Journal of Colloid and Interface Science 652 (2023) 405–417

Contents lists available at ScienceDirect

Journal of Colloid And Interface Science

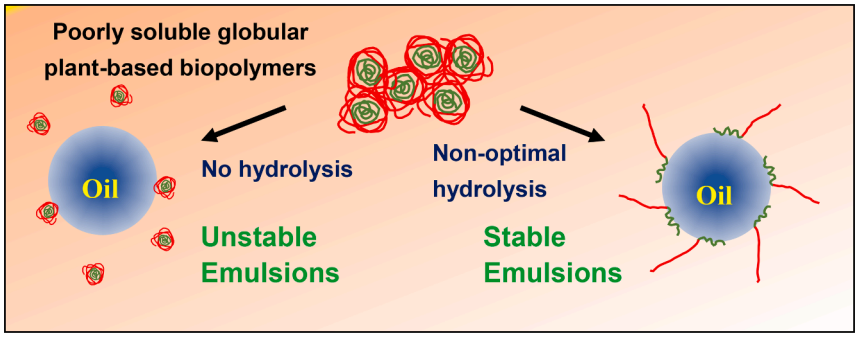
journal homepage: www.elsevier.com/locate/jcis

Surface adsorption properties of peptides produced by non-optimum pH pepsinolysis of proteins: A combined experimental and self-consistent-field calculation study

Cuizhen Chen [a], Brent S. Murray [a,*], Rammile Ettelaie [a,*]

[a] Food Colloids & Bioprocessing Group, School of Food Science & Nutrition, University of Leeds, Leeds LS2 9JT, UK

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Protein hydrolysates
Adsorption
Self-consistent field calculations (*SCFC*)
Emulsion stability
Partial hydrolysis

ABSTRACT

Hypothesis: Partial hydrolysis of large molecular weight (M_w), highly aggregated plant proteins is frequently used to improve their solubility. However, if this hydrolysis is extensive, random or nonselective, it is unlikely to improve functional properties such as surface activity, emulsion, or foam-stabilising capacity.

Experiments and simulation: Soy protein isolate (*SPI*) was hydrolysed by pepsin under optimal (pH 2.1) and non-optimal (pH 4.7) conditions. The surface activity and emulsion stabilising capacity of the resultant peptides were measured and compared. The colloidal interactions between a pair of emulsion droplets were modelled via Self-Consistent-Field Calculations (*SCFC*).

Findings: Hydrolysis at pH 2.1 and 4.7 resulted in a considerable increase in measured surface activity compared to the native (non-hydrolysed) *SPI*, but the hydrolysate from pH 2.1 was not as good an emulsion stabiliser as the hydrolysate (particularly the fraction $M_w > 10$ kDa) at pH 4.7. Furthermore, peptide analysis of the latter suggested it was dominated by a fragment of one of the major soy proteins β-conglycinin, with $M_w \approx 25$ kDa. *SCFC* calculations confirmed that interactions mediated by adsorbed layers of this peptide point to it being an excellent emulsion stabiliser.

* Corresponding authors.
E-mail addresses: b.s.murray@leeds.ac.uk (B.S. Murray), r.ettelaie@leeds.ac.uk (R. Ettelaie).

https://doi.org/10.1016/j.jcis.2023.08.040
Received 30 March 2023; Received in revised form 1 August 2023; Accepted 6 August 2023
Available online 9 August 2023

1. Introduction

In a recent publication, several authors speculated together on what may make "the perfect hydrocolloid stabiliser" [1]. This was not merely an idle exercise because, although the principles of colloidal stability are well established, in reality, a stabiliser has to fulfil or impart a number of related capabilities. For example, if one is concerned with biocompatible colloidal products, as in foods, pharmaceuticals, cosmetics and agrochemicals, then the stabiliser should ideally be based on natural, non-covalently modified materials. In the realms of hydrocolloids, this principally means protein and polysaccharide biopolymers, which also have the advantage of being based on renewable resource materials. In terms of proteins with a lower carbon and more general environmental footprint, this has also recently led many researchers to investigate more widely the potential of plant proteins to replace animal-based proteins (e.g., from milk or meat sources) as alternatives [2].

Proteins are the surface-active biopolymer par excellence. They adsorb to almost any surface, so providing good interfacial coverage and anchoring to the surface. At the same time, in polar media (e.g. water), sufficiently long polar and ionisable regions of the structure may protrude away from the surface to induce sufficient steric and/or electrostatic stabilisation to the colloidal particles, to the surface of which the proteins are adsorbed. Purely polysaccharide-based structures are rarely surface active unless they happen to have some protein naturally associated with this structure (as in the cases of gum arabic [3] or sugar beet pectin [4]). In other examples, the polysaccharide and protein can become effectively cross-linked in some way, as in the case of Maillard conjugates [5] generated by simple heating, or strong electrostatic binding [6], or enzymatic [7] or chemical [8] cross-linking. However, the latter two types of cross-linking begin to somewhat deviate from the requirements of having 'natural' stabilisers.

Apart from the basic requirement of producing a sufficiently strong repulsive interaction force to keep the colloidal particles apart, the other capabilities often referred to include a host of factors that are not so frequently (or easily) addressed in their entirety. These include the dynamic capability of the adsorbed layer to form and heal under the wide range of stresses and strains experienced during processing or application of the product. At the same time, the adsorbed layer might have to maintain stability under a wide range of pH, salt and temperature conditions, due to environmental factors or digestive conditions, if the product is ingested. In addition, the stabiliser might also have the responsibility of imparting the correct mouthfeel or skin feel, i.e., lubrication properties [9,10].

The wide-ranging nature of these different desired capabilities means that, not surprisingly, no single type of protein-polysaccharide combination can entirely meet the varied requirements in all cases. Therefore, the search for new stabilisers that provide the additional advantages continues, as indicated above. This almost naturally brings one back to consider plant proteins and their associated polysaccharides, because so far these have been relatively under-utilized (or underappreciated) as surface active ingredients. There are a number of reasons for this, but one generic fact is the relative insolubility of plant proteins compared to dairy or meat proteins. This should not be surprising, given that plant proteins are largely storage proteins of the seeds of cereals, legumes, nuts, etc., designed not to be easily dispersed or digested until utilised by the seed in germination. Attempts to improve the solubility and therefore the functional (gelling, film-forming, emulsifying, foaming) properties of plant proteins include extensive thermal and mechanical processing [11], chemical modification (such as succinylation and acetylation) [12] and enzyme treatment [13]. Of the latter, hydrolysis via proteolytic or glycosidic enzymes remains one of the most commonly utilized techniques.

It was mentioned above that the principles of colloidal stability are relatively well established [14,15]. In other words, it is now possible to predict to a reasonable degree the equilibrium conformation of proteins at interfaces and hence the effect that this will have on the mediated colloidal interaction potential between two surfaces (particles) when covered by such biopolymers. Most extant amongst these predictive theories is the Scheutjens-Fleer methodology involving self-consistent field calculations (*SCFC*) [16], particularly suited to dense adsorbed polymer layers. One strong suggestion from these studies is that relatively minor degrees of random peptide bond breakage (say 10%) of any protein will result in fragments that are simply too short to provide sufficient steric (and possibly also electrostatic) stabilisation, as we shall demonstrate later below. At the same time these short fragments may be more surface active and displace (longer) less surface active fragments [17], where the latter would have provided a sufficiently large repulsive interaction energy if they had been able to remain on the interface. Most experimental data so far seem to suggest that too much hydrolysis is not conducive to achieving fragments with suitable emulsion stabilizing characteristics [18], despite further possible improvements in solubility. For example, a relatively recent study by Liu et al, involving hydrolysis of native Fava bean protein isolate, found that a degree of hydrolysis (D_H) of 4% resulted in polypeptides that had superior emulsification properties compared to the intact protein, or those that were 9% or 15% hydrolysed [19]. A similar study by Yue et al on the impact of hydrolysis of soy protein isolate, prior to conjugation of fragments with maltodextrin, found an optimum D_H of 8%. Further hydrolysis beyond this optimum value caused a deterioration of the emulsification ability of the conjugates [5]. More evidence for this trend is also reported by several other studies (e.g. [20,21]). However, it must be noted that the difficulty in controlling D_H and the varieties of different proteins and enzymes involved, mean that there are also a few exceptions to this trend. Though not often observed, in some studies no tangible improvement in emulsification capacity of the protein was found at any level of hydrolysis [22]. Yet, in some other work it is reported that the fragments continue to become better emulsifiers with increasing hydrolysis, with no apparent optimum D_H value reached [23].

In the present work we explore a strategy of using enzymes at their sub-optimal pH for hydrolysis of plant proteins to generate fragments that are on one hand sufficiently soluble, but on the other remain large enough to act as reasonable steric and electrostatic emulsion stabilisers. This is exemplified here by the use of the soya protein β-conglycinin. The deliberate use of sub-optimal conditions is to make the enzyme more selective and thus prevent extensive hydrolysis. The experimental work here is in part supported by *SCFC*. The great advantage of the *SCFC* methodology is that the properties of many and multiple (i.e., mixtures [24] of) sequences can be predicted easily and quickly under a wide range of physical conditions (e.g., ionic strength, pH) that would otherwise take an enormous amount of time and effort to test experimentally. In the case of the caseins from milk the correspondence between theory and experiment have been shown to be quite remarkable [25–27], but partly because these individual proteins possess relatively little secondary or tertiary structure. One expects that enzymatic fragmentation of plant proteins will also serve to destroy much of their tertiary and a large part of the secondary structures, thus making *SCFC* applicable to the present problem.

The described study demonstrates the effects of limited hydrolysis of a plant protein (from soy) in improving its colloid (emulsion) stabilising properties and provides a plausible explanation of this via *SCFC* calculations, based on some of the actual peptide fragments produced from β–conglycinin. Thus, one may be able to move further down the road of predicting what type of treatment, and for which plant proteins, is most likely to produce materials with the optimum hydrocolloid stabilising properties, applying the proposed strategy to many other varieties of plant proteins in future.

2. Materials and methods

2.1. Materials

Commercial soybean protein isolate (*SPI*) powder with 90% purity

was purchased from Pulsin (Gloucester, United Kingdom). Micellar Casein with 83% purity was purchased from the company BulkTM (London, United Kingdom). Pepsin from porcine gastric mucosa (≥250 units/mg), alcalase with specific activity > 2.4 U/g (P4860), *n*-tetradecane (purity 99.0 %), ultrafiltration discs (3 kDa and 10 kDa NMW), picrylsulfonic acid solution 5 wt% (2,4,6-trinitrobenzenesulfonic acid solution, TNBS), ammonium bicarbonate, iodoacetic acid, trifluoroacetic acid and Kromasil C18 media were all purchased from Sigma-Aldrich (Gillingham, United Kingdom). Trypsin, chymotrypsin and glutamic protease were purchased from Promega – (Madison, WI, USA). Acetonitrile and formic acid were from Fisher Scientific (Loughborough, UK) and dithiothreitol (DTT) was from Roche Diagnostics (Welwyn Garden City, UK). Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) chemicals: Laemmli sample buffer, Mini Protean Tris-Glycine eXtended (TGX) 4–20% gel, Precision Plus Protein dual color standards and tris/glycine/SDS were from BioRad (Watford, UK); InstantBlue Coomassie stain was from AbCam (Cambridge, UK).

