RSC Advances

PAPER

Cite this: *RSC Adv.*, 2024, **14**, 15542

Evaluating the capability of soybean peptides as calcium ion carriers: a study through sequence analysis and molecular dynamics simulations†

Jiulong An,[a] Yumei Wang,[a] Wenhui Li,[a] Wanlu Liu,[a] Xiangquan Zeng,[ab] Guoqi Liu,[b] Xinqi Liu[a] and He Li *[ab]

Received 19th April 2024
Accepted 6th May 2024

DOI: 10.1039/d4ra02916j

rsc.li/rsc-advances

Calcium homeostasis imbalance in the body can lead to a variety of chronic diseases. Supplement efficiency is essential. Peptide calcium chelate, a fourth-generation calcium supplement, offers easy absorption and minimal side effects. Its effectiveness relies on peptide's calcium binding capacity. However, research on amino acid sequences in peptides with high calcium binding capacity (HCBC) is limited, affecting the efficient identification of such peptides. This study used soybean peptides (SP), separated and purified by gel chromatography, to obtain HCBC peptide (137.45 μg mg^{-1}) and normal peptide (≤95.78 μg mg^{-1}). Mass spectrometry identified the sequences of these peptides, and an analysis of the positional distribution of characteristic amino acids followed. Two HCBC peptides with sequences GGDLVS (271.55 μg mg^{-1}) and YEGVIL (272.54 μg mg^{-1}) were discovered. Molecular dynamics showed that when either aspartic acid is located near the N-terminal's middle, or glutamic acid is near the end, or in cases of continuous Asp or Glu, the binding speed, probability, and strength between the peptide and calcium ions are superior compared to those at other locations. The study's goal was to clarify how the positions of characteristic amino acids in peptides affect calcium binding, aiding in developing peptide calcium chelates as a novel calcium supplement.

1 Introduction

Calcium is the most abundant mineral element in the human body. Calcium homeostatic imbalances can lead to a variety of diseases such as rickets, osteomalacia and osteoporosis.[1] In addition, calcium is closely linked to the activity of many metabolic enzymes in humans as well as to the nervous system. Oral administration is one of the most effective ways of supplementing calcium, so the efficient development of calcium supplements is essential for the effective prevention of chronic diseases of the skeleton.[2]

Current research indicates that peptides can spontaneously bind with calcium ions to form peptide calcium chelates.[3] In recent years, these chelates, as the fourth category of calcium supplements, have increasingly become a research focus. As a novel type of calcium supplement, they offer advantages over traditional calcium supplement products, including fewer side effects, higher absorption efficiency, and better stability.[3] Obtaining high calcium binding capacity (HCBC) peptides is essential for peptide calcium chelate preparation, and many studies have shown that HCBC peptides can be derived from various food proteins. For example, casein phosphopeptides from milk were initially found to bind Ca^{2+} ions. And then, HCBC peptides have been extracted from plant and animal protein such as wheat, black beans, egg yolk proteins and octopus processing by-products.[4–7]

Although the sources of HCBC peptides are diverse. However, determining the sequence of a HCBC peptide requires the use of two or three different types of chromatographic purification techniques. Subsequently, sequence identification was performed using mass spectrometry. This reduces the efficiency of peptide acquisition.[8] The structure–activity relationship of peptides is currently a hot topic in the field of bioactive peptides. However, the structure–activity relationship of calcium binding ability, especially the influence of characteristic amino acid positions, still needs further study.[8]

Molecular dynamics, is a technique that uses computers to simulate interactions between molecules on a certain time scale. It has been shown to be a good judge of peptide and other molecule interactions.[9] Utilizing molecular dynamics simulations, the binding principles and mechanisms between peptides and small molecules can be predicted and elucidated.[10] There are already some studies that have used molecular dynamics to identify Ca^{2+} binding sites on peptides. For

[a]Key Laboratory of Geriatric Nutrition and Health (Beijing Technology and Business University), Ministry of Education, Beijing 100048, China. E-mail: lihe@btbu.edu.cn
[b]Key Laboratory of Green and Low-carbon Processing Technology for Plant-based Food of China National Light Industry Council, Beijing Technology and Business University, Beijing 100048, China. E-mail: 20210803@btbu.edu.cn; liuguoqi-006@163.com

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4ra02916j

example, the study of Xu *et al.*, based on molecular dynamics findings, discovered a chelating interaction between Ca^{2+} and the aspartic acid (Asp, D) of a mussel peptide (LGKDQVRT), confirming the interaction between carboxyl groups and calcium ions.[11] Additionally, Gan *et al.* employed molecular dynamics to analyze the conformational stability differences between the peptide CBP (DEDEQIPSHPPR) and its chelate, CBP-Ca, and found that the chelate exhibited a more stable conformation.[12] However, studies using molecular dynamics to determine the contributions of different characteristic amino acids to calcium binding ability are still relatively scarce.[13]

Soybean peptide (SP) is obtained by enzymatic digestion of soy protein, which has the advantages of a comprehensive amino acid, easy absorption, and has been shown to have HCBC.[14] The use of SP as a carrier to study the structure–active relationship can exclude the influence of other protein source carriers that do not contain a complete range of amino acids or unbalanced content.[14]

This study aimed to use SP prepared by enzymatic digestion. SP with HCBC were separated and purified by gel chromatography. After that, mass spectrometry sequencing and positional quantification methods were used to explore the types and positional distribution frequencies of calcium binding characteristic amino acids in the peptides, in order to identify the key peptides and the sequence characteristics of the HCBC peptides. The mechanism was verified by molecular dynamics. We expect that this study will provide a new theoretical basis for the study of the structure–activity relationship of the calcium binding capacity of peptides and the development of food-derived peptide calcium chelates.

2 Materials and methods

2.1 Materials

Soybean protein (92% protein) was purchased from Yuxin Biotechnology Co. (Shandong, China). Alkaline protease and papain were provided by Novozymes (Beijing, China). Molecular weight standards (glycine–glycine–glycine, bacillus peptide, thymic peptide, insulin) were purchased from Yuanye Company (Shanghai, China). Calcium chloride solution (1 mol L^{-1}) was bought from Solarbio(Beijing, China). Formic acid (chromatographic grade) was sourced from Macklin Company (Shanghai, China). Acetonitrile (chromatographic grade) was purchased from ThermoFisher Company (Massachusetts America). Sequence-specific peptides were synthesized from HongTide Biotechnology Co., Ltd (Shanghai, China).

2.2 Preparation of soybean peptide (SP)

Soybean protein was dissolved in deionized water to achieve a 10% protein solution. This solution underwent hydrolysis with a 1% solution of alkaline protease and papain (1 : 1, mass ratio), mixed, at 50 °C for four hours. To stop the hydrolysis, the mixture was heated to 85 °C for 15 minutes, deactivating the enzymes.[15] Subsequently, the peptides were isolated using ultrafiltration membranes with a cutoff of 3000 Da. The filtrate was then spray-dried at 170 °C, resulting in fine, light yellow particles of soybean peptides, ready for future experiments.

2.3 Molecular weight measurement

The method of Li *et al.* was referenced, with slight modifications.[16] The molecular weight distribution of SP was analysed using a 1260 Infinity II LC system and a TSK-GEL G2000SWXL (5 μm, 7.8 × 300 mm^2) column. The concentration of SP and four standards (glycine–glycine–glycine, bacillus peptide, thymus peptide, and tryptophan) were all at a concentration of 1 mg mL^{-1}, and were filtered through a 0.22 μm membrane prior to sampling. Samples were filtered through a 0.22 μm membrane before sampling. The sample volume was 10 μL, and the mobile phase of water/acetonitrile/trifluoroacetic acid (45 : 55 : 1) was used for elution. The flow rate was set at 0.5 mL min^{-1} and the detection wavelength was 220 nm.

2.4 Purification of SP

The soybean peptides were configured as a 1 mg mL^{-1} solution and purified using an Akta Pure protein purifier (Cytiva Biotechnology, USA) loaded with a Superdex 30 Increase 10/300 GL gel column.[17] The sample volume was 500 μL per sample, and isocratic elution was performed using ultrapure water at a flow rate of 0.5 mL min^{-1}. The eluate was detected by UV absorbance at 220 nm.[17] The collector was set to collect 1 mL of eluate per tube. A total of 1.5 times the column volume of eluate was collected. SP was divided into five fractions (labeled F1–F4), which were freeze-dried and prepared for use.

