and 344, as they possess other useful bioactivities and structural stability.

4.1 Strength, limitations and future perspectives

This study provides insights into the potential biomaterial applications of lentil hydrolysate peptides, particularly their self-assembly properties and bioactivities. A key strength is the identification of peptide 1,137, which balances moderate self-assembly, bioactivity, and structural

stability, making it a promising candidate for applications like scaffolds or drug delivery systems. The computational approach efficiently narrows down candidates for further experimental exploration, saving time and resources. Poor solubility, often seen as a limitation, provides benefits in biomaterial applications. It allows for controlled and sustained release of bioactive compounds, enhancing therapeutic outcomes. Additionally, it supports the development of localized treatments, such as coatings for implants or tissue-specific drug carriers, where gradual biodegradation is advantageous. However, several limitations must be addressed. The computational predictions, while valuable, require experimental confirmation to validate self-assembly behaviors and bioactivity profiles under physiological conditions. Future work should include experimental studies to confirm peptide assembly and functional properties, along with optimization of solubility and stability through chemical modifications or nanocarrier integration. Developing biomaterials that leverage peptides like 1,137, with balanced stability and bioactivity, could lead to innovative applications in drug delivery, tissue engineering, and antimicrobial surfaces. Simulations integrating material-specific conditions could refine these applications further.

References

Akbarian, M., Khani, A., Eghbalpour, S., & Uversky, V. N. (2022). Bioactive Peptides: Synthesis, Sources, Applications, and Proposed Mechanisms of Action. *International Journal of Molecular Sciences*, *23*(3), 1445.

Bayarri, G., Andrio, P., Hospital, A., Orozco, M., & Gelpí, J. L. (2022). BioExcel Building Blocks Workflows (BioBB-Wfs), an integrated web-based platform for biomolecular simulations. *Nucleic Acids Research*, *50*(W1), W99–W107.

Bellesia, G., & Shea, J. E. (2009). Effect of beta-sheet propensity on peptide aggregation. *The Journal of Chemical Physics*, *130*(14), 145103.

Ben-Shushan, S., & Miller, Y. (2021). Neuropeptides: Roles and activities as metal chelators in neurodegenerative diseases. *Journal of Physical Chemistry B*, 125, 2796–2811.

Boachie, R. T., Okagu, O. D., Abioye, R., Hüttmann, N., Oliviero, T., Capuano, E., . . . Udenigwe, C. C. (2022). Lentil protein and tannic acid interaction limits in vitro peptic hydrolysis and alters peptidomic profiles of the proteins. *Journal of Agricultural and Food Chemistry*, *70*(21), 6519–6529.

Cabri, W., Cantelmi, P., Corbisiero, D., Fantoni, T., Ferrazzano, L., Martelli, G., . . . Tolomelli, A. (2021). Therapeutic peptides targeting PPI in clinical development: Overview, mechanism of action and perspectives. *Frontiers in Molecular Biosciences*, *8*, 697586.

Das, U., Hariprasad, G., Ethayathulla, A. S., Manral, P., Das, T. K., Pasha, S., . . . Srinivasan, A. (2007). Inhibition of protein aggregation: supramolecular assemblies of arginine hold the key. *PloS One*, *2*(11), e1176.

Dobson C. M. (2003). Protein folding and misfolding. *Nature*, *426*(6968), 884–890.

Du, Z., Ding, X., Xu, Y., & Li, Y. (2023). UniDL4BioPep: A universal deep learning architecture for binary classification in peptide bioactivity. *Briefings in bioinformatics*, *24*(3), bbad135.

Fatoki T. H. (2023). Effect of pH on structural dynamics of HMG-CoA reductase and binding affinity to β-sitosterol. *Journal of Biomolecular Structure and Dynamics*, *41*(10), 4398–4404.

Fatoki T. H., Chukwuejim, S., Udenigwe, C. C., & Aluko, R. E. (2023). *In silico* exploration of metabolically active peptides as potential therapeutic agents against amyotrophic lateral sclerosis. *International Journal of Molecular Sciences*, *24*(6), 5828.

Fatoki, T. H., Aluko, R. E., & Udenigwe, C. C. (2022). *In silico* investigation of molecular targets, pharmaco kinetics, and biological activities of chicken egg ovalbumin protein hydrolysates. *Journal of Food Bioactives*, *17*, 34-48.