2.2. Methods

2.2.1. Preparation of soybean protein hydrolysates (SPH)

The preparation method for soybean protein hydrolysates followed the method of Han et al. [28] with slight modifications. Commercial soybean protein powder (5 wt%) was dissolved in Milli-Q water and the pH adjusted to 1.3, 2.1 and 4.7. Pepsin, at an enzyme to substrate ratio of 1: 25, was added to the solution and stirred for 15 min at room temperature. Then, the mixtures were immersed in a 37℃ water bath for 6 h, followed by heating in a boiling water bath for a further 15 min to denature and deactivate the pepsin. When the mixtures were cooled back down to room temperature, any precipitates were removed by centrifugation (4000*g* X 20 min 4 °C), with pH adjusted to 7. After centrifugation, the mixtures were filtered through 10 or 3 kDa ultrafiltration membranes. Thus, a series of *SPH* samples with three different molecular weight (*Mw*) ranges: ≤3 kDa, 3 to 10 kDa and ≥10 kDa, were obtained. Finally, the *SPH* samples were freeze-dried and stored at 4 °C in a fridge for further use.

2.2.2. Preparation of emulsions stabilised by SPH

SPH-stabilized emulsions were made with 10% *n*-tetradecane and 90% aqueous solution containing 1% protein/peptides. The two phases were blended at room temperature (20 to 25 °C) using an IKA homogeniser (12,000 rpm, 10 min) to form coarse droplets, followed by 3 passes through the Leeds Jet Homogenizer [29] at 500 bar. The emulsions also had 0.02 wt% sodium azide added as an antimicrobial agent.

2.2.3. Surface tension measurements

The surface tension γ of various *SPH* solutions (10^{-3} wt%) in 0.05 M phosphate buffer at pH 7 was measured via an Ez-Pi Plus Kibron tensiometer (Kibron Inc, Finland). Measurements were taken every 6 s at room temperature. The calibration reference was the γ of pure water at 20 °C – taken as 72.8 mN m^{-1}. Data was collected until γ became constant (±0.3 mN m^{-1}), which was after no more than 10 min.

2.2.4. Droplet size and zeta potential measurements

The particle size distribution (*PSD*) of protein solutions prior to emulsification and for the emulsions stabilised by the same protein solutions were measured via a Mastersizer 3000 (Malvern Panalytical Instrument, USA). All emulsions were shaken by hand to ensure homogenous sampling before the size measurements. PSDs are shown in the Supplementary data (Figure S2) and summarised via the surface weighted and volume weighted means $D_{3,2}$ and $D_{4,3}$, respectively. The refractive index for soybean protein was taken as 1.45 (for the PSDs of the protein dispersions), whilst the refractive index of the oil was taken as 1.43. All zeta potentials were measured using a Zetasizer Nano ZS (Malvern Panalytical Instrument, USA) and calculated using Henry's equation with Smolukowski approximation for polar media. The results reported are the averages of 3 separate measurements ± the standard deviation.

2.2.5. Degree of hydrolysis measurements

The degree of hydrolysis (D_H) was estimated by a modified TNBS method [28]. The principle behind this method is the reaction between the TNBS reagent and *N*-terminal amino acid groups. In brief, concentrated TNBS reagent (5 wt%) was diluted 25 times in 0.1 M $NaHCO_3$ buffer at pH 8.5. Leucine solutions of a range of concentrations were used as the standard solutions for the calibration curve. In this study, 40 μL diluted TNBS was added to 80 μL protein solutions or the standard solutions at 37 °C for 2 h. Then, 80 μL of 1 mol dm^{-3} HCl and 40 μL of 10 wt% SDS buffer solution were added to the samples to terminate the hydrolysis reaction. Finally, the absorbance of each sample was measured at 330 nm with water as a reference. The D_H value of the samples were calculated from:

$$D_H = 100 \times (C_H - C_0)/C_T \quad (1)$$

where C_H is the equivalent leucine concentration (as obtained from the TNBS method above) of the hydrolysed sample; C_0 is the equivalent leucine concentration of the non-hydrolysed sample; C_T is the *theoretical* equivalent leucine concentration if the protein was completely hydrolysed to its constituent amino acids, calculated from the total amine nitrogen groups in *SPI*.

It should be noted that the values of D_H calculated in this way will probably always be <100% because it will be almost impossible to achieve 100% hydrolysis of the soy protein in practice by pepsin under any conditions, so that the denominator in Eq. (1) is artificially high compared to the true value when practical pepsin action may be said to be 'complete'.

2.2.6. Isolation and identification of peptide sequences

2.2.6.1. SDS-PAGE electrophoresis of peptides and further peptide band digestion. The peptide sample was mixed 1:1:1 with 150 mM DTT and Laemmli sample buffer and heated to 95 °C for 5 min. 10 uL of each sample were loaded into a separate well of a Mini Protean TGX 4–20% gel. Precision Plus Protein dual color standards were used as molecular weight markers. For gel separation a constant voltage of 200 V was applied for 30 ± 5 min - until the dye front reached the lower edge of the gel. The running buffer used was 1x tris/glycine/SDS.

The gels were then rinsed with distilled water and stained with InstantBlue Coomassie stain for ~ 15 min until bands were visible. The gels were stored in water until band excision. Gel bands were excised and chopped into ~ 1 mm^3 pieces with a scalpel. Bands were de-stained by covering them with 30 % ethanol and incubating them at 70 °C for 30 min with shaking. The supernatant was then discarded. This was repeated until all the stain was visibly removed. The gel pieces were washed by covering with 25 mM ammonium bicarbonate/50% acetonitrile and incubating in this solution for 10 min with shaking, then the supernatant was discarded. To reduce cystine residues, 100 μL of 10 mM DTT solution was added and the pieces incubated at 57 °C for 1 h with shaking. The supernatant was discarded and the gel pieces allowed to re-equilibrate to room temperature. Cysteine residues were then alkylated by addition of 100 μL 55 mM iodoacetic acid and incubated at room temperature in the dark for 45 min with shaking. Then the supernatant was discarded. Gel pieces were dehydrated by addition of 100% acetonitrile for 5 min at room temperature. The gel pieces were removed from the acetonitrile and left to dry in a laminar flow cabinet for 60 min. Once dry, the gel slices were cooled on ice then covered by the addition of ice-cold trypsin, chymotrypsin or glutamic proteases solution (20 ng μL^{-1} in 25 mM ammonium bicarbonate) and left on ice for 10 min to rehydrate. Excess enzyme solution was removed and the gel pieces covered with 25 mM ammonium bicarbonate. After briefly vortexing and centrifuging, the gel pieces were then incubated at 37 °C with shaking for 18 h. The

resulting digest was (a) vortexed, centrifuged and then the supernatant added to an Eppendorf tube containing 5 μL acetonitrile/ water/ formic acid (60/35/5; v/v) to quench protease activity. Then (b) 50 μL acetonitrile/ water/ formic acid (60/35/5; v/v) was added to centrifugate, vortexed for 10 min and re-centrifuged to produce a second supernatant that was combined with the first. The supernatant was pooled with the previous supernatant. Step (b) was then repeated once more to produce a pool of 3 supernatants that was then dried by vacuum centrifugation. The peptides were reconstituted in 20 μL 0.1% aqueous trifluoroacetic acid.

2.2.6.2. LC-MS-MS of peptides. The main variables affecting the degree of hydrolysis - time, temperature, stirring and enzyme concentration were easily and tightly controlled. However, the state of dispersion of the SPI throughout the digestion is possibly another variable. For this reason, 3 separate hydrolysates were prepared and these pooled before the following lengthy and detailed peptide analysis. A 3 μL sample of the pooled hydrolysates (approx. 0.6 ug of protein) was injected onto an in house-packed 20 cm capillary column (inner diameter 75 μm, 3.5 μm Kromasil C18 media). An Ultimate 3000 nano liquid chromatography system was used to apply a gradient of 2–30% acetonitrile in 0.1% formic acid over 30 min at a flow rate of 300 nL/min. Total acquisition time was 60 min including column wash and re-equilibration. Peptides were eluted from the column and into an Orbitrap Exploris 240 Mass Spectrometer (ThermoFisher Scientific, Hemel Hempstead, UK) via a nanospray flex ion source using a capillary voltage of 2.7 kV. Precursor ion scans were acquired in the Orbitrap with a resolution of 60000. EASY-IC internal calibration was used for precursor ion scans. Up to 20 ions per precursor scan (charge state 2+ and higher) were selected for HCD fragmentation using a normalised collision energy of 30%. Fragments were measured in the Orbitrap at a resolution of 15000. Dynamic exclusion of 30 s was used.

2.2.6.3. Data analysis. Peptide MS/MS data were processed with PEAKS Studio XPro (Bioinformatic Solutions Inc, Waterloo, Ontario, Canada) and searched against the soy database. Carbamidomethylation was selected as a fixed modification, variable modifications were set for oxidation of methionine and deamidation of glutamine and asparagine. MS mass tolerance was 20 ppm, and fragment ion mass tolerance was 0.05 Da. The peptide false discovery rate was set to 1%.

2.2.7. Self-consistent field (SCF) calculations

This study applied a self-consistent field (SCF) calculation scheme to estimate the interactions between emulsion droplets mediated by adsorbed protein layers on their surfaces. In the context of polymers at interfaces, SCF calculations were first used by Dolan and Edwards [30] to calculate the interactions induced by such overlapping layers of polymers on two adjacent surfaces. Later, Scheutjens and Fleer [31] introduced a new more efficient scheme for implementing such calculations. They were able to identify sections of the chains divided into tails, loops and trains for the first time to describe the behavior of polymers on the interfaces more accurately. They were also able to determine free energy changes resulting from the overlap of such polymer layers, which in turn can influence the emulsifying and colloid stabilising capacity of macromolecules being studied. The SCF calculations used in the present study are based on the general framework of the so called Scheutjens-Fleer theory [31]. The central aspect of this scheme has been described in many earlier papers [25,26,32–34]. Therefore, we only provide a limited discussion of the more critical parameters and the information that is more specific to the current work.

The gap between two adjacent parallel surfaces is divided into a set of lattice sites with each site occupied by one of eight possible monomers: i.e. five kinds of possible amino acids (see Table 1), two type of ions or solvent molecules.