2.5 Determination of calcium binding capacity of peptides

Huang *et al.*'s method was adapted with minor modifications.[18] 1 mL of peptide solution (1 mg mL^{-1}) was mixed with 2 mL of $CaCl_2$ solution (5 mmol L^{-1}) and reacted at 37 °C for 30 minutes. This was followed by the addition of sodium phosphate buffer (20 mmol L^{-1}, pH 7.8) and the reaction continued for another 30 minutes. Post-reaction, the mixture was centrifuged to precipitate free calcium (5000 rpm, 10 min), and the supernatant containing bound calcium was collected. The bound calcium content in the supernatant was then quantified using colorimetric methods.[18]

2.6 Identification of peptide sequences by LC-MS/MS

The different soy peptide samples, purified by gel chromatography and dissolved after ultrafiltration, were loaded onto an EASY-nLC 1200 nanoscale high-performance liquid chromatography. The chromatographic column used was the Acclaim PepMap™ RSLC (50 μm × 150 mm). The mobile phase A was 0.1% formic acid aqueous solution, and the mobile phase B was an 80% acetonitrile solution with 0.1% formic acid.

Following separation by high-performance liquid chromatography, the samples were subjected to mass spectrometric analysis using a Thermo QE HF mass spectrometer (Thermo Fisher) with an analysis duration of 120 minutes. After each full scan, 20 fragment ion spectra were collected. The raw mass spectrometry files were analyzed using the Proteome Discoverer

2.5 software to search the corresponding database, ultimately yielding the identification of the peptides.

2.7 Analysis of soybean peptide sequences

2.7.1 Hydrophilicity scores and theoretical charge calculations. The grand average of hydropathy (grave) and the theoretical charge under physiological conditions (pH = 7.4) was calculated for each peptide, using the Biopython module under the computer Python.[19]

2.7.2 Amino acid frequency and position score statistics. Using Python to count the frequency of occurrence of each amino acid (number of occurrences of each amino acid/total number of amino acids), and positional scoring based on the amino acid sequence. Recorded the following data, the position of a particular amino acid in the peptide (di) (recorded as 0 for the first position of the nitrogen terminus, 1 for the second position, and so on). The total number of that amino acid in the peptide (An), and if the amino acid is more than one in the peptide, the respective di is recorded, followed by the total number of amino acids in the peptide (length).

Additionally, the study involved screening and counting the peptides in the sequence that contained a specific amino acid. It also recorded the number of these peptides in which the specific amino acid appeared in either a continuous (*e.g.*, AA) or discontinuous (*e.g.*, AXA) form.

2.8 Molecular dynamics

Peptide structure files were obtained by the online peptide structure constructor (**https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD4/**).[20] The reaction system was constructed using Gromacs 2022.3 software. The Charmm36 force field and TIP3P water model were selected and each peptide was placed in a cubic box.[21] The gmx genion command was used to add Ca^{2+} ions for the reaction to the box, and Cl^{-} ions was supplemented to ensure that the total charge of the reaction system was zero. The ratios of calcium ions and peptides were set as in the literature, and adequate amounts of calcium ions were ensured.[21] The total energy of the system was reduced using steepest descent minimization (STEEP). Subsequently, the system was simulated by *NVT* and *NPT* at 100 ps, respectively. The temperature of the system is controlled to be around 310 K and the pressure is controlled to be around 1 bar. Then 100 ns simulations were performed on a supercomputing platform (Beijing Parallel Technology Corporation). The simulated energy and trajectory files were analysed using the software package that comes with Gromacs and the VMD software.[22]

2.9 Statistical analysis

The results were presented as the mean ± standard deviation (SD). Statistical analysis was performed using one-way analysis of variance (ANOVA) and Tukey's posthoc test with SPSS 24 software. Statistical significance was denoted by a P value of less than 0.05.

3 Results and discussion

3.1 Molecular weight of SP

As shown in Table 1, the molecular weight of the enzymatic soybean peptides was mainly concentrated below 1000 Da (81.27%). This indicates that soybean proteins were sufficiently hydrolysed under the action of the two enzymes. Related studies have shown that small molecular weight peptides bind calcium ions more readily. In mung bean protein hydrolysate, small peptides had more chelating activity than large peptides. Lin *et al.* found that calcium-binding peptides derived from blue food proteins (marine proteins) ranged from 200–2000 Da.[23] It can be seen that the peptides prepared in the present study match the characteristics of calcium-binding peptides in molecular weight. So it can be used as a follow up study.

In this study, two proteases were selected for the preparation of soybean peptides based on the following considerations. Firstly, the combined action of the two enzymes can make the hydrolysis more adequate.[24] In addition, the selection of proteases with specific action can obtain peptides with specific sequence characteristics.[25] Alkaline proteases mainly act on the carboxyl side of amino acids with aromatic or hydrophobic properties.[26] Papain acts on basic amino acids, leucine or glycine peptide bonds.[27] It has been shown that acidic amino acids are important in the binding of peptides to Ca^{2+} ions.[28] The use of these two proteases avoids acidic amino acids as enzymatic cleavage sites and reduces their over distribution at peptide endpoints.[25]

3.2 Purification of HCBC fractions of SP

According to previous reports, the molecular weight of protein hydrolysates or peptides has an effect on their metal binding capacity.[29] Therefore, separation of soybean peptides was carried out using gel chromatography. Its elution curve is shown in Fig. 1A. Four fractions (F1–F4) were collected. After the lyophilised powder of each component was made into a solution, the amount of calcium binding was determined by competing with phosphate for the precipitation of calcium ions. The calcium binding amounts of the different fractions of soybean peptides are shown in Fig. 1B. Fraction F1 exhibited the

Table 1 Molecular weights of soybean peptides[a]

	Soybean peptide	
Molecular weight (Da)	Integrated area (%)	Composite ratio (%)
<500	54.35[a]	54.35[a]
500–1000	26.92[b]	26.92[b]
1000–1500	0.02[e]	9.54[c]
1500–2000	9.52[c]	
2000–3000	4.29[d]	9.20[c]
>3000	4.92[d]	

[a] Integrated area refers to the peak area corresponding to each 500 Da interval, and composite ratio refers to the sum of the proportions of 1–2 molecular weight intervals; letters with different corners represent significant differences, $p < 0.05$.

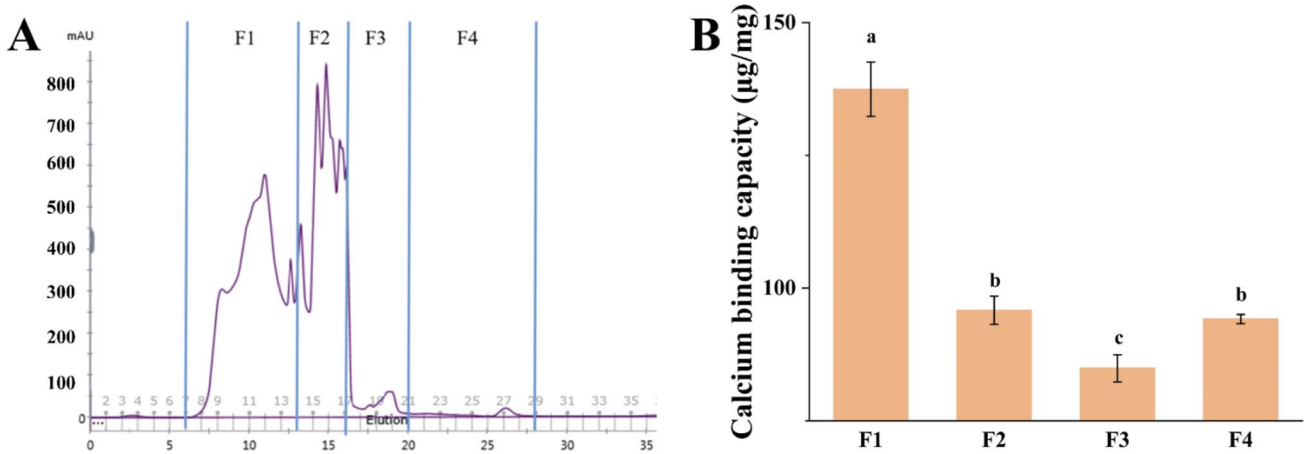

Fig. 1 Separation and purification of soybean peptides and determination of their calcium binding capacity (A) gel chromatogram of soybean peptides; (B) determination of the calcium binding capacity of different fractions of soybean peptides. *Vertical units in (B) represent the mass of calcium bound per milligram of peptide; different small-case letters represent significant differences, $p < 0.05$.

highest calcium binding amount, 137.45 μg mg^{-1}. The calcium binding of the peptide extracted from soybean-based yoghurt by Gan *et al.* was 36.64 mg g^{-1}.[30] This suggests that F1 has a high calcium binding capacity and can be followed up as a high calcium binding capacity soybean peptide.

3.3 Identification of the sequence of SP

In addition to molecular weight, amino acid sequence characteristics may also affect on calcium binding capacity. Ke *et al.* identified 26 peptides from a stickwater peptide calcium chelate (stickwater: protein solution from fishmeal processing), of which nineteen peptides contained a Glu residue, twenty peptides contained an Asp residue, and seven peptides contained the DDYVE sequence.[31] Katimba *et al.* found that the amino acid type and position had an effect on the zinc binding capacity of peptides, which was greater than that of molecular weight.[32] Therefore, it is necessary to further investigate the effect of amino acid sequence on calcium binding capacity.