Hiew, S. H., Guerette, P. A., Zvarec, O. J., Phillips, M., Zhou, F., Su, H., . . . Miserez, A. (2016). Modular peptides from the thermoplastic squid sucker ring teeth form amyloid-like cross-β supramolecular networks. *Acta Biomaterialia*, *46*, 41–54.

Hjalte, J., Hossain, S., Hugerth, A., Sjögren, H., Wahlgren, M., Larsson, P., & Lundberg, D. (2022). Aggregation behavior of structurally similar therapeutic peptides investigated by ^{1}H NMR and all-atom molecular dynamics simulations. *Molecular Pharmaceutics*, *19*(3), 904–917.

Høgstedt, U. B., Østergaard, J., Weiss, T., Sjögren, H., & van de Weert, M. (2018). Manipulating aggregation behavior of the uncharged peptide carbetocin. *Journal of Pharmaceutical Sciences*, *107*(3), 838–847.

Hollingsworth, S. A., & Dror, R. O. (2018). Molecular dynamics simulation for all. *Neuron*, *99*, 1129–1143.

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., & Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*, *65*(3), 712–725.

Hospital, A., Goñi, J. R., Orozco, M., & Gelpí, J. L. (2015). Molecular dynamics simulations: advances and applications. *Advances and Applications in Bioinformatics and Chemistry*: *AABC*, *8*, 37–47.

Huang, G., Li, F., Zhao, X., Ma, Y., Li, Y., Lin, M., . . . Xu, F. (2017). Functional and Biomimetic Materials for Engineering of the Three-Dimensional Cell Microenvironment. *Chemical Reviews*, *117*(20), 12764–12850.

Josefsson, L., Ye, X., Brett, C. J., Meijer, J., Olsson, C., Sjögren, A., . . . Lendel, C., (2020). Potato protein nanofibrils produced from a starch industry sidestream. *ACS Sustainable Chemistry and Engineering*, *8*, 1058−1067.

Ke, P. C., Zhou, R., Serpell, L. C., Riek, R., Knowles, T. P. J., Lashuel, H. A., . . . Mezzenga, R., (2020). Half a century of amyloids: past, present and future. *Chemical Society Reviews*, *49*(15), 5473–5509.

Langyan, S., Khan, F. N., Yadava, P., Alhazmi, A., Mahmoud, S. F., Saleh, D. I., . . . Kumar, A.

(2021). *In silico* proteolysis and analysis of bioactive peptides from sequences of fatty acid desaturase 3 (FAD3) of flaxseed protein. *Saudi Journal of Biological Sciences*, *28*(10), 5480–5489.

Levin, A., Hakala, T. A., Schnaider, L., Bernardes, G. J. L., Gazit, E., & Knowles, T. P. J. (2020). Biomimetic peptide self-assembly for functional materials. *Nature Reviews Chemistry*, *4*, 615–634.

Maqbool, A., Bakhsh, A. & Aksoy, E. (2021). *Effects of natural variations on biofortification in wild germplasm for genetic improvement in crop plants*. Cambridge, MA: Academic Press.

Maupetit, J., Tuffery, P., & Derreumaux, P. (2007). A coarse-grained protein force field for folding and structure prediction. *Proteins*, *69*(2), 394–408.

Miserez, A., Yu, J., & Mohammadi, P. (2023). Protein-based biological materials: Molecular design and artificial production. *Chemical Reviews*, *123*(5), 2049-2111.

Mookherjee, N., Anderson, M. A., Haagsman, H. P., & Davidson, D. J. (2020). Antimicrobial host defence peptides: functions and clinical potential. *Nature Reviews Drug Discovery*, *19*(5), 311–332.

Okeke, E. B., Abioye, R. O., Ventura-Santana, E., Sun, X., & Udenigwe, C. C. (2022). TNKPVI, a putative bioaccessible pharmacophore of anti-inflammatory potato patatin-derived decapeptide DIKTNKPVIF. *Molecules* (*Basel, Switzerland*), *27*(12), 3869.

Pavlicevic, M., Marmiroli, N., & Maestri, E. (2022). Immunomodulatory peptides - a promising source for novel functional food production and drug discovery. *Peptides*, *148*, 170696.

Payne, R. W., & Manning M.C. (2009). Peptide formulation: Challenges and strategies. *Innovations in Pharmaceutical Technology*, *28*, 64–68.

Peredo-Lovillo, A., Hernández-Mendoza, A., Vallejo-Cordoba, B., & Romero-Luna, H. E. (2021). Conventional and *in silico* approaches to select promising food-derived bioactive peptides: A review. *Food Chemistry*: *X*, *13*, 100183.