The distance between two opposite interfaces, L, is considered as consisting of equally separated layers parallel to the surfaces, each with a thickness equal to the size of a grid point, a_0 (the nominal value of a_0 is taken to be the size of a peptide bond ~0.3 nm here). We can vary the distance between the surfaces from 2 to 120 layers (z = 120) in the present study, giving a maximum separation distance of ~36 nm. This is sufficient in most cases for the two surfaces to be far enough to be considered as isolated from each other. All internal interactions between different types of monomers, monomers and solvent and those with ions are expressed by the potential of mean forces in SCF calculations. These potentials are to be calculated at each layer and for each type of monomer (as well as solvent and ions). The mean potentials are in turn themselves depend on the concentration profiles $\{\varnothing_i^\alpha(z)\}$ of various monomers. Here, $\varnothing_i^\alpha(z)$ represents the volume fraction of monomers of kind α, belonging to molecules (chains) of type i, residing in layer z. Normalised in units of k_BT, where k_B is the Boltzmann constant and T the temperature ($T = 298$ K is assumed in this study), the mean potential for monomers of type α has three components as given by the equation below:

$$\psi^\alpha(z) = \psi_{hc}(z) + q_\alpha\psi_{el}(z) + \psi^\alpha_{\text{int}}(z) \tag{2}$$

First of these, $\psi_{hc}(z)$ is a hard-core potential term acting equally on all monomers in a given layer, irrespective of their kind. This is the interaction arising from the crowding of different monomers in the same layer, which enforces and ensures the incompressibility of the system. The term $q_\alpha\psi_{el}(z)$ is a long-ranged electrostatic interaction between charged species, only present if the monomers of type α possess an electrical charge. Its value is proportional to the charge of monomer of type α, i.e.q_α, and to the electric potential at layer z, $\psi_{el}(z)$. The latter is in turn given by the solution to Poisson's equation $\varepsilon_r\varepsilon_0\nabla^2\psi_{el}(z) = -\sum_i\sum_\alpha q_\alpha\phi_i^\alpha(z)$ in which vacuum permittivity $\varepsilon_0 = 8.85\text{x}10^{-12}$F/m and the dimensionless parameter $\varepsilon_r = 78.5$ is the relative permittivity of water. The final component in equation (2) above, i.e. $\psi^\alpha_{int}(z)$, is a short-ranged contribution arising from the interaction of a monomer with other neighboring monomers around it. Its strength is specified using the Flory-Huggins interaction parameter $\chi_{\alpha\beta}$ between two dissimilar monomer types α and β. The various values of $\chi_{\alpha\beta}$ used in our calculations are those following the work of Leermakers et al [25] and are shown in Table 2.

In general, the calculation of the density profiles $\{\varnothing_i^\alpha(z)\}$ can be carried out if the values of the corresponding fields $\{\psi^\alpha(z)\}$ are known [17,25,35–37]. However, the set of fields $\{\psi^\alpha(z)\}$ are not available a priori and in turn are themselves dependent on the spatial distribution of various monomers comprising the chains, viz. equation (2). To overcome this issue, an iterative procedure is implemented in which, starting with a guess set of fields $\{\psi^\alpha(z)\}$, the density distributions are calculated. Then using these calculated values of $\{\varnothing_i^\alpha(z)\}$ a new set of fields is obtained from equation (2). This is done with the aid of the segment density functions as has been discussed in many articles and reviews. The process is repeated until convergence is obtained. At this point, the fields lead to a set of density profiles which in turn gives the same values of the fields $\psi^a(z)$ and $\psi_{el}(z)$. Once these values are available, one can proceed further to determine the free energy change in the system. This is given by [17,33]

Table 1
Classification of different types of amino acid residues into five groups [17].

1	Hydrophobic	Pro, Ile, Gly, Leu, Val, Phe, Ala Met, Trp
2	Polar (non-charged)	Gln, Asn, Ser, Thr, Tyr
3	Positive	Arg, Lys, *N*-terminus
4	Positive	His
5	Negative	Glu, Asp, C-terminus

Table 2
Flory-Huggins interaction parameters between different kinds of monomers used in our calculations.

Monomer type	0	1	2	3	4	5	6	Ion+	Ion-
0 - solvent	0	1	0	0	0	0	0	−1	−1
1 – hydrophobic residues	1	0	2.0	2.5	2.5	2.5	2.5	2.5	2.5
2 – polar residues	0	2.0	0	0	0	0	0	0	0
3 - positive residues	0	2.5	0	0	0	0	0	0	0
4 - histidine (His)	0	2.5	0	0	0	0	0	0	0
5 - negative residues	0	2.5	0	0	0	0	0	0	0
6 - positive ions	−1	2.5	0	0	0	0	0	0	0
7 - negative ions	−1	2.5	0	0	0	0	0	0	0
s - surface	0	−2.0	0	0	0	0	0	0	0

$$F(L) = \left[-\sum_i \frac{1}{N_i} \sum_{z=1}^{L} \sum_{\alpha} \left(\phi_i^{\alpha}(z) - \Phi_i^{\alpha}\right) \right] - \sum_i \sum_{z=1}^{L} \sum_{\alpha} \psi_{\alpha}(z)\phi_i^{\alpha}(z) + \left[\frac{1}{2} \sum_{z=1}^{L} \sum_{\alpha,\beta} \sum_{i,j} \chi_{\alpha\beta} \left(\phi_i^{\alpha}(z) - \Phi_i^{\alpha}\right) \left(\left\langle \phi_j^{\beta}(z) \right\rangle - \Phi_j^{\beta} \right) \right] + \left[\frac{1}{2} \sum_{z=1}^{L} \psi_{el}(z) \sum_i \sum_{\alpha} q_{\alpha} \phi_i^{\alpha}(z) \right] + \sum_i \sum_{\alpha} \chi_{as} \left[\phi_i^{\alpha}(1) + \phi_i^{\alpha}(L) \right] \quad (3)$$

for two surfaces at a separation distance L. Here N_i is the degree of polymerisation of the polymer chains (where for ions and solvent molecules we simply have $N_i = 1$). The quantity $< \phi_j^{\beta}(z) >$ represents the average volume fraction of the neighboring monomers of kind β surrounding a monomer placed in layer z, whereas Φ_j^{β} is the corresponding value in the bulk solution. The averaged values $< \phi_j^{\beta}(z) >$, are calculated as follows:

$$\left\langle \phi^{\beta}(z) \right\rangle = \lambda_{-1}\phi_{\beta}(z-1) + \lambda_0\phi_{\beta}(z) + \lambda_{+1}\phi_{\beta}(z+1) \quad (4)$$

with the weight factors $\lambda_{-1} = \lambda_{+1} = 1/6$ and $\lambda_0 = 4/6$ for the cubic lattice adopted in our calculations here. The weight factors reflect the number of adjacent sites in each neighboring layer to any given grid point. The monomers residing in layers 1 and L also interact with one or the other surface. The last term in equation (3) accounts for such interactions. The difference $F(L) - F(\infty)$ between the free energy when the surfaces are a distance L apart, and when they are isolated (i.e. far apart), is precisely the interaction potential induced between them due to the presence of polymers. If required, it is possible to convert these results to interactions between two spherical colloidal particles, using the well-known Derjaguin approximation [38], as follows:

$$V(z) = \pi R \int_z^{\infty} [F(z') - F(\infty)] dz' \quad (5)$$

where R is the radius of the droplets, z the closest approach distance between the surface of drops, and we assume that for separation distances of interest $z \ll R$.

2.2.8. Protein identification and PTM mapping

SPH samples were analysed (Mass Spec Facility, Faculty of Biological Science, University of Leeds) to confirm their protein identity and determine the primary structures.

2.2.9. Statistical analysis

All the measurements, unless specifically stated otherwise, were conducted on three identical samples in triplicate. The obtained data were averaged, and the results provided as mean values. The error bars in graphs, and the quoted standard errors in the tables, were the calculated standard deviations. All the calculations were analysed by Microsoft Excel 2016.

3. Results and discussion

3.1. Predicted average molecular weight change as function of hydrolysis conditions

Since this study investigates the emulsification properties of fragmented proteins, it is very revealing to understand the *Mw* change of fragments in the hydrolysate at different D_H values. The degree of hydrolysis, D_H, is defined here as the number of bonds cleaved by the protease divided by the total number of bonds on the protein backbone.

$$D_H = \frac{N_{broken}}{N_{total}} \quad (6)$$

Therefore, for a given D_H value, we can obtain the relative average *Mw* of fragments compared to the non-hydrolysed protein:

$$\frac{\left\langle MW_{\text{fragment}} \right\rangle}{MW_{\text{intact}}} = \frac{1}{D_H \times (N_i - 1) + 1} \quad (7)$$

where N_i is the total number of amino acid residues of the intact protein; $\left\langle MW_{\text{fragment}} \right\rangle$ is the average *Mw* of fragments in the hydrolysate; MW_{intact} is the molecular weight of the intact protein. Note that for a selective enzyme, an alternative definition for D_H is also often used. This is defined as the number of broken bonds relative to the total bonds that are susceptible to breakage, as opposed to all the peptide bonds. For the case displayed in Fig. 1, involving an indiscriminate enzyme, the two definitions are of course identical.

As hinted in Fig. 1, the average *Mw* declines rapidly with D_H and roughly as $1/D_H$ for chains with a large degree of polymerisation. This is

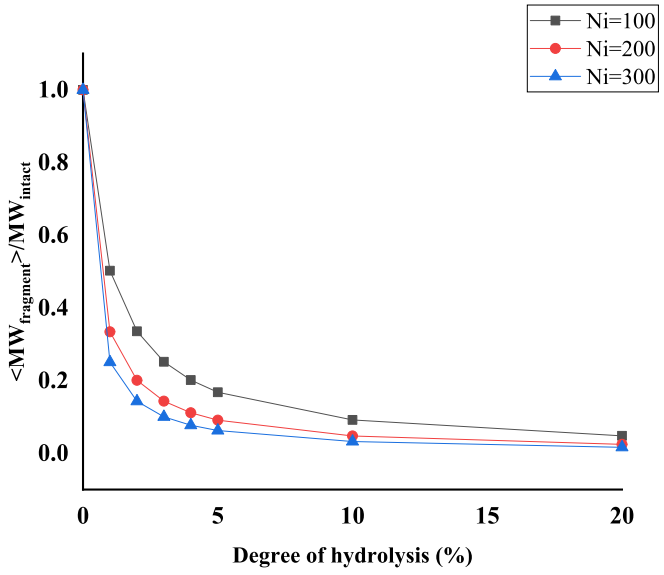


Fig. 1. Calculated average *Mw* of hydrolysates relative to that of the intact polymers, plotted against the degree of hydrolysis (D_H).

most notable in the D_H interval from 0% to 10%. Take $N_i = 100$ as an example; when the D_H value increases to 10% and further 20%, the average *Mw* of the resulting fragments drops approximately to 0.09 and 0.04 of the original intact chains, respectively. This means that most of the fragments in such hydrolysate consist of short chains – with less than ten residues. Moreover, for larger protein chains, the average *Mw* will decline even more rapidly with D_H (see the curves for $N_i = 200$ and $N_i = 300$ in Fig. 1). On average, and irrespective of its size, the *Mw* of a chain drops by approximately a factor of 2 each time a bond on its backbone is broken. However, for a given value of D_H the number of broken bonds scales with the size of the chain. Therefore, it is clear that for larger chains the drop will be more significant for the same D_H. While this may seem an obvious point to make, it is emphasized here due to its practical importance. In our experiments, it is required that the fragmented proteins will provide enough steric and electrostatic repulsion to stabilise the O/W emulsion droplets. Such fragments must contain a relatively long hydrophilic block to form a thick enough protruding layer and a hydrophobic block of sufficient size to strongly absorb at the oil–water interface. On the other hand, electrostatic repulsion is partly dominated by the magnitude of charges close to the droplet surfaces, according to DLVO theory [39]. Nevertheless, it is clear that limited hydrolysis is preferred otherwise the vast majority of the peptides will be too short to provide sufficient electro-steric stabilisation. This is why we deliberately chose a range of pH values during hydrolysis that will have different efficiencies of hydrolysis and also used filtration to separate out a range of different M_w for each hydrolysis conditions.