3.3.1 Hydrophilicity and theoretical charge of peptides. The hydrophilicity and charge characteristics of peptides are fundamental in the interaction with other substances, and both are related to the amino acid sequence.[33] Therefore we counted Grand average of hydropathicity (Gravy) and theoretical charges of the peptide in the four fractions. The peptide sequences and related indexes are recorded in Table S1.†

Table 2 Hydrophilicity and charge characteristics of peptides in different SP fractions[a]

	Hydrophobicity (%)	Hydrophilicity (%)	Cationic (%)	Anionic (%)
F1	56.00^{c}	44.00^{a}	12.00^{a}	88.00^{a}
F2	63.67^{b}	36.33^{b}	10.50^{a}	89.50^{a}
F3	77.50^{a}	19.17^{c}	14.00^{a}	86.00^{a}
F4	62.67^{b}	37.33^{b}	15.00^{a}	85.00^{a}

[a] F1–F4 represent different fractions of SP, different corner letters in the same column represent significant differences ($p < 0.05$). Letters with different corners represent significant differences, $p < 0.05$.

In terms of hydrophilicity, Gravy, is an index to evaluate the hydrophilicity of peptides based on primary structure.[34] As shown in Table 2, the proportion of hydrophilic peptides in F1 (HCBC fraction) was the highest ($p < 0.05$). Therefore, it can be assumed that hydrophilic peptides have better calcium ion chelating ability than hydrophobic peptides. Hydrophilic amino acids have been reported to confer better metal chelating ability to peptides.[35]

The theoretical charge is an indicator of the charge of a protein or peptide calculated based on the amino acid sequence under physiological conditions (pH = 7.4). As shown in Table 2, most of the peptides in the four fractions had negative theoretical charges (>80%). This indicates that the soybean peptides are negatively charged under physiological conditions, and can still ensure the binding of positively charged calcium ions *in vivo*.[36] The stability of the peptide calcium chelate *in vivo* can be ensured.[37]

3.3.2 Calcium binding characteristic amino acid determination. Different kinds of amino acids contribute differently to calcium binding due to side chains, spatial structure, and other factors.[13] Amino acid frequencies of the four fractions SP were counted in Table 3. Aspartic acid (Asp) and glutamic acid (Glu) were higher in the F1 fraction than in the other three fractions ($p < 0.05$).

Asp and Glu are acidic amino acids. They have additional carboxyl groups and therefore carry a negative charge at physiological pH. The formation of peptide calcium chelates is dependent on the electron donors of peptide surface, and the negative charge carried by the acidic amino acids, as well as the extra carboxyl groups, can act as electron donors during chelate formation.[38] In the study by Huang *et al.*, it was found that the amount of Asp and Glu in egg white peptide calcium chelates was higher than that of normal egg white peptides before chelation.[39] And for single peptides with known sequences (*e.g.*, lemon basil peptide YDSSGGPTPWLSPY and sea cucumber

Table 3 Frequency of amino acids in the sequences of each SP fraction[a]

	F1 (%)	F2 (%)	F3 (%)	F4 (%)
A	4.48^{d}	6.38^{b}	6.29^{b}	7.48^{a}
C	0.32^{a}	0.31^{a}	0.20^{a}	0.49^{a}
D	6.46^{a}	3.98^{b}	1.95^{c}	3.45^{b}
E	12.33^{a}	6.65^{b}	1.61^{d}	3.39^{c}
F	6.56^{b}	7.69^{b}	11.17^{a}	10.22^{a}
G	6.43^{b}	9.44^{a}	8.74^{ab}	10.68^{a}
H	0.94^{a}	1.02^{a}	0.74^{a}	1.65^{a}
I	11.30^{b}	12.19^{b}	20.12^{a}	9.65^{b}
K	3.00^{a}	3.10^{a}	3.71^{a}	3.30^{a}
L	5.18^{ab}	4.56^{bc}	5.87^{a}	3.67^{c}
M	3.57^{b}	5.88^{a}	4.81^{ab}	6.30^{a}
N	2.96^{a}	2.98^{a}	2.39^{a}	2.20^{a}
P	9.52^{a}	8.38^{a}	8.02^{a}	13.34^{b}
Q	3.07^{a}	2.57^{a}	1.62^{b}	1.74^{ab}
R	2.48^{a}	1.80^{a}	2.04^{a}	2.01^{a}
S	5.45^{ab}	6.14^{a}	4.18^{b}	4.52^{ab}
T	3.95^{a}	4.20^{a}	3.21^{a}	3.21^{a}
V	9.55^{a}	9.14^{ab}	10.10^{a}	6.78^{b}
W	0.84^{b}	0.84^{b}	0.70^{b}	3.00^{a}
Y	1.59^{b}	2.76^{ab}	2.52^{ab}	2.90^{a}

[a] F1–F4 represent different fractions of soybean peptides, different corner letters in the same line represent significant differences, $p < 0.05$.

peptide QEELISK), it has also been shown that Asp and Glu in them are the main sites for Ca^{2+} ions binding.[40,41] Therefore, Asp and Glu were identified as characteristic amino acids for calcium binding capacity.

3.4 Characteristic amino acid positional tendencies

After determining the types of amino acids in the high calcium-binding capacity soybean peptides, the distribution of the characteristic amino acids in the peptides should be further determined. Because the same kind of amino acid in different positions of the peptide segment has different effects on the peptide activity.

The distribution of key amino acids in peptides with high calcium-binding capabilities is likely not random but exhibits a certain pattern. Identifying this pattern could facilitate rapid screening for peptides with high calcium-binding potential. Currently, there is limited research on the specific positioning of calcium-binding characteristic amino acids, although some studies have been conducted on zinc-binding peptides. For example, in studies of rapeseed protein peptides, asparagine residues at the N-terminus were found to have superior metal chelating ability. This explains why the peptide NCS (73.8%) has a higher chelating capacity than LAN (49%).[42] In addition, histidine located at the N-terminus of a peptide shows stronger zinc chelation than when located at the C-terminus, which could explain the higher chelation capacity of HNAPNPGLPYAA (91.7%) in contrast to NAPLPPPLKH (15.2%).[43] This observation suggests that critical positions in peptides can be categorised as N/C-termini and the remaining middle positions. Therefore, we analysed all the peptide sequences containing aspartic acid and glutamic acid in F1 and counted the positional distribution of these two amino acids in the peptides. We propose to divide peptide segments into four categories: N-terminal end (N_{end}), C-terminal end (C_{end}), N-terminal middle (N_{mid}) and C-terminal

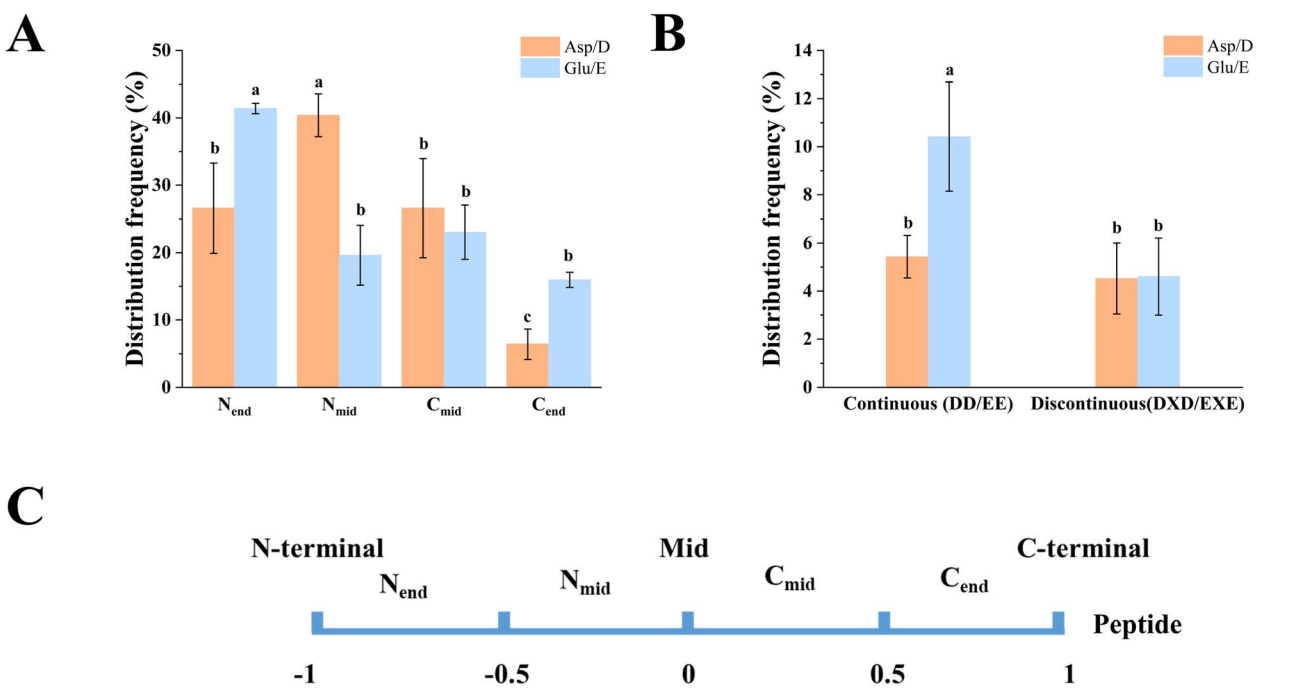

Fig. 2 Distribution of aspartic acid (Orange, Asp, D) and glutamic acid (Green, Glu, E) in High Calcium Binding Soy Peptides. (A) Distribution frequency of amino acids across four positions in peptides; (B) occurrence frequency of the two amino acids in continuous (AA) or discontinuous (AXA) sequences within peptides; (C) schematic of peptide positional categorization. *Frequency calculation: number of segments per position × 100%/total segments with specific amino acid. Different letters on bars indicate significant differences, $p < 0.05$.

middle (C_{mid}). A schematic representation of this positional categorisation is shown in Fig. 2C.