Prabakaran, R., Rawat, P., Kumar, S., & Gromiha, M. M. (2021). Evaluation of in silico tools for the prediction of protein and peptide aggregation on diverse datasets. *Briefings in Bioinformatics*, *22*(6), bbab240. Retreived from https://doi.org/10.1093/bib/bbab240

Rauscher, S., Baud, S., Miao, M., Keeley, F. W., & Pomès, R. (2006). Proline and glycine control protein self-organization into elastomeric or amyloid fibrils. *Structure*, *14*, 1667−1676.

Rey, J., Murail, S., de Vries, S., Derreumaux, P., & Tuffery, P. (2023). PEP-FOLD4: a pH-dependent force field for peptide structure prediction in aqueous solution. *Nucleic Acids Research*, *51*(W1), W432–W437.

Rezvankhah, A., Yarmand, M.S., Ghanbarzadeh, B., & Mirzaee, H. (2023). Development of lentil peptides with potent antioxidant, antihypertensive, and antidiabetic activities along with umami taste. *Food Science and Nutrition*, *11*(6):2974-2989.

Samaranayaka, A. (2017). *Lentil*: *Revival of poor man*'*s meat in sustainable protein sources*. Cambridge, MA: Academic Press.

Sehgal, A., Sita, K., Rehman, A., Farooq, M., Kumar, S., Yadav, R., . . . Siddique, M. (2021). Lentil in crop physiology case histories for major crops. Cambridge, MA: Academic Press.

Sharma, S., Sarkar, S., Paul, S. S., Roy, S., & Chattopadhyay, K. (2013). A small molecule chemical chaperone optimizes its unfolded state contraction and denaturant-like properties. *Scientific Reports*, *3*, 3525.

Shevkani, K., Singh, N., Chen, Y., Kaur, A., & Yu, L. (2019). Pulse proteins: Secondary structure, functionality and applications. *Journal of Food Science and Technology*, *56*(6), 2787−2798.

Shire, S. J., Shahrokh, Z., & Liu, J. (2004). Challenges in the development of high protein concentration formulations. *Journal of Pharmaceutical Sciences*, *93*(6), 1390–1402.

Singh, G., Brovchenko, I., Oleinikova, A., & Winter R. (2008). Peptide aggregation in finite systems. *Biophysical Journal*, *95*, 3208–3221

Thangakani, A. M., Kumar, S., Nagarajan, R., Velmurugan, D., & Gromiha, M. M. (2014). GAP: towards almost 100 percent prediction for β-strand-mediated aggregating peptides with distinct morphologies. *Bioinformatics*, *30*(14), 1983–1990.

Thavarajah, D., Ruszkowski, J., & Vandenberg, A. (2008). High potential for selenium biofortification of lentils (*Lens culinaris* L.). *Journal of Agricultural and Food Chemistry*, *56*(22), 10747–10753.

Tuffery, P., & Derreumaux, P. (2023). A refined pH-dependent coarse-grained model for peptide structure prediction in aqueous solution. *Frontiers in Bioinformatics*, *3*, 1113928.

Udenigwe, C. C., Abioye, R. O., Okagu, I. U., & Obeme-Nmom, J. I. (2021). Bioaccessibility of bioactive peptides: recent advances and perspectives. *Current Opinion in Food Science*, *39*, 182–189.

Ullah, S., & Chen, X. (2020). Fabrication, applications and challenges of natural biomaterials in tissue engineering. *Applied Materials Today*, *20*(100656): 100656.

Walsh, I., Seno, F., Tosatto, S. C., & Trovato, A. (2014). PASTA 2.0: an improved server for protein aggregation prediction. *Nucleic Acids Research*, *42*(Web Server issue), W301–W307.

Ye, X., Lendel, C., Langton, M., Olsson, R. T., & Hedenqvist, M. S. (2019). Protein nanofibrils: Preparation, properties, and possible applications in industrial nanomaterials. In Industrial Applications of Nanomaterials. Amsterdam, The Netherlands: Elsevier.

Yu, J., Kan, Y., Rapp, M., Danner, E., Wei, W., Das, S., . . . Israelachvili, J. N. (2013). Adaptive hydrophobic and hydrophilic interactions of mussel foot proteins with organic thin films. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(39), 15680–15685.

Zeece, M. (2020). Food systems and future directions in introduction to the chemistry of food. Cambridge, MA: Academic Press.