Our experiments considered three different pH values (1.3, 2.1 and 4.7) to hydrolyse soybean protein by pepsin, because pepsin's specificity and activity are significantly different at these pH values. Pepsin is more specific but with relatively low activity at pH 1.3. The specificity tends to be lost when the pH is higher than 2. At pH 2.1, pepsin obtains the maximum activity with a broad specificity. We also chose a sub-optimal value of pH 4.7, where we expect broader specificity but only moderate activity [40] to generate more varied but also perhaps longer fragments.

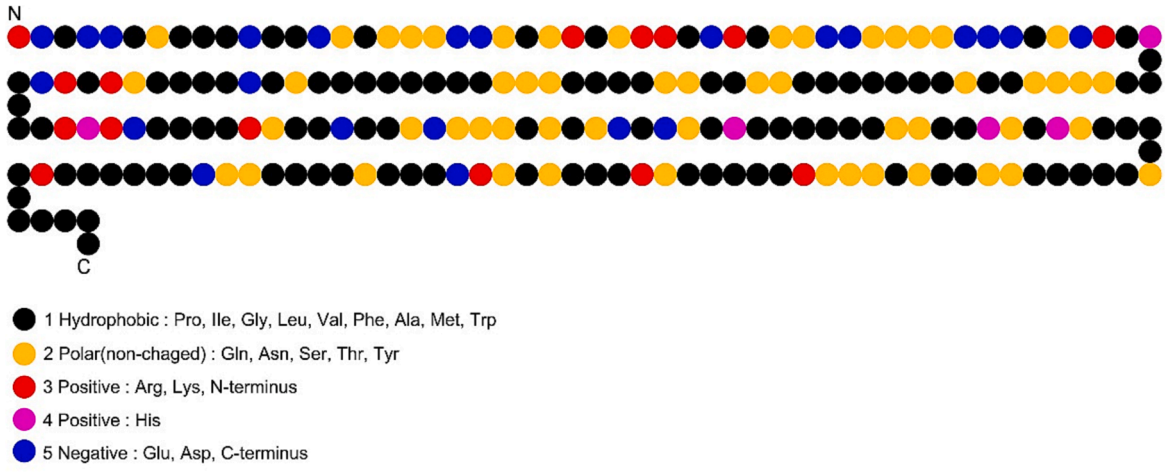
3.2. Calculation of the interaction potential

Section 3.1 above stresses the importance of D_H in determining the M_w of the fragments obtained. In this section, we use *SCFC* to illustrate the effect of M_w on colloidal interaction forces. Previous research has shown that the milk protein β-casein displays excellent emulsification properties in practice [41–43]. Moreover, β-casein has been considered as one of the few natural proteins with an approximately di-block-type structure [44]. Fig. 2 shows that β–casein consists of a predominantly hydrophilic *N*-terminus side and a mostly hydrophobic C-terminus end. Therefore, even without the contribution of electrostatic repulsion from charged amino acid residues, the di-block structure of β-casein is expected to induce a reasonable steric repulsion between droplets. Though probably not quite enough to stabilise the emulsions by itself, the provision of such steric repulsion can greatly enhance the good emulsification ability of this protein.

However, it should also be noted that a diblock–like structure is not the only important criteria for the provision of strong repulsive forces and thus for colloidal stabilisation ability. As mentioned in section 3.1, the length of adsorbed fragments is another critical factor. This point can be illustrated more clearly using theoretical calculations for the induced inter-droplet interaction potentials between β-casein covered emulsion droplets and those involving adsorbed layers of an ideal short di-block fragment (see Fig. 3). The term ideal here is used to refer to a fragment that consists of only two consecutive purely hydrophilic and solely hydrophobic blocks [1].

At pH = 7, β–casein has a net charge of −6.14e. We make the charge density carried by our short fragment comparable with that for β–casein. This is done by setting the total charge q_{short} of the hypothetical fragment equal to $q_{short} = q_{\beta-casein}(N_{short}/N_{\beta-casein}) = (-6.14e)(15/209) = -0.44e$.

Fig. 3 shows the interaction potentials between two polymer-coated droplets, where in one case the polymer in question is β–casein and in the other the ideal short sized di-block fragment. Here the interaction potential is expressed in the unit of k_BT, and as mentioned in section 2.26, its value between two droplets is obtained from the potential calculated between two flat parallel plates, converted to that for spheres with the aid of equation (5). Decreasing values of the interaction potential with increasing separation indicate a repulsion between two droplets. Generally, flocculation happens due to a weak attraction between droplets when the minimum well in the interaction potential is



RELEELNVPGEIVESLSSSEESITRINKKIEKFQSEEQQQTEDELQDKIHP	1 - 51
FAQTQSLVYPFPGPIPNSLPQNIPPLTQTPVVVPPFLQPEVMGVSKVKEAM	52 - 102
APKHKEMPFPKYPVEPFTESQSLTLTDVENLHLPLPLLQSWMHQPHQPLPP	103 - 153
TVMFPPQSVLSLSQSKVLPVPQKAVPYPQRDMPIQAFLLYQEPVLGPVRGP	154 - 204
FPIIV	205 - 209

Fig. 2. The primary structure of β-casein with the constituent amino acids grouped into five different groups as indicated.

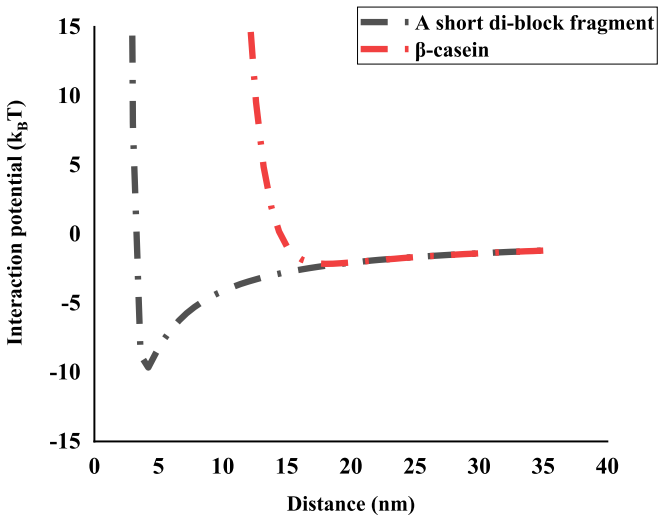


Fig. 3. Comparison of interaction potentials plotted vs separation distance, as obtained from SCF calculations, between two droplets covered with β–casein and with a short di-block fragment. The results also include direct van der Waals interactions operating between the droplets. The graphs were obtained at pH = 7, with the volume fraction of salt ions at 0.001 (~0.01 mol/l) and for droplets of size 1 μm.

more negative than around –5 k_BT. At less negative values, thermal agitation and Brownian motion is sufficient to separate two weakly aggregated droplets. Nevertheless, as the attraction increases, flocculation will further develop and cause stronger irreversible aggregates that can lead into coalescence.

In Fig. 3, apart from forces mediated by polymer layers we have also included the more direct van der Waals interactions. These were given by the following equation [14,15]

$$V_{vw} = -\frac{A_H R}{12r} \tag{8}$$

where A_H is the composite Hamaker constant for the dispersed phase in the continuous medium with a typical value of 1 k_BT for edible oils dispersed in water; R is the radius of the two spherical droplets and r is the distance between the droplets.

The primary structure of β-casein was taken from Farrell et al [45]. The short di-block fragment comprises of a hydrophobic domain with five hydrophobic residues and a hydrophilic section with 15 non-charged polar residues. Though not entirely impossible, this fragment is not very likely to arise from fragmentation of any common food proteins. However, our calculations aim to compare the interaction potential of these two di–block like macromolecules with quite different sizes (where β–casein consists of 209 amino acid residues).

As shown in Fig. 3, the minimum value of the interaction potential curve induced by adsorbed layers of β-casein is found to be around −2.0 k_BT. This occurs at a droplet separation distance ~ 18 nm. The interaction potential curve drops rapidly in the range 0 to 18 nm, but with the interaction potential still higher than zero at $r < 14.4$ nm. Therefore, according to these calculations, β-casein produces a repulsive effect at a suitably long range of separation distances. This explains why β-casein exhibits excellent emulsion stability properties, as also found in many previous theoretical studies [26,46], especially at pH values away from its isoelectric point. As for the curve obtained for the short di-block fragment, we found that the magnitude of the minimum well in the inter-droplet interaction potential was 9.64 k_BT. This is, much deeper than that obtained for β-casein. Moreover, the minimum value occurs at a substantially closer distance between the droplets, at $r = 4.2$ nm. In this case of stabilisation by short fragments, the van der Waals forces seem to dominate any repulsion produced by the overlap of polymer layers; thus, the emulsion prepared with our hypothetical ideal short di-block fragment is predicted not be colloidally stable.

From the above calculation, we can compare the surface behavior of two roughly similar "di-block" chains with different lengths. For β–casein we find a marked improvement in the provision of induced repulsive interactions as compared to the short di-block fragment. This result emphasizes the fact that any peptide-based emulsifier should contain fragments comprising of a sufficiently large number of amino acid residues, even when possessing the most favorable structure, i.e. a di-block–like fragment.

3.3. Surface tensions of soybean protein hydrolysates

The level of fall in the value of γ at the air–water (A−W) interface of a surfactant solution is also a good indicator of its surface activity at an oil–water (O−W) interface. Following hydrolysis and filtration etc. to obtain the various M_w fractions, because the amounts of samples available were limited, and since it is far easier to reliably measure the interfacial tension at an A−W rather than the O−W interface, we opted to measure γ at the A−W interface. A summary of the final values of γ (i.e., after 10 min – see methods section) is shown in Table 3, measured at 10^{-3} wt% protein to hopefully accentuate any differences between the samples, rather than the higher concentration used to prepare the emulsions (see later), where values tend to converge more at the short adsorption times relevant to emulsion formation. The values of γ given are the average values of 20 measurements every 6 s after 10 min, by which time plateau values appeared to have been reached. The Supplementary data gives further details on γ versus time, but since the *SCFC* cannot give kinetic but only equilibrium data, only the final values of γ are relevant here, to compare surface activity of the various hydrolysates and the intact proteins from which they were derived.