After determining the positional division, we need a reasonable indicator to determine where a particular amino acid in a peptide belongs to the position. Describing the position of an amino acid within a peptide can be inaccurate and non-intuitive due to the varying lengths of peptides. For example, the same amino acid in the third position of the nitrogen terminus is located in the middle position in a pentapeptide, whereas in an octapeptide or a longer peptide it may be more biased towards the nitrogen terminal end position. A more accurate method of positional description is therefore needed. So we designed the variable S based on the principles of normalisation and data comparability. Its calculation formula is

$$S = \frac{\sum_{i=1}^{n}\left[2 \times \left(\frac{d_i}{\text{length} - 1} - 0.5\right)\right]}{A_n}$$

If the peptide is divided into N_{end}, N_{mid}, C_{mid}, C_{end}. The S can be used to represent the position of a key amino acid in the peptide as a number between −1 and 1. The four positions correspond to the S as follows: if the number is between −1 and −0.5, we consider it to be N_{end} position; if the number is between −0.5 and 0, we consider it to be in the middle of the N_{mid} position, and so on. The position of the amino acid in the peptide can be expressed more intuitively and accurately by S. The calculation script for S is shown in ESI Code S2.†

3.4.1 Positional tendency of characteristic amino acids in peptides. As shown in Fig. 2A, Asp (D) and Glu (E) tended to be distributed at the N-terminus ($p < 0.05$). It suggests that amino acids at the N-terminus may better promote the binding of peptides to Ca^{2+} ions. Xie *et al.* found that the characteristic amino acids located at the nitrogen terminus were able to better bind Zn^{2+}.[42]

In F1, the probability of ASP at N_{mid} was significantly higher than in other positions ($p < 0.05$). In peptides containing the aspartic amino acid (Asp), aspartic acid was located in the middle of the nitrogen terminus in 40.4% of the peptides. Previously mentioned F1 is a high calcium binding capacity soybean peptide Asp located at N_{mid} may be more favourable for binding to Ca^{2+} ions. The Asp in wheat protein peptide (FVDVT) was located at N_{mid}, and the results of NMR hydrogen spectroscopy indicated that it played a key role in the calcium binding process.[44] As for Glu, the amino acid was more preferred to be distributed at the peptide N_{end} compared to other positions (41.41%, $p < 0.05$). This may be the location that favours Glu binding of calcium ions.

There are two positional relationships between two identical amino acids in a peptide segment, continuous (AA) and discontinuous (AXA). It has been suggested that these two positional relationships may have different abilities to bind metal ions.[40] As shown in Fig. 2B. For Asp (D), there was no significant difference in the frequency of occurrence of adjacent and discrete cases ($p > 0.05$). For Glu (E) the frequency of peptides in the adjacent position was significantly higher than in the separated position ($p < 0.05$). According to previous studies, phosphopeptides are peptides with good calcium binding ability, and it was found that the sequence "$S_PS_PS_PEE$" frequently appeared in peptides with strong binding ability.[45] The amino acids in this sequence are all continuous. The two characteristic amino acids in continuous positions may bind Ca^{2+} ions better.

In summary, the distribution of aspartic acid and glutamic acid in F1 does show a certain pattern, however, whether this pattern actually affects the calcium binding capacity of the peptides needs to be verified by further experiments using specific peptides. We selected two peptides (GGDLVS and YEGVIL) that matched the positional statistics results. By varying the key amino acid position design.[40] A total of 10 peptides were obtained as Group D: GGDLVS (D–N_{mid}); GGLDVS (D–C_{mid}); GGLVSD (D–C_{end}); GDDLVS (DD); GDLDVS (DXD), Group E: YEGVIL (E–N_{mid}); YGEVIL (E–N_{mid}); YGVIEL (E–C_{end}); YEEVIL(EE); YEVEIL (EXE) for subsequent validation. The Asp

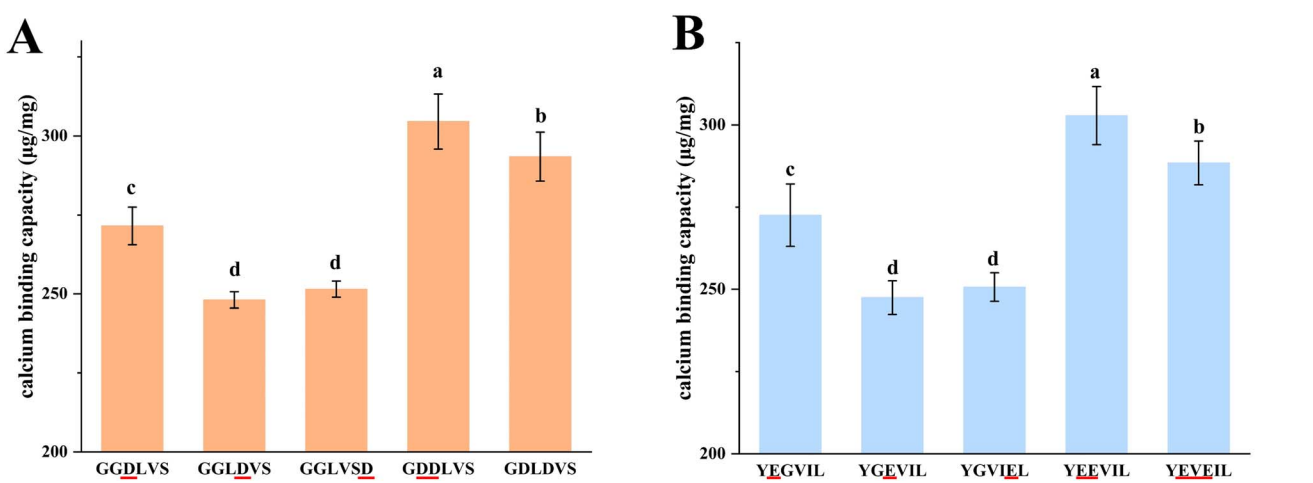

Fig. 3 Calcium binding amounts of peptides. (A) Peptides containing Asp(D). (B) peptides containing Glu(E). *Vertical coordinates represent the mass of calcium bound per milligram of peptide; different lower case letters represent significant differences, $p < 0.05$.

or Glu acid in four of these peptides, GGDLVS, YEGVIL, GDDLVS, and YEEVIL, fit the pattern we found.

3.5 Amount of calcium binding of synthetic peptides

Peptide were synthesized according to amino acid sequences, and their calcium binding amount were evaluated. As shown in Fig. 3, the calcium binding amounts of the two peptides GGDLVS and YEGVIL identified from F1 reached 271.55 μg mg^{-1} and 272.54 μg mg^{-1} respectively. According to previous studies, the peptide YDSSGGPTPWLSPY, extracted from lemon basil, demonstrated a calcium-binding amounts of 151.88 μg mg^{-1}.[41] And the calcium binding amount of the peptide VLSGGTTMAMYTLV extracted from the bone of Alaska cod was 160 μg mg^{-1}.[46] It can be seen that GGDLVS and YEGVIL have better calcium binding capacity. They have the potential to be used as raw materials for the preparation of peptide calcium chelates.

As shown in Fig. 3A, the three peptides, GGDLVS, GGLDVS, and GGLVSD, had Asp (D) located at different positions in the peptides and had different calcium binding capacities. GGDLVS had the highest amount of calcium binding ($p < 0.05$), where Asp was located at the N_{mid} position. Similarly, KGDPGLSPGK, a calcium binding peptide identified from cod fish bone, had the highest calcium binding activity among cod bone hydrolysates.[47] Based on the *S*, the Asp (D) is located at the N_{mid} position too. The Asp at the N_{mid} position enables the peptide to have a higher calcium binding amount compared to other positions.