As observed in Table 3, none of the casein hydrolysates produced via alcalase- or pepsin-treatment gave a lower γ than the original (non-hydrolysed) casein solution (γ = 55 mN m^{-1}). Except for the pepsin-treated samples with a M_w of 3–10 kDa, which had a slightly reduced γ, the casein-based hydrolysates produced under the other conditions had a higher γ than the non-hydrolysed protein, i.e., were *less* surface active. This is probably because the hydrolysis destroyed the already close to ideal primary structure of the two main components of casein: β- and α_{s1}-casein (ideal in terms of block copolymer structure and surface activity). In other words, there is probably little advantage in hydrolysing the casein to smaller protein fragments, since casein is already an excellent emulsifier. In theory, fragments of the constituent proteins might be slightly better, but to produce them would require much more selective hydrolysis than used here and finding enzymes that could do this is quite challenging. Even if some suitable small fragments did exist in the samples (such in the 3–10 kDa fraction), the presence of higher concentrations of other less favorable fragments might kinetically

Table 3
Average surface tensions γ at 10^{-3} wt% protein and pH 7, of different *SPH* samples produced by different enzyme treatments. The 'whole' values refer to the measurements on whole hydrolysates, i.e., before fractionating into the 3 different M_w ranges. The corresponding γ for non-hydrolysed soy protein and casein were 71 and 55 mN m^{-1}, respectively.

	Different M_w fractions/kDa			
	<3	3 ~ 10	>10	whole
	Surface tensions (γ) of fractions/ mN m^{-1}			
Enzyme treatments	**Casein**			
Alcalase pH 8	63 ± 0.62	69 ± 0.1.36	70 ± 3.85	70 ± 2.38
Pepsin pH 1.3	59 ± 2.45	51 ± 3.07	61 ± 1.09	62 ± 2.77
Pepsin pH 2.1	72 ± 1.07	54 ± 3.95	68 ± 0.68	67 ± 1.02
	SPI			
Alcalase pH 8	65 ± 4.24	59 ± 5.72	56 ± 1.05	60 ± 2.33
Pepsin pH 1.3	49 ± 2.11	50 ± 1.34	58 ± 1.49	60 ± 1.16
Pepsin pH 2.1	69 ± 2.75	54 ± 2.03	66 ± 2.00	67 ± 1.64
Pepsin pH 4.7	–	–	52 ± 0.70	–

inhibit more surface active fragments from adsorbing [17].

In contrast to the casein, all the soybean protein hydrolysates had a lower γ (i.e., were more surface active) than the non-hydrolysed *SPI* (γ = 71 mN m^{-1}). Indeed, under the conditions of measurement used, the highly aggregated and relatively insoluble *SPI* seemed to be hardly surface active at all. The increase in surface activity upon hydrolysis must therefore be attributed to increased solubilisation and/or the production of smaller soy protein polypeptides that have greater tendency for adsorption, which therefore might be better emulsifiers. The γ results suggest that lower (<3 kDa) M_w fraction produced by pepsin treatment at pH 1.3, might be the best emulsifier, since this has the lowest γ (49 mN m^{-1}) of all the *SPH*, even lower than γ for casein (55 mN m^{-1}) under the same conditions. The use of pH 1.3 is significant because this at the optimum pH for the action of this enzyme [40,47]. For *SPI*, the lowest γ at pH 1.3 was closely followed by that of the medium M_w (3 – 10 kDa) *SPH* produced by pepsin at pH 1.3, which gave a similar γ (50 mN m^{-1}), again lower than the value for non-hydrolysed casein. Although these low γ values at this bulk concentration indicate strong adsorption, as explained earlier a low γ does not necessarily guarantee a good colloidal stabiliser. As well as being strongly adsorbed to the interface, the material must provide sufficient steric and/or electrostatic stabilisation, via sufficiently long chains protruding way from the interface into the aqueous phase, and/or chains that provide a sufficiently high charge density at the oil droplet surface. The more surface-active peptides might be too low in M_w to provide either, so further experiments were required to evaluate this, as reported below. In this respect, it is interesting that the high (>10 kDa) *Mw* range hydrolysate produced at pH 4.7, far from the pH optimum, produced a γ value only slightly higher (52 mN m^{-1}) than the above two fractions produced at pH 1.3. This is discussed further below.

3.4. Characterisation of selected SPHs and O/W emulsions stabilised by them

Emulsions were prepared using 9 different *SPHs* as in Table 3 involving pepsin treatment. This pH is further away from the optimum pH for pepsin action than pH 1.3 and was used to see if this might generate at least equally surface active fragments but perhaps even better stabilising material, i.e., higher M_w and fewer types of fragments. In fact, out of these 9 samples, only the *SPH* produced at pH values 2.1 and 4.7 were able to stabilise emulsions with any measurable stability at all – and in both cases when the M_w faction > 10 kDa was considered. Therefore, in the remaining characterisation of the *SPHs*, and the corresponding emulsions stabilised by them, we have focused on only those samples stabilised by these two *SPHs*, hereafter denoted as *P*2 and *P*4 for simplicity.

The degree of hydrolysis, D_H, for the *P*2 and *P*4 *SPHs* was calculated as 17.0 ± 4.6 and 11.1 ± 0.4%, respectively, substantiating the expectation that hydrolysis is less efficient at the higher pH, away from the optimum. Both values are still much lower than the theoretical degree of hydrolysis that can be achieved by pepsin (maximum 25 %, calculated by a tool PeptideCutter from Expasy) [48], if one only considers the peptide bonds that pepsin is traditionally considered capable of breaking. Similar conclusions were reached in a recent experimental study [28]. In fact, there are at least two factors that can lead to a lower D_H than expected. Firstly, the continuous hydrolysis of the protein by the enzyme leads to an increase in the pH of the solution, which then changes both the activity and the specificity of the enzyme; secondly, the native compact globular form of the soy protein means that some potential cleavage sites remain inaccessible to the enzyme.

Fig. 4 shows a comparison in appearance between 1 wt% dispersions of the *P*2 and *P*4 samples and the non-hydrolysed *SPI*. The *SPI* is quite turbid whilst the *P*4 sample preparation is almost completely transparent, whilst the turbidity of *P*2 solution is somewhere in between. This in itself suggested that the solubility of both hydrolysates might significantly be higher than that of the intact protein, but particularly for the pH 4.7 treatment with lower D_H.

Mastersizer measurements of these dispersions showed that the D_{32} of the non-hydrolyzed *SPI* was significantly reduced from 128.0 ± 21.0 to 4.4 ± 0.30 and 2.9 ± 0.24 μm for *P*2 and *P*4, respectively. This represents a significant decrease in mean particle size, whilst objects of these dimensions are clearly still not single soy protein molecules. However, these values are averages of distributions and the greater optical clarity of the *P*2 and *P*4 dispersions is a good indication that the majority of the proteinaceous material has a much smaller particle size than even these D_{32} values suggest, and therefore may be expected to give significantly superior performance in terms of emulsion droplet coverage compared to non-hydrolysed *SPI*.

Measurements of ζ–potential were carried out for *P*2 and *P*4 solutions using the method outlined in section 2.2.4, with a background electrolyte concentration of 10 mM NaCl and a pH of 7. Both samples exhibited a negative charge, but the ζ–potential of *P*4 (−40.3 ± 6.1 mV) was significantly higher in magnitude than that of *P*2 (−15.6 mV ± 2.4 mV). It must be emphasised that both samples will contain many different fragments, having varying charges, so these values are just average values of all these fragments. Nonetheless, the differences between the two samples are significant and suggest that all other things being equal, adsorption of all the *P*4 material might lead to a higher emulsion droplet surface charge than adsorption of all the *P*2 material. Of course, preferential adsorption by certain fractions within each material could alter this prediction entirely.

Emulsions were prepared using 1 wt% *P*2 and *P*4 as stabiliser according to the procedure described in section 2.2.2, and emulsion stability was monitored for up to 30 days of storage. Fig. 5 shows visual appearance of the emulsions and Fig. 6 shows the changes in the D_{32} value measured over the first 30 days.

Considering the freshly prepared (i.e., day 1) emulsions, it is clear that the emulsion stabilised by *P*4 has a more uniform milky-white appearance than that stabilised by *P*2. After two days the visual difference was even more marked: the *P*2 emulsion formed a distinctly separate cream and aqueous layers, whilst the *P*4 emulsion became slightly less cloudy at the bottom of the tube and with a thicker cream layer just starting to become visible. At longer times (e.g. > 1 week), the appearances started to converge, but the earlier observations did suggest slower creaming of the *P*4 emulsions and therefore probably smaller mean droplet sizes. It is also worth stating that if the *P*4 emulsion was

Fig. 4. Visual images (from left to right) of 1 wt% non-hydrolysed *SPI*, *P*2 and *P*4 protein suspensions.

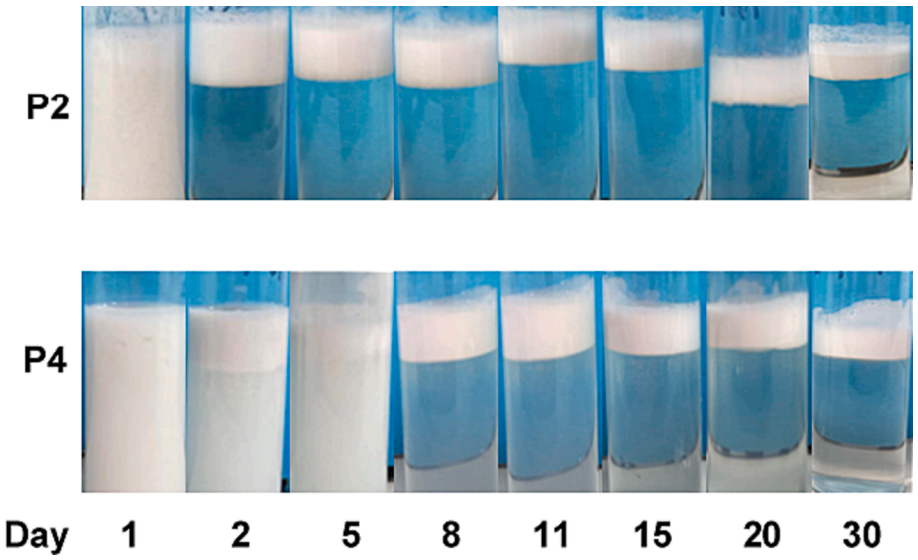


Fig. 5. Visual appearance of emulsions prepared using 1 wt% *P*2 or *P*4.

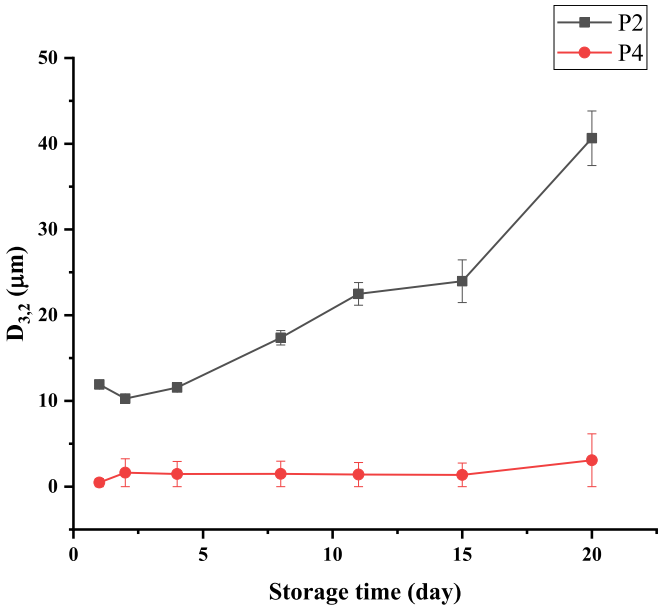


Fig. 6. The change of droplet size $D_{3,2}$ of O/W emulsions, stabilised by *P*2 and *P*4 hydrolysates, with storage time. Each measurement was repeated five times and the calculated standard deviation was used as the error bar. The samples were stored at pH 7 and a temperature of 4℃.

hand shaken at any time it easily reverted back to the appearance of the fresh sample, i.e., there was no strong irreversible droplet aggregation. From day 8 to day 60 (data not shown), the appearance of both emulsions remained largely unchanged, with no appreciable increase in the thickness of the cream layer.

That the *P*4 sample is better able to stabilise smaller droplets is confirmed in the *PSD* measurements shown in Fig. 6. The D_{32} value of the fresh *P*4 emulsion is lower (0.47 ± 0.0013 μm) than that of the *P*2 emulsion (11.92 ± 0.23 μm), but moreover, after a slight increase in the first 1 or 2 days (to 1.62 ± 0.051 μm) D_{32} was essentially constant until a slight increase was again observed between day 15 and day 20 (to 3.07 ± 0.036 μm). This might indicate the onset of some coalescence or flocculation, or the latter possibly leading to slight sampling errors if the samples are not shaken well enough before taking an aliquot for size measurements. In contrast, D_{32} of the *P*2 emulsion increases more or less continuously after day 5, reaching a value of 40.64 ± 3.43 μm after 20 days. Note that for these droplet sizes in a low viscosity medium as in this case (essentially water), creaming at this droplet volume faction (10 wt%) and droplet size would be expected to be relatively rapid if the droplets were not strongly flocculated. Rapid creaming and/or flocculation may in fact explain the apparent slight decrease in D_{32} for the *P*2 emulsion between day 1 and 4 (from 11.92 ± 0.23 μm to 11.56 ± 0.39 μm), which may be due to a slight sampling error if the larger droplets cream out of the measurement path during the Mastersizer measurements).

In summary, the *P*4 soy protein hydrolysates, obtained by pepsin treatment at the sub-optimal hydrolysis pH of 4.7 for the enzyme, exhibited significantly better emulsification and stability properties than those for *P*2 obtained at optimum pepsin hydrolysis pH (2.1).

3.5. Protein sequencing results

The emulsion studies suggested significant superiority of the *P*4 sample and that this might be due to the presence of fewer types and/or higher M_w peptides with more suitable amino acid sequences for emulsion stabilisation. Therefore, it was worthwhile investigating this fraction in more detail, to see if amino acid sequence information could be obtained. Gel electrophoresis results (Fig. 7) showed that the *P*4 sample was still very complex in composition, with a wide range of small fragments. Thus, bands F, G and H are faint and broad, but there was one band (E) at 25 kDa, that was relatively clear and more intense. Moreover, this M_w is in a range capable of imparting reasonable steric stabilisation – provided it adsorbs in the appropriate configuration. At the same time, the *P*4 sample was prepared by filtration through a 10 kDa ultrafiltration membrane, so that only a small number of fragments with a *Mw* of<10 kDa should be present. Sections 3.1 and 3.2 demonstrated the influence of the size (*Mw*) of the peptide on its emulsification properties. The results seem to agree with the predictions that below a lower M_w size limit the peptides are not effective stabilisers. Overall, the dominant 25 kDa band seemed a strong contender for the premier species imparting the good stabilising characteristics of the *P*4 hydrolysate. Hence, a full sequencing of the peptides in this band was attempted.

Unfortunately, the sequencing instrument/technique used can only measure fragments of lower size. Therefore, the already hydrolysed *P*4 sample was further broken down into shorter sequences via selective proteolysis, with the aim that the original sequence(s) could be reconstructed from the measured sequences, their overlaps and the known amino acid sequences of the constituent soy proteins. Trypsin,

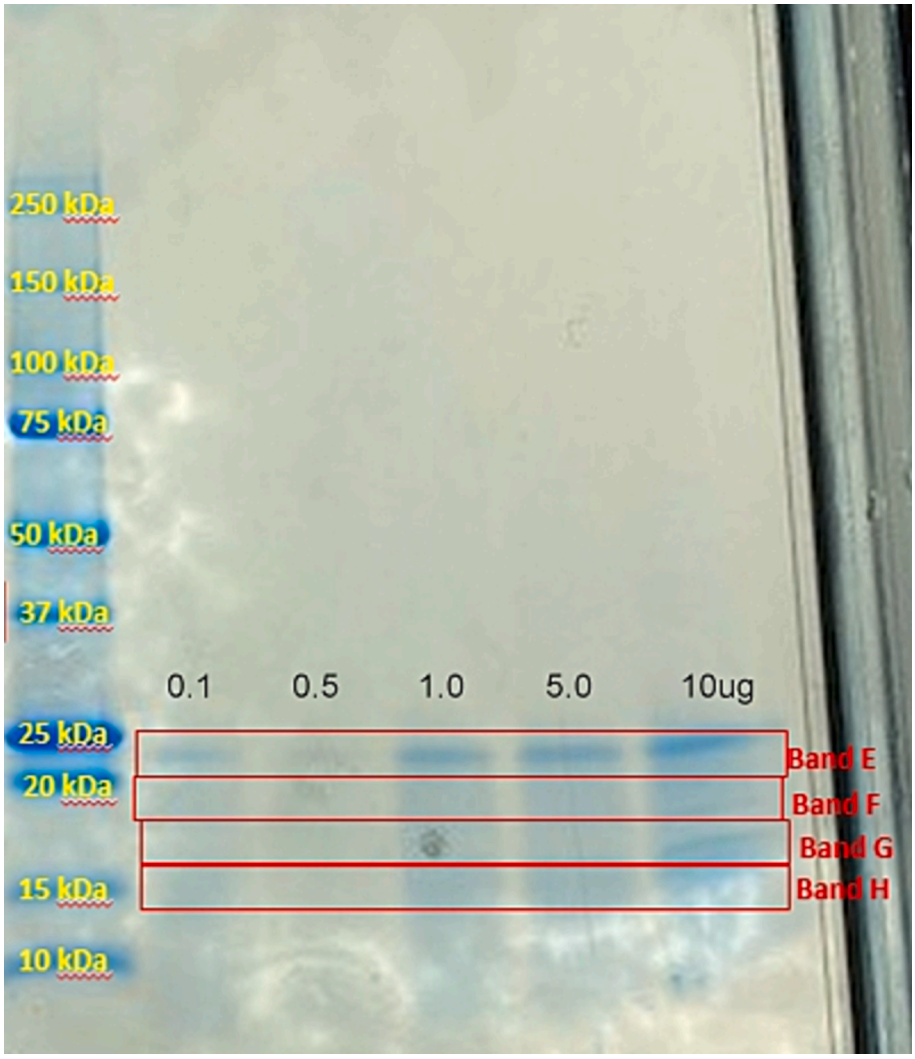


Fig. 7. SDS-PAGE pattern of the *P*4 hydrolysate.

chymotrypsin and glutamic proteases were therefore also used as described in the methods.

In the final analysis, the protein sequencing identified a total of 38 protein fragments from about six different proteins, including 24 fragments from the β-conglycinin alpha subunit, and the rest from glycinin (G4 and G5), β-conglycinin alpha' subunit, trypsin inhibitor (A and B) and kunitz trypsin inhibitor. Thus, the conclusion was that most of the detected fragments (see the blue lines in Fig. 8) originated from the β-conglycinin alpha subunit, which constitutes one of the main components of the soy protein. As mentioned above, it is difficult to measure the complete structure of all components in a band due to the limitations of the sequencing method. Predicted sequence 1 (PS1, the red line in Fig. 8) contains all of the detected fragments and has a $Mw =$ 12 kDa. This is obviously somewhat smaller than the supposed M_w of band E (25 kDa) but it should be remembered that there will be some error on this 25 KDa M_w due to the M_w markers in the standard not necessarily aligning perfectly with the sample gel.

Based on PS1, we can guess at sequences that have *Mw* closer to the

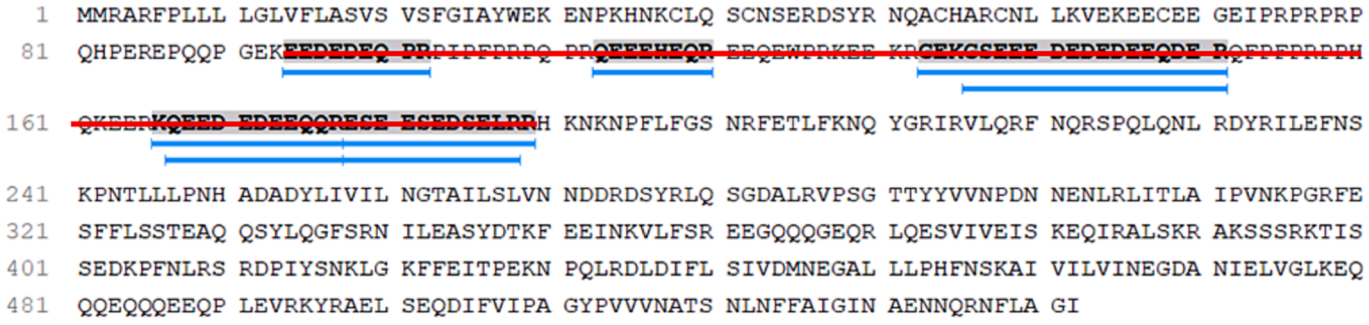


Fig. 8. Sequence results of the fragments detected in the *P*4 sample for the β-conglycinin alpha subunit, involving the hydrolysates in the 25 kDa band. The blue lines represent the sequences found in the sample that match this part of the subunit sequence; the red line represents the predicted sequence 1 (PS1) that contains all of the above but also the intervening parts of the sequence.

observed 25 kDa, by including appropriate extensions either side of the PS1 sequence. Obviously, there are a large number of possibilities, depending on what fraction of the 'missing' mass is allocated to the amino or carboxyl end of the PS1 chain. However, as a first attempt, we chose the 3 possibilities indicated in Fig. 9, denoted PS2, PS3 and PS4.