In Fig. 3B the Glu(E) of YEGVIL is located at the N_{end} position, the Glu(E) of YGEVIL is located at the N_{mid} position and the Glu(E) of YGVIEL is located at the C_{end} position. According to the calcium binding amount results, the peptide YEGVIL possessing Glu(E) at the N_{end} position possessed the highest calcium binding amount ($p < 0.05$). Charoenphun *et al.* purified tilapia protein hydrolysates using gel chromatography and identified the tetrapeptide EPAH from the fraction with the highest calcium binding capacity.[48] It is a peptide in which the Glu(E) is in the first position at the N-terminus and belongs to the N_{end} position. For Glu(E), binding to calcium ions may be more favourable at the N_{end}.

Calcium binding was higher for GDDLVS than GDLDVS in Fig. 3A. YEEVIL had higher calcium binding than YEVEIL in Fig. 3B ($p < 0.05$). Continuous Asp and Glu are more advantageous for calcium binding. Sun *et al.* found that continuous Glu(E) in sea cucumber peptide (QEELISK) is more likely to bind Ca^{2+}.[40] The calcium binding amount of QEELISK was significantly higher than that of QELEISK and QAALISK. Si *et al.* measured the calcium binding amount of the phosphopeptide EDDSpSp (which contains a continuous Asp), and found that the calcium binding capacity of the peptide was up to 468 mg g^{-1}, which was much higher than that of peptides without continuous Asp.[49] This suggests that consecutively arranged characteristic amino acids are indeed more favourable for binding calcium ions.

In summary, Asp located in the N_{mid} position of the peptide, Glu located in the N_{end} position and continuous Asp and Glu are more advantageous for binding Ca^{2+} ions, resulting in a higher amount of calcium binding to the peptide. It also demonstrates that there is a correlation, not a coincidence, between the pattern in the distribution of calcium-binding characteristic amino acid positions and the calcium-binding capacity of the peptides. It means that peptides that conform to the pattern do have higher calcium binding capacity than peptides that do not conform to the pattern. It is reliable to use this pattern to screen peptides with high calcium binding capacity.

3.6 Molecular dynamics simulations

In order to investigate the mechanisms behind the phenomenon that peptides conforming to the distribution pattern have a higher calcium binding capacity. Molecular dynamics simulations of 100 ns were performed to analyse the reaction process of 10 peptides with calcium ions (all factors were consistent within each reaction box except for different peptides). Conformational changes during the reaction were first analysed. Subsequently, we analyse the simulation results in terms of combining speed, combining magnitude and combining strength. The speed of binding was analysed with the visual treatment, the magnitude of binding probability was verified by radial distribution function, and the strength of binding was judged using the Coul-SR.

3.6.1 Root mean square deviation. RMSD can reflect the difference between the conformation of a protein or peptide and its initial conformation at a certain point in the simulation process Fig. 4 shows the RMSD changes of 10 peptides during the reaction with calcium ions. It can be seen that the conformations of all 10 peptides changed during the reaction with calcium ions, and the green squares in the figure indicate the more significant changes in the peptide conformations. It is worth noting that the range of such changes is between 1–4 Å. In the previous study, Gan *et al.* found that peptide conformational changes of less than 6 Å during the reaction with calcium ions were reasonable and the peptide conformations were stable.[50] Therefore, we believe that the changes in the 10 peptides are due to the chelating effect of calcium ions and not due to problems with the simulation system and process parameter settings. Our simulation process is reliable. Therefore, the simulation results can be used for subsequent radial distribution function analysis and energy analysis.

3.6.2 Visualising snapshot results. To determine the binding rate of peptides to calcium ions, the simulation results after visualisation were viewed frame by frame. Record a snapshot of the keyframe when the calcium ion bound to the characteristic amino acid of each peptide for the first time in a bidentate mode and the corresponding time. Calcium binding characteristic amino acids located in a dominant position are able to bind calcium ions faster. As shown in Fig. 5, it can be seen that the binding of Asp to Ca^{2+} ions occurs earliest when Asp is in the N_{mid} position, and the first binding of the two occurs at 8.1 ns after the start of the reaction. The Glu located at the N_{end} position of the peptide occurred binding to the Ca^{2+} ion at 16.65 ns, which was able to bind Ca^{2+} ions faster than the other positions. In addition, two continuous Asp/Glu acids were able to bind calcium ions faster than two discontinuous Asp/Glu.

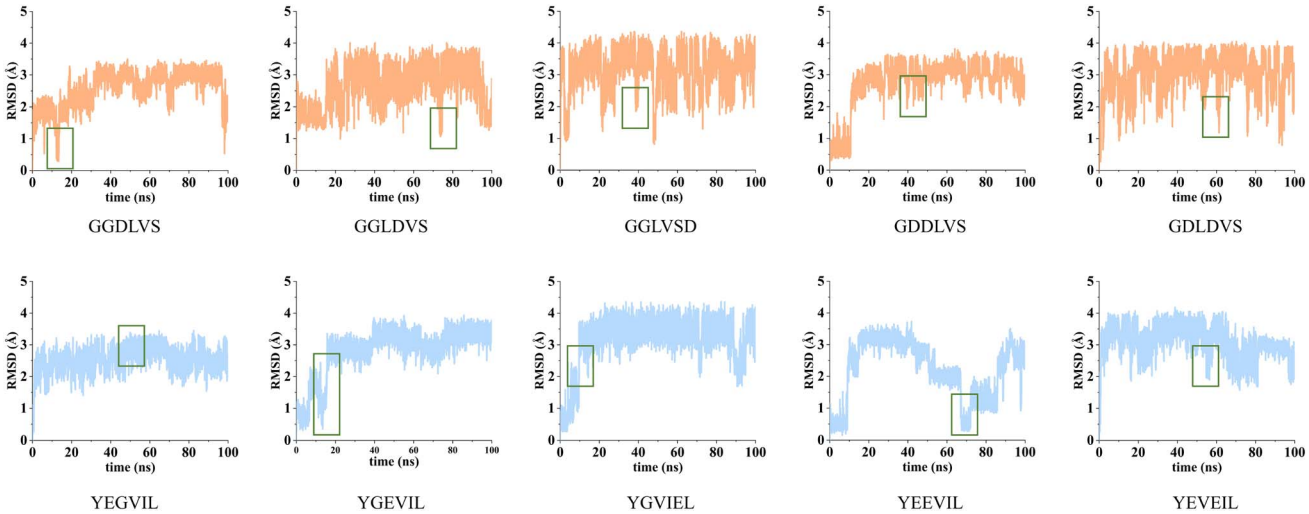

Fig. 4 RMSD *versus* time for ten peptides in molecular dynamics.

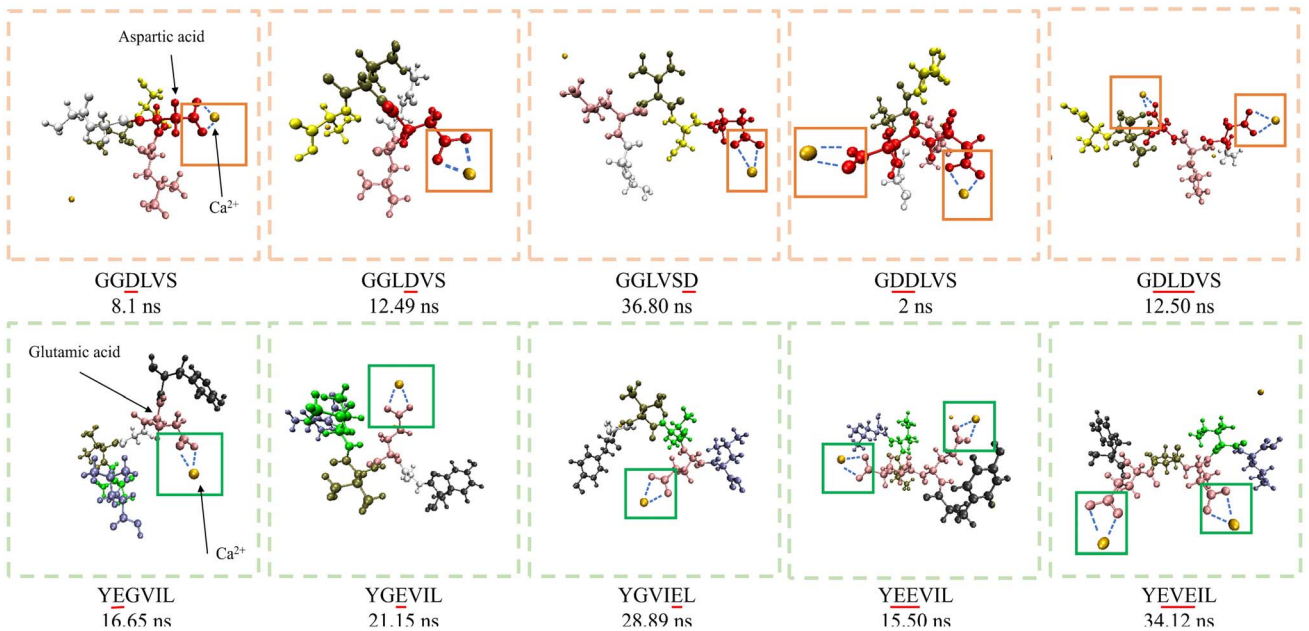

Fig. 5 Molecular dynamics snapshot of the ten peptides and their response times when they first bind to calcium ions in a bidentate mode. *The peptide sequences and times are shown below each figure; the red spheres are the atoms representing Asp acid, the pink spheres are the atoms representing Glu acid, and the gold spheres are the atoms representing calcium ions.