From PS1 and the 3 possible sequences PS2, PS3 and PS4, one can now start to calculate a variety of surface properties. Using the same SCF calculations as illustrated in section 3.2, one can now also predict the adsorbed configurations and colloidal interactions induced by these different peptides adsorbed layers of these fragments. All the calculations were carried out at pH = 7 to match the conditions in the experimental section. Some of the relevant properties can be found in Table 4.

As is seen in Fig. 10, the depth of the interaction potential minimum mediated by the PS1 fragment is predicted to be 4.25 k_BT. This occurs at a droplet separation distance of 8.4 nm. This is marginally deeper than the potential well obtained for the interaction potential produced by β–casein layers discussed in Section 3.2 (~2.14 k_BT). The shapes of the interaction potential induced by the PS2, PS3, and PS4 fragments, which all contain PS1 as part of their structure, were found to be broadly similar to PS1. Since pH 7 is far from the isoelectric point of the predicted sequences, all the systems can provide electrostatic, as well as potentially steric, repulsion. However, in order of increasing long-range repulsive character we can see that PS2 > PS4 ≫ PS3 ≫ PS1. PS1 has the highest magnitude of net negative charge but is predicted to have the lowest adsorbed amount (see Table 4) and the shortest range repulsive interaction. However, it should be remembered that PS1 is roughly half the M_w of the other 3 peptides and therefore might be expected to not extend as far away from the surface. PS3 has the lowest net charge, a similar adsorbed amount to PS2 and PS4, but gives a repulsion somewhere intermediate between PS1 and PS2, with PS2 ≈ PS4. The sample PS2 results in only slightly more repulsive force than PS4, whilst the net charge on it is marginally less negative than on PS4 (and their M_w are identical). Nonetheless, PS2 is predicted to have a larger adsorbed amount than PS4. These relative values indicate the overriding significance of M_w and adsorbed amount, whilst the effect of charge really depends on its distribution along the chains. These trends are further substantiated by the predicted configurations of the adsorbed peptides, shown in Fig. 11. It is indeed the case that PS2 and PS4 show much greater numbers of segments extending away for the surface than PS1 or PS3, whilst both PS2 and PS4 also have an anchoring section of residues at the surface. In fact, PS4 seems to have an almost 'perfect' diblock configuration of a closely adhering train of segments at the interface, attached to a long dangling tail extending away from the surface. In comparison, the lower M_w PS1 has an adsorbed configuration that extends not nearly so far away from the surface, whilst PS3 (M_w 24.5 kDa, very slightly lower than that of PS2 and PS4 at 25.0 kDa) is slightly better but forms more of a train–loop type configuration.

Remember that PS1 is the only fragment that was definitely identified as being present in some form in the fraction *P*4 that proved to be an effective emulsifier. However, if PS1 was actually present as part of the sequence in the higher M_w forms of PS4 or PS2, then the above calculations explain why *P*4 might indeed perform as an excellent emulsifier.

Taken all together then, the qualitative agreement between the behavior of these predicted fragments and that of the real hydrolysates is remarkable, even considering the discrepancy in the apparent M_w of band E and the M_w of the only fragments identified.

Table 4
Properties of the 4 possible peptides dominating the interface. Γ = adsorbed amount calculated from *SCFC* at pH 7; σ = calculated net charge on peptide at pH 7.

	PS1	PS2	PS3	PS4
Mw (kDa)	12.0	25.0	24.5	25.0
pI	4.4	4.9	5.0	4.8
σ/ e	−25.2	−19.7	−16.7	−20.4
$\Gamma \times 10^3$ /chains/a_0^2	0.338	8.44	4.24	6.98

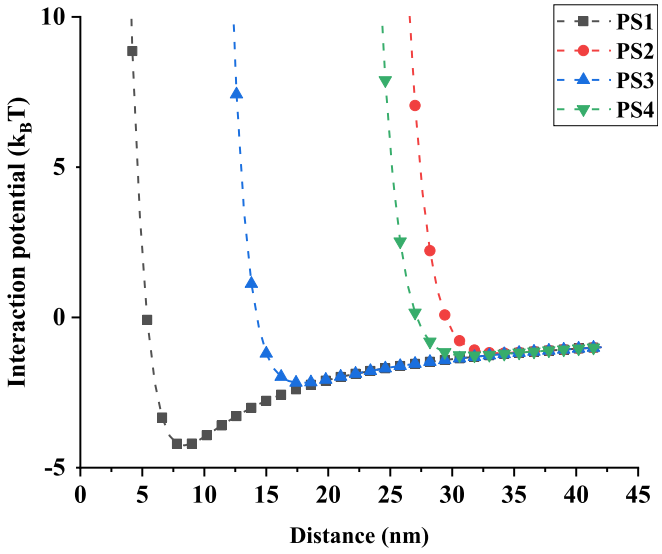


Fig. 10. Calculated colloidal interaction potentials induced between two droplets of size 1 μm, covered by adsorbed layers of each of the four possible predicted fragments (PS1 to PS4) present in the *P*4 hydrolysate sample. All results also included van der Waals interactions and were calculated at a pH = 7. The volume fraction of the background electrolyte was set at 0.001 (~0.01 mol/l).

4. Conclusions

Plant-based proteins tend to be storage proteins, with poor solubility,

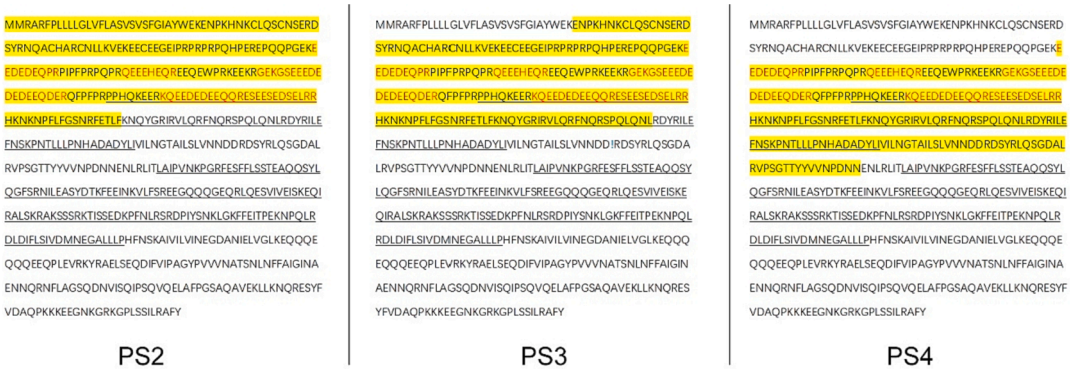


Fig. 9. Three predicted protein fragments that may be derived from beta-conglycinin and present in Band E of the *P*4 hydrolysate.

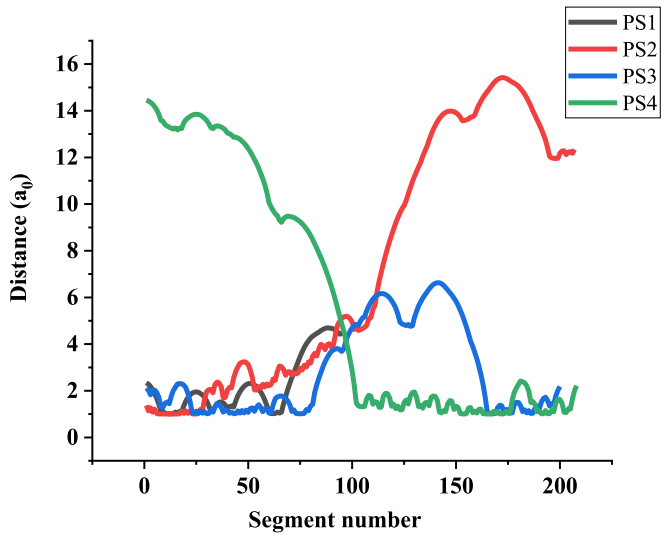


Fig. 11. Predicted configuration of four peptides at pH 7 upon adsorption, calculated by *SCFC*. The graphs show the average distance from the surface of each amino acid in the sequence of the polypeptide. The volume fraction of the background electrolyte was set to 0.001 (~0.01 mol/l) in all cases.

compact globular structures, slow surface adsorption kinetics and primary amino acid sequences that are far from ideal for use as emulsifier and colloid stabilisers [1]. To improve the solubility of such vegetable proteins and achieve better emulsion stabilising properties, many researchers have explored hydrolysis of protein as a useful strategy [5,18,21–23]. However, these studies suggest that to obtain an emulsion stabiliser anywhere as good as say the milk protein β-casein, let alone some of the synthetic macromolecular counterparts, one either must carefully control the level of hydrolysis, or else use a highly selective enzyme that only cleaves the chains at a few well defined and very specific places. Both of these requirements can be quite difficult and/or expensive to achieve. Indeed, the rapid change in the distribution of fragments upon hydrolysis by an indiscriminate enzyme, even with a marginal alteration in the degree of hydrolysis, may explain the frequently contrasting results reported in the literature. In fact, using self-consistent-field calculations, we have shown that indiscriminate proteolytic enzymes under their optimal conditions of action, are more likely to produce a range of fragments far too low in molecular weight to be of any use as stabilisers (even though they may be surface active). This can already be the case even when the degree of hydrolysis, (i.e. the percentage of peptides bonds cleaved) is relatively low ~10%. In the current study we have explored an alternative strategy of deliberately moving away from optimal conditions for the activity of enzymes, to make them both more selective and ensure that the degree of hydrolysis remains limited, even when not so rigorously controlled. The pepsinolysis of soy protein isolate (*SPI*) was used as an example to illustrate this approach. This leads to a range of higher M_w fragments that were found to have much superior surface activity and O/W emulsions stabilising properties compared to both the native non-hydrolysed protein, as well as the polypeptides generated under optimum pH conditions for pepsin activity. This can be explained by the M_w fraction > 10 kDa having peptide sequences, derived from β-conglycinin, that our *SCF* calculations predicted to have the right combination of hydrophobic and hydrophilic segments, giving both a good level of adsorption and strong inter-droplet repulsion forces. Therefore, the technique of considering enzymes that are commonly used to help solubilise plant proteins, but under non-optimal conditions, is demonstrated to be a feasible one for deriving suitable emulsifiers from plant-based proteins. In future work the method will be exploited further as a general technique for producing longer, more functional, emulsion stabilizing proteins (or indeed other biopolymers). It will be applied to a range of different plant materials using a variety of commonly available enzymes. This promising procedure has also the advantage that it can be largely guided by theoretical considerations, such as those presented here.

CRediT authorship contribution statement

Cuizhen Chen: Conceptualization, Visualization, Methodology, Software, Investigation, Formal analysis, Writing – original draft. **Brent S. Murray:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Rammile Ettelaie:** Conceptualization, Methodology, Software, Validation, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors are indebted to Dr. James Ault and Miss Rachel George of the Mass Spectroscopy Facility, of Faculty of Biological Science, University of Leeds, for separation and identification of the peptide sequences. They also thank Dr. Ruixian Han, for providing advice and expertise on the initial protein hydrolysis conditions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jcis.2023.08.040.

References

[1] B.S. Murray, R. Ettelaie, A. Sarkar, A.R. Mackie, E. Dickinson, The perfect hydrocolloid stabilizer: Imagination versus reality, Food Hydrocolloids 117 (2021), 106696.
[2] T.G. Burger, Y. Zhang, Recent progress in the utilization of pea protein as an emulsifier for food applications, Trends in Food Science & Technology 86 (2019) 25–33.
[3] M.P. Yadav, J.M. Igartuburu, Y. Yan, E.A. Nothnagel, Chemical investigation of the structural basis of the emulsifying activity of gum arabic, Food Hydrocolloids 21 (2) (2007) 297–308.
[4] S.J. Stubley, O.J. Cayre, B.S. Murray, I.C. Torres, Emulsifying properties of sugar beet pectin microgels, Food Hydrocolloids 137 (2023), 108291.
[5] Y. Ding, L. Chen, Y. Shi, M. Akhtar, J. Chen, R. Ettelaie, Emulsifying and emulsion stabilizing properties of soy protein hydrolysates, covalently bonded to polysaccharides: The impact of enzyme choice and the degree of hydrolysis, Food Hydrocolloids 113 (2021), 106519.
[6] C. Schmitt, S.L. Turgeon, Protein/polysaccharide complexes and coacervates in food systems, Advances in Colloid and Interface Science 167 (1) (2011) 63–70.
[7] T. Wu, C. Liu, X. Hu, Enzymatic synthesis, characterization and properties of the protein-polysaccharide conjugate: A review, Food Chemistry 372 (2022), 131332.
[8] M. Hattori, Functional improvements in food proteins in multiple aspects by conjugation with saccharides: case studies of β-lactoglobulin-acidic polysaccharides conjugates, Food Science Technology Research 8 (4) (2002) 291–299.
[9] K.-M. You, B.S. Murray, A. Sarkar, Tribology and rheology of water-in-water emulsions stabilized by whey protein microgels, Food Hydrocolloids 134 (2023), 108009.
[10] S. Soltanahmadi, B.S. Murray, A. Sarkar, Comparison of oral tribological performance of proteinaceous microgel systems with protein-polysaccharide combinations, Food Hydrocolloids 129 (2022), 107660.
[11] D. Bogahawaththa, N.H.B. Chau, J. Trivedi, M. Dissanayake, T. Vasiljevic, Impact of controlled shearing on solubility and heat stability of pea protein isolate dispersed in solutions with adjusted ionic strength, Food Research International 125 (2019), 108522.
[12] T.A. El-Adawy, Functional properties and nutritional quality of acetylated and succinylated mung bean protein isolate, Food Chemistry 70 (1) (2000) 83–91.
[13] J. Davis, D. Doucet, E. Foegeding, Foaming and interfacial properties of hydrolyzed β-lactoglobulin, Journal of Colloid Interface Science 288 (2) (2005) 412–422.
[14] R.J. Hunter. Foundations of colloid science, Oxford University Press, 2001.

[15] W.B. Russel, D.A. Saville, W.R. Schowalter. Colloidal dispersions, Cambridge University Press, 1991.
[16] G.J. Fleer, M.A. Cohen Stuart, J.M.H.M. Scheutjens, T. Cosgrove, B. Vincent. Polymers at interfaces, Chapman & Hall, 1993.
[17] R. Ettelaie, A. Zengin, H. Lee, Fragmented proteins as food emulsion stabilizers: A theoretical study, Biopolymers 101 (9) (2014) 945–958.
[18] Y. Wang, Z. Li, H. Li, C. Selomulya, Effect of hydrolysis on the emulsification and antioxidant properties of plant-sourced proteins, Current Opinion in Food Science 48 (2022), 100949.
[19] C. Liu, M. Bhattarai, K.S. Mikkonen, M. Heinonen, Effects of Enzymatic Hydrolysis of Fava Bean Protein Isolate by Alcalase on the Physical and Oxidative Stability of Oil-in-Water Emulsions, Journal of Agricultural and Food Chemistry 67 (23) (2019) 6625–6632.
[20] I. Celus, K. Brijs, J.A. Delcour, Enzymatic Hydrolysis of Brewers' Spent Grain Proteins and Technofunctional Properties of the Resulting Hydrolysates, Journal of Agricultural and Food Chemistry 55 (21) (2007) 8703–8710.
[21] J.A. Lopes-da-Silva, S.R. Monteiro, Gelling and emulsifying properties of soy protein hydrolysates in the presence of a neutral polysaccharide, Food Chemistry 294 (2019) 216–223.
[22] N.A. Avramenko, N.H. Low, M.T. Nickerson, The effects of limited enzymatic hydrolysis on the physicochemical and emulsifying properties of a lentil protein isolate, Food Research International 51 (1) (2013) 162–169.
[23] A. Nisov, D. Ercili-Cura, E. Nordlund, Limited hydrolysis of rice endosperm protein for improved techno-functional properties, Food Chemistry 302 (2020), 125274.
[24] R. Ettelaie, E. Dickinson, B.S. Murray, Self-consistent-field studies of mediated steric interactions in mixed protein+ polysaccharide solutions, in: E, Dickinson (Ed.), Food Colloids: Interactions, Microstructure and Processing (2005) 74–84.
[25] F.A.M. Leermakers, P.J. Atkinson, E. Dickinson, D.S. Horne, Self-consistent-field modeling of adsorbed β-casein: effects of pH and ionic strength on surface coverage and density profile, Journal of Colloid Interface Science 178 (2) (1996) 681–693.
[26] E. Dickinson, V.J. Pinfield, D.S. Horne, F.A.M Leermakers, Self-consistent-field modelling of adsorbed casein interaction between two protein-coated surfaces, Journal of the Chemical Society, Faraday Transactions 93 (9) (1997) 1785–1790.
[27] R. Ettelaie, A. Zengin, S.V. Lishchuk, Novel food grade dispersants: Review of recent progress, Current Opinion in Colloid, Interface Science 28 (2017) 46–55.
[28] R. Han, A.J. Hernández Álvarez, J. Maycock, B.S. Murray, C. Boesch, Comparison of alcalase- and pepsin-treated oilseed protein hydrolysates – Experimental validation of predicted antioxidant, antihypertensive and antidiabetic properties, Current Research in Food Science 4 (2021) 141–149.
[29] I. Burgaud, E. Dickinson, P. Nelson, An improved high-pressure homogenizer for making fine emulsions on a small scale, International journal of food science technology 25 (1) (1990) 39–46.
[30] A.K. Dolan, S.F. Edwards, Theory of the stabilization of colloids by adsorbed polymer, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences 337 (1975) 509–516.
[31] J.M.H.M. Scheutjens, G.J. Fleer, Statistical theory of the adsorption of interacting chain molecules. 1. Partition function, segment density distribution, and adsorption isotherms, The Journal of Physical Chemistry 83 (12) (1979) 1619–1635.
[32] O.A. Evers, J.M.H.M. Scheutjens, G.J. Fleer, Statistical thermodynamics of block copolymer adsorption. 1. Formulation of the model and results for the adsorbed layer structure, Macromolecules 23 (25) (1990) 5221–5233.
[33] R. Ettelaie, A. Akinshina, E. Dickinson, Mixed protein–polysaccharide interfacial layers: a self consistent field calculation study, Faraday discussions 139 (2008) 161–178.
[34] R. Ettelaie, N. Khandelwal, R. Wilkinson, Interactions between casein layers adsorbed on hydrophobic surfaces from self consistent field theory: κ-casein versus para-κ-casein, Food Hydrocolloids 34 (2014) 236–246.
[35] A.C. Balazs, C. Singh, E. Zhulina, Modeling the interactions between polymers and clay surfaces through self-consistent field theory, Macromolecules 31 (23) (1998) 8370–8381.
[36] T. Cosgrove, T. Heath, B. Van Lent, F. Leermakers, J. Scheutjens, Configuration of terminally attached chains at the solid/solvent interface: self-consistent field theory and a Monte Carlo model, Macromolecules 20 (7) (1987) 1692–1696.
[37] D. Irvine, A. Mayes, L. Griffith-Cima, Self-consistent field analysis of grafted star polymers, Macromolecules 29 (18) (1996) 6037–6043.
[38] C.R. O'Melia, Chapter 18 - Fundamentals of particle stability, in: G. Newcombe, D. Dixon (Eds.), Interface Science and Technology, Elsevier, 2006, pp. 317–362.
[39] E. Dickinson. Introduction to Food Colloids, Oxford University Press, 1992.
[40] D. Piper, B.H. Fenton, pH stability and activity curves of pepsin with special reference to their clinical importance, Gut 6 (5) (1965) 506.
[41] E. Dickinson, M.G. Semenova, A.S. Antipova, Salt stability of casein emulsions, Food Hydrocolloids 12 (2) (1998) 227–235.
[42] E. Dickinson, Properties of emulsions stabilized with milk proteins: overview of some recent developments, Journal of Dairy Science 80 (10) (1997) 2607–2619.
[43] E. Dickinson, Milk protein interfacial layers and the relationship to emulsion stability and rheology, Colloids and Surfaces B: Biointerfaces 20 (3) (2001) 197–210.
[44] E. Dickinson, Interfacial, Emulsifying and Foaming Properties of Milk Proteins, in: P.F. Fox, P.L.H. McSweeney (Eds.), Advanced Dairy Chemistry—1 Proteins: Part A / Part B, Springer, US, Boston, MA, 2003, pp. 1229–1260.
[45] H.M. Farrell, R. Jimenez-Flores, G.T. Bleck, E.M. Brown, J.E. Butler, L.K. Creamer, C.L. Hicks, C.M. Hollar, K.F. Ng-Kwai-Hang, H.E. Swaisgood, Nomenclature of the Proteins of Cows' Milk—Sixth Revision, Journal of Dairy Science 87 (6) (2004) 1641–1674.
[46] E. Dickinson, Structure and composition of adsorbed protein layers and the relationship to emulsion stability, Journal of the Chemical Society, Faraday Transactions 88 (20) (1992) 2973–2983.
[47] B. Keil, Essential substrate residues for action of endopeptidases, Springer, Specificity of proteolysis, 1992, pp. 43–228.
[48] PeptideCutter, ExPASy, SIB Swiss Institute of Bioinformatics, 2023. https://www.expasy.org/resources/peptidecutter. (Accessed 11 August 2022).