It can be seen that peptides characterised by high calcium binding capacity in the experiments always bind calcium ions earlier in the simulation process at the key binding sites above the peptide. This phenomenon is also found in other peptides with high calcium binding capacity. For example. Hou *et al.* found that arginine and glutamine were the major calcium ion binding sites on the krill peptide VLGYIQIR by mass spectrometry.[21] During molecular dynamics simulations of the peptide with calcium ions within 100 ns, it was found that the end position arginine had already bound calcium ions in a snapshot at 2.0 ns and that the glutamine had not yet bound.

3.6.3 Radial distribution function (RDF) results. In order to determine the probability of peptide binding to calcium ions, the metric radial distribution function (RDF) was chosen for the analysis. This is one of the common analytical indexes in molecular dynamics. RDF means the average distribution density of other particles around a certain particle within a specific distance range (usually between 1–10 Å, *i.e.*, within 1 nm).[48] The RDF provides an intuitive understanding of the spatial structure and interaction characteristics between various particles in a system. The peak of the RDF usually corresponds to the average stable distance between particles (*e.g.*, between protein and ligand).[51] The curve based on RDF

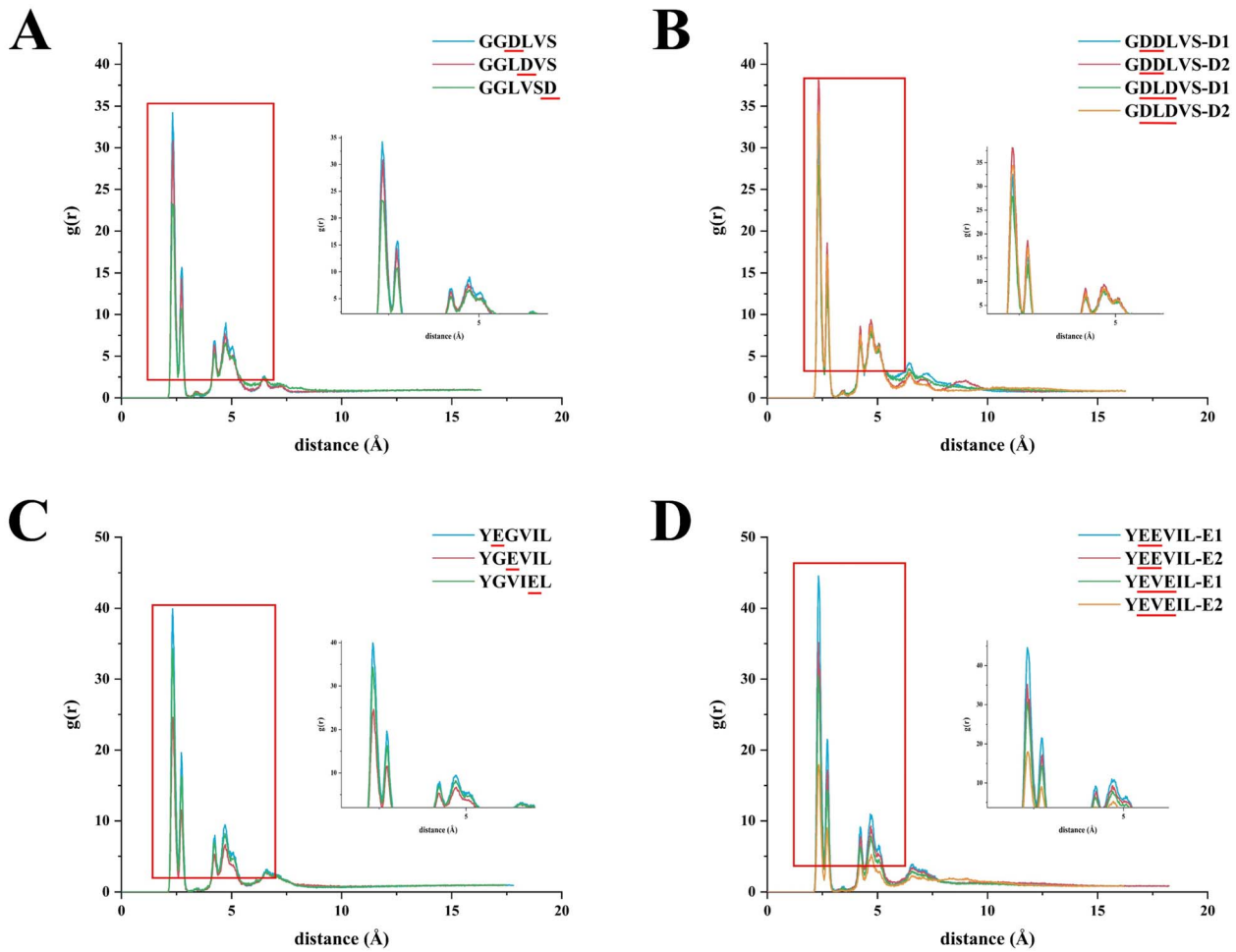

Fig. 6 RDF curves of different peptides after molecular dynamics simulation; (A) peptide with one Asp; (B) peptide with Asp; (C) peptide with one Glu; (D) peptide with two Glu. *D1/E1 represents amino acid at the front position; D2/E2 represents amino acid at the back position.

can make a judgement on the distribution of calcium ions around the peptide segment, which can be used to analyse the peptide and calcium ion binding mechanism.

As shown in Fig. 6, the RDF curves for all peptides, showed peaks at positions around 2.5 Å and 5 Å. This indicates a higher probability that calcium ions are distributed at distances of about 2.5 Å and 5 Å around the characteristic amino acids. The distance of 2.5 Å is usually recognised in molecular dynamics as the range of coordinate covalent bond interactions of metal ions, while interactions at the 5 Å position may be attributed to other forces.[52] By comparing the heights of the two peaks, it can be concluded that the ligand interaction plays a dominant role in the binding of the metal ion to the peptide.

For the three peptides containing one Asp. As shown in Fig. 6A, the distribution of the peak heights of its two peaks was D–N_{mid} > D–C_{mid} > D–C_{end}, indicating that calcium ions appeared more frequently around the Asp at the N_{mid} position. For two peptides containing two Asp, calcium ions appeared more frequently around continuous Asp acid as shown in Fig. 6B.

For the three peptides with one Glu, as shown in Fig. 6C, the Glu located at the N_{end} contains a higher RDF peak. This indicates that peptides with Glu present at the N_{end} end have a higher chance of binding calcium ions. For the two peptides containing two Glu, as shown in Fig. 6D, it is evident from the RDF peaks that the two Glu located in the continuous position have a higher chance of binding calcium ions compared to the Glu located in the discontinuous position.

RDF analyses indicate that peptides, especially those with amino acids that favour calcium binding (*e.g.* Asp at the middle N position, Glu at the end of the N terminus and continuous Asp or Glu), are more likely to attract calcium ions. Simulations by Zhu *et al.* showed that the addition of soy isolate protein hydrolysate reduced the RDF peaks of water molecules around starch.[53] This reduction indicates a decrease in water molecules at specific distances, suggesting that changes in the RDF peak could reflect changes in the number of particles. Consequently, these peptides also exhibited high calcium binding in previous chemical assays, demonstrating a correlation between RDF peak patterns and changes in calcium binding. This suggests that peptides with high calcium binding capacity may exhibit this property due to an increased frequency of calcium ion encounters.

3.6.4 Coulomb interaction-short rang. In terms of peptide–calcium binding strength, the Coulomb interaction-short range (Coul-SR) metric was selected for analysis, and we used the "gmx

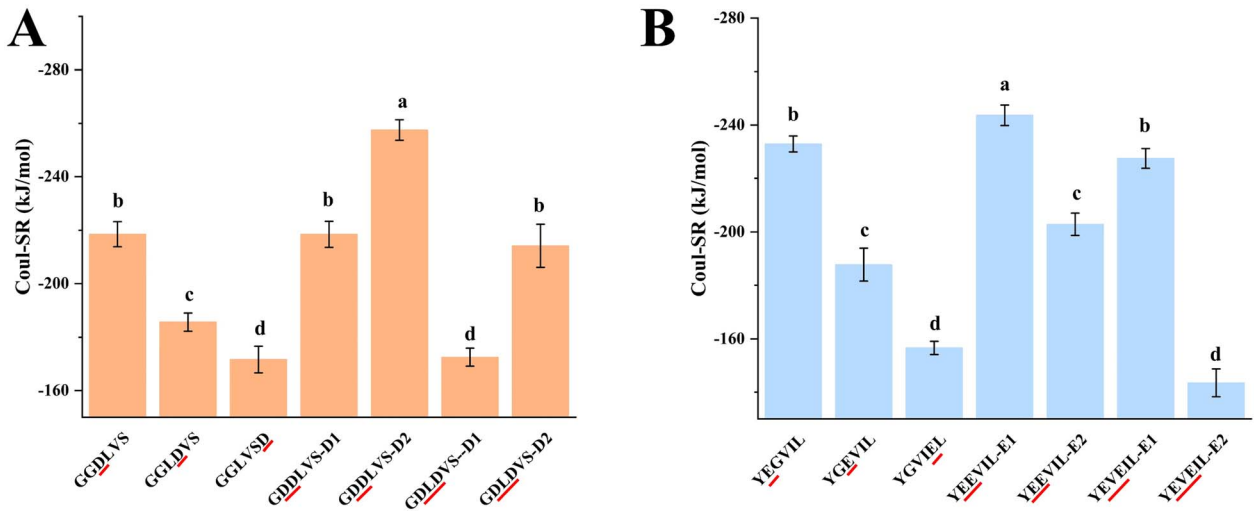

Fig. 7 Mean values of short-range coulombic interactions short-range (Coul-SR) of key amino acids ((A) ASP; (B) Glu) with calcium ions in different peptides. *Two different aspartic acids/glutamates in the peptides are distinguished by D1,D2,E1,E2; different lower case letters represent significant differences, $p < 0.05$.

energy" module in Gromacs to count the Coul-SR of each peptide glutamic acid and aspartic acid with calcium ions during the simulation process.[51] The larger the absolute value of the negative Coul-SR value is, the stronger the binding of the two is.[54] Mishra *et al.* simulated the Coul-SR values for the binding of uranyl ions and zinc ions to human serum proteins and found that zinc ions have a higher binding energy, indicating that zinc ions are more tightly bound to human serum proteins and uranyl ions could not displace zinc ions from binding human serum proteins.[55] This suggests that Coul-SR can be used to determine the binding of metal ions to proteins reliably.

As shown in Fig. 7A and B. The Coul-SR when aspartic acid is located in the N_{mid} position (-218.559 kJ mol^{-1}) and when glutamic acid is located in N_{end} (-232.88 kJ mol^{-1}) are significantly higher ($p < 0.05$), indicating that the binding strength between aspartic acid located in the N_{mid} position or glutamic acid located in N_{end} and calcium ions is higher. For peptides containing two aspartic acids or two glutamic acids, the peptides with contiguous aligned positional features have stronger binding between them and calcium ions. Stronger binding means that the key amino acids are more tightly bound to the calcium ions.

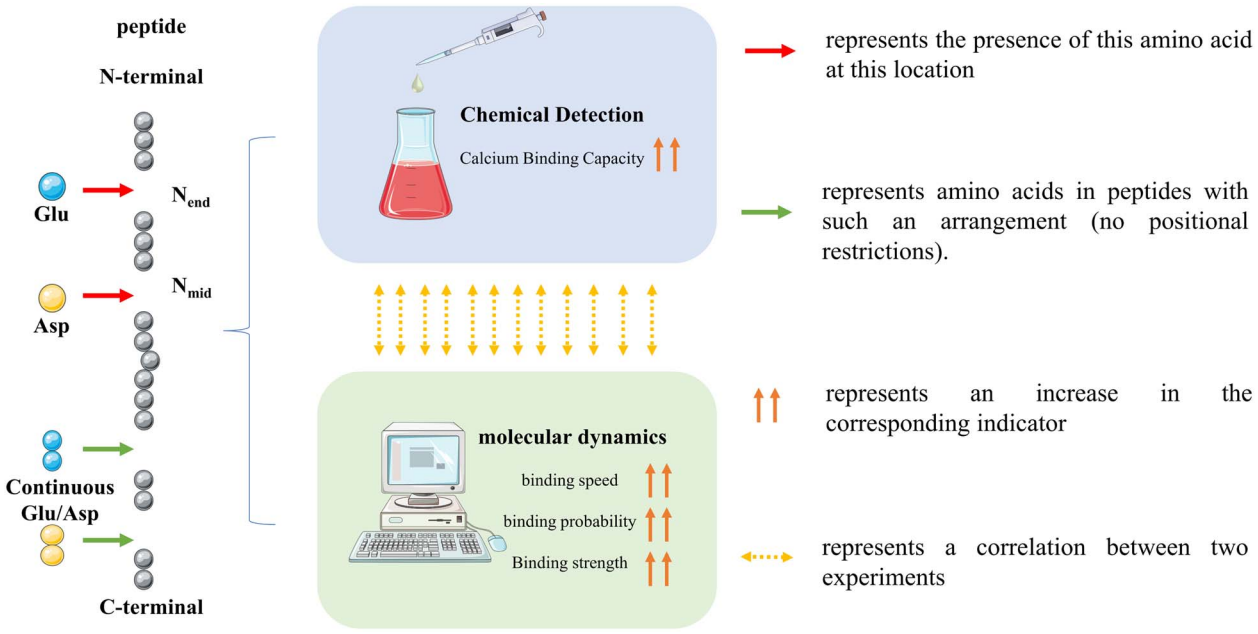

Fig. 8 Schematic representation of the correlation between calcium binding and molecular dynamics results.

3.7 Relationship between molecular dynamics results and calcium binding capacity

Combining the results of calcium binding and molecular dynamics, it can be seen that the peptides with higher calcium binding under the chemical experimental method have higher binding spped, binding probability and binding energy in the simulation process with molecular dynamics. The peptides with high calcium binding have the following structural features: aspartic acid position and N_{mid} position (GGDLVS), glutamic acid in N_{end} position, continuous aspartic acid and glutamic acid. This structure is in accordance with the pattern we have statistically established from the peptide sequences. The relationship between the corresponding indicators involved in the research process is shown in Fig. 8.

4 Conclusions

In this study, soybean peptides were prepared by co-hydrolysis of soy protein using two enzymes. Subsequently, HCBC soy peptide fractions were obtained after purification. Mass spectrometry sequencing showed that in HCBC peptides, the predominant position of aspartate (Asp) for calcium binding was near the N-terminus (N_{mid}), while glutamate (Glu) was near the N-terminus (N_{end}). And there is a higher probability that these two amino acids are present consecutively. Based on these patterns, we screened two representative peptides (GGDLVS and YEGVIL) from. By changing the key amino acid positions and preparing the corresponding peptides. We found that the amino acid arrangement conforming to the laws of HCBC peptides could lead to peptides with higher calcium binding capacity. This suggests that the patterns we found can be used as characteristics of peptides with high calcium binding capacity. The molecular dynamics results show that Asp/Glu with these characteristics is superior to Asp/Glu without these characteristics in terms of binding rate, binding probability and binding strength. The molecular dynamics results correlate with the calcium binding amount determined by chemical method. This further explains the mechanism of its high calcium binding. These findings are important for the efficient preparation of calcium-binding peptides. It also facilitates the research and development of novel and efficient calcium supplements in the future.

In the present experiments, which focused on the rapid screening of soybean peptides, the aim was to discover the link between the primary structural features of the peptides and the calcium binding capacity. Subsequent studies should use relevant instruments for further structural analyses between the obtained peptides and the chelates formed by calcium ions.

Author contributions

Jiulong An: conceptualization; methodology; formal analysis; software development; data curation; writing the original draft and Investigation. Yumei Wang: conceptualization; methodology; validation and writing the original draft. Wenhui Li writing the original and visualization. Wanlu Liu: validation and visualization. Xiangquan Zeng: writing—review and editing. Guoqi Liu: investigation. Xinqi Liu supervision; funding acquisition. He Li writing—review and editing; funding acquisition.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors thank the National Key Research and Development Program of China for its support (No. 2021YFD2100402), and thank Hechi Jinxing Biotechnology Co., Ltd for providing the testing conditions.

Notes and references

1 N. Sun, H. Wu, M. Du, Y. Tang, H. Liu, Y. Fu and B. Zhu, *Trends Food Sci. Technol.*, 2016, **58**, 140–148.
2 Q. Tian, Y. Fan, L. Hao, J. Wang, C. Xia, J. Wang and H. Hou, *Crit. Rev. Food Sci.*, 2021, **63**, 4418–4430.
3 H. Zong, L. J. Peng, S. S. Zhang, Y. Lin and F. Q. Feng, *Eur. Food Res. Technol.*, 2012, **235**, 811–816.
4 F. Liu, L. Wang, R. Wang and Z. Chen, *J. Agric. Food Chem.*, 2013, **61**, 7537–7544.
5 M. Wang, Z. J. Zheng, C. H. Liu, H. Sun and Y. F. Liu, *Food Funct.*, 2020, **11**, 8724–8734.
6 Y. Liu, S. H. Lu, J. Meng, H. Xiang, S. A. Korma, I. Cacciotti and C. Cui, *LWT–Food Sci. Technol.*, 2023, **184**, 114986.
7 J. H. Wu, X. X. Cai, M. R. Tang and S. Y. Wang, *J. Sci. Food Agric.*, 2019, **99**, 536–545.
8 J. L. An, Y. X. Zhang, Z. W. Ying, H. Li, W. L. Liu, J. R. Wang and X. Q. Liu, *Foods*, 2022, **11**, 2762.
9 A. Vidal-Limon, J. E. Aguilar-Toalá and A. Liceaga, *J. Agric. Food Chem.*, 2022, **70**, 934–943.
10 X. Hu, Z. Zeng, J. Zhang, D. Wu, H. Li and F. Geng, *Food Chem.*, 2023, **405**, 134824.
11 Z. Xu, S. Han, H. Chen, Z. Zhu, L. Han, X. Dong, M. Du and T. Li, *Front. Nutr.*, 2022, **9**, 840638.
12 K. Wang, X. Kong, M. Du, W. Yu, Z. Wang, B. Xu, J. Yang, J. Xu, Z. Liu, Y. Cheng and J. Gan, *Nutrients*, 2022, **14**, 1940.
13 M. Zhang and K. Liu, *Food Sci. Biotechnol.*, 2022, **31**, 1111–1122.
14 L. Wen, Y. Jiang, X. Zhou, H. Bi and B. Yang, *Food Chem.*, 2021, **359**, 129970.
15 J. Zhang, W. Li, Z. Ying, D. Zhao, G. Yi, H. Li and X. Liu, *Food Nutr. Res.*, 2020, **64**, 3677.
16 W. Li, Y. Zhang, H. Li, C. Zhang, J. Zhang, J. Uddin and X. Liu, *RSC Adv.*, 2020, **10**, 16737–16748.
17 X. Chen, W. Liu, J. Zhang, H. Li and X. Liu, *Food Funct.*, 2023, **14**, 7882–7896.
18 S. L. Huang, L. N. Zhao, X. Cai, S. Y. Wang, Y. F. Huang, J. Hong and P.-F. Rao, *J. Dairy Res.*, 2015, **82**, 29–35.
19 P. J. A. Cock, T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski and M. J. L. de Hoon, *Bioinformatics*, 2009, **25**, 1422–1423.

20 J. Rey, S. Murail, S. de Vries, P. Derreumaux and P. Tuffery, *Nucleic Acids Res.*, 2023, **51**, 432–437.
21 H. Hou, S. Wang, X. Zhu, Q. Li, Y. Fan, D. Cheng and B. Li, *Food Chem.*, 2018, **243**, 389–395.
22 V. Galiano and J. Villalaín, *Biochim. Biophys. Acta, Biomembr.*, 2015, **1848**, 2849–2858.
23 S. Lin, J. Li, X. Hu, S. Chen, H. Huang, Y. Wu and Z. Li, *Trends Food Sci. Technol.*, 2024, **145**, 104364.
24 S. Budseekoad, C. T. Yupanqui, N. Sirinupong, A. M. Alashi, R. E. Aluko and W. Youravong, *J. Funct. Foods*, 2018, **49**, 333–341.
25 W. Liu, X. Chen, H. Li, J. Zhang, J. An and X. Liu, *Foods*, 2022, **11**, 2361.
26 V. G. Tacias-Pascacio, R. Morellon-Sterling, E.-H. Siar, O. Tavano, Á. Berenguer-Murcia and R. Fernandez-Lafuente, *Int. J. Biol. Macromol.*, 2020, **165**, 2143–2196.
27 P. Setthaya, S. Jaturasitha, S. Ketnawa, T. Chaiyaso, K. Sato and R. Wongpoomchai, *Foods*, 2021, **10**, 2994.
28 J. Lin, X. Cai, M. Tang and S. Wang, *J. Agric. Food Chem.*, 2015, **63**, 9704–9714.
29 F. Toldrá, M. Reig, M. C. Aristoy and L. Mora, *Food Chem.*, 2018, **267**, 395–404.
30 J. Gan, Z. Q. Xiao, K. Wang, X. Kong, M. D. Du, Z. H. Wang, B. Xu and Y. Q. Cheng, *J. Sci. Food Agric.*, 2023, **103**, 2939–2948.
31 H. Ke, R. Ma, X. Liu, Y. Xie and J. Chen, *LWT–Food Sci. Technol.*, 2022, **168**, 113947.
32 H. A. Katimba, R. Wang and C. Cheng, *Crit. Rev. Food Sci. Nutr.*, 2023, **63**, 3959–3979.
33 B. Wang, N. Xie and B. Li, *J. Food Biochem.*, 2019, **43**, e12571.
34 A. Falanga, R. Bellavita, S. Braccia and S. Galdiero, *J. Pept. Sci.*, 2024, **30**, e3558.
35 N. Sun, P. Cui, Z. Jin, H. Wu, Y. Wang and S. Lin, *Food Chem.*, 2017, **230**, 627–636.
36 J. Shao, M. Wang, G. Zhang, B. Zhang and Z. Hao, *Int. J. Food Prop.*, 2022, **25**, 2198–2210.
37 L. Zhang, Y. Lin, S. Wang and J. Aquat, *Food Prod.*, 2018, **27**, 518–530.
38 W. Wu, B. Li, H. Hou, H. Zhang and X. Zhao, *Food Funct.*, 2017, **8**, 4441–4448.
39 W. Huang, Y. Lan, W. Liao, L. Lin, G. Liu, H. Xu, J. Xue, B. Guo, Y. Cao and J. Miao, *LWT–Food Sci. Technol.*, 2021, **149**, 112035.
40 N. Sun, S. Hu, D. Wang, P. Jiang, S. Zhang and S. Lin, *J. Agric. Food Chem.*, 2022, **70**, 2018–2028.
41 N. Kheeree, K. Kuptawach, S. Puthong, P. Sangtanoo, P. Srimongkol, P. Boonserm, O. Reamtong, K. Choowongkomon and A. Karnchanatat, *Sci. Rep.*, 2022, 12.
42 N. Xie, J. Huang, B. Li, J. Cheng, Z. Wang, J. Yin and X. Yan, *Food Chem.*, 2015, **173**, 210–217.
43 K. X. Zhu, X. P. Wang and X. N. Guo, *J. Funct. Foods*, 2015, **12**, 23–32.
44 L. Wang, Y. Ding, X. Zhang, Y. Li, R. Wang, X. Luo, Y. Li, J. Li and Z. Chen, *Food Chem.*, 2018, **239**, 416–426.
45 E. Miquel, A. Alegría, R. Barberá and R. Farré, *Int. Dairy J.*, 2006, **16**, 992–1000.
46 W. K. Jung, R. Karawita, S. J. Heo, B.-J. Lee, S. K. Kim and Y. J. Jeon, *Process Biochem.*, 2006, **41**, 2097–2100.
47 K. Zhang, J. Li, H. Hou, H. Zhang and B. Li, *J. Funct. Foods*, 2019, **52**, 670–679.
48 N. Charoenphun, B. Cheirsilp, N. Sirinupong and W. Youravong, *Eur. Food Res. Technol.*, 2012, **236**, 57–63.
49 K. Si, T. Gong, S. Ding, H. Liu, S. Shi, J. Tu, L. Zhu, L. Song, L. Song and X. Zhang, *Food Chem.*, 2023, **404**, 134567.
50 J. Gan, X. Kong, Z. Xiao, Y. Chen, M. Du, Y. Wang, Z. Wang, Y. Cheng and B. Xu, *Foods*, 2022, **11**, 3290.
51 F. Pan, J. Li, L. Zhao, T. Tuersuntuoheti, A. Mehmood, N. Zhou, S. Hao, C. Wang, Y. Guo and W. Lin, *J. Food Biochem.*, 2020, **45**, e13570.
52 M. S. Cates, M. L. Teodoro and G. N. Phillips, *Biophys. J.*, 2002, **82**, 1133–1146.
53 Z. J. Zhu, C. J. Sun, C. H. Wang, L. P. Mei, Z. X. He, S. Mustafa, X. F. Du and X. Chen, *Int. J. Biol. Macromol.*, 2024, **255**, 128213.
54 B. Huang, M. Li, H. Mo, C. Chen and K. Chen, *Int. J. Mol. Sci.*, 2022, **23**, 10144.
55 V. Mishra, A. K. Pathak and T. Bandyopadhyay, *J. Biomol. Struct. Dyn.*, 2022, **41**, 7318–7328.