Recent advances in the production of emulsifying peptides with the aid of proteomics and bioinformatics

Moreno, Pedro Jesús García; Yesiltas, Betül; Echers, Simon Gregersen; Marcatili, Paolo; Overgaard, Michael T.; Hansen, Egon B.; Jacobsen, Charlotte

Published in:
Current Opinion in Food Science

DOI (link to publication from Publisher):
10.1016/j.cofs.2023.101039

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Food Science

Recent advances in the production of emulsifying peptides with the aid of proteomics and bioinformatics

Pedro J García-Moreno[1], Betül Yesiltas[2], Simon Gregersen Echers[3], Paolo Marcatili[4], Michael T Overgaard[3], Egon B Hansen[2] and Charlotte Jacobsen[2]

Food industry aims to develop novel protein-based emulsifiers from sustainable sources (e.g. plants, seaweed/microalgae, microbial, and insects) to satisfy the clean-label demand by consumers. Enzymatic hydrolysis releases peptides with enhanced surface properties compared with the parent alternative proteins. Traditionally, a trial-and-error top-down approach, which requires extensive costs in screening analyses, has been carried out to produce emulsifying peptides. This review presents the recent advances in a novel and fundamentally orthogonal bottom-up strategy, facilitated by quantitative proteomics and bioinformatic functional prediction, to produce emulsifying peptides by targeted enzymatic hydrolysis based on *in silico* proteolysis. Moreover, new insights on the relation between interfacial properties of peptides and emulsifying activity, as well as impact on stability of wet and dried emulsions, are discussed.

Addresses
[1] Department of Chemical Engineering, University of Granada, Spain
[2] National Food Institute, Technical University of Denmark, Denmark
[3] Department of Chemistry and Bioscience, Aalborg University, Denmark
[4] Department of Health Technologies, Technical University of Denmark, Denmark

Corresponding author: García-Moreno, Pedro J (pjgarcia@ugr.es)

Current Opinion in Food Science 2023, **51**:101039

This review comes from a themed issue on **Foodomics Technologies**

Edited by **Alberto Valdés**

For complete overview of the section, please refer to the article collection, "Foodomics Technologies 2023"

Available online 24 April 2023

https://doi.org/10.1016/j.cofs.2023.101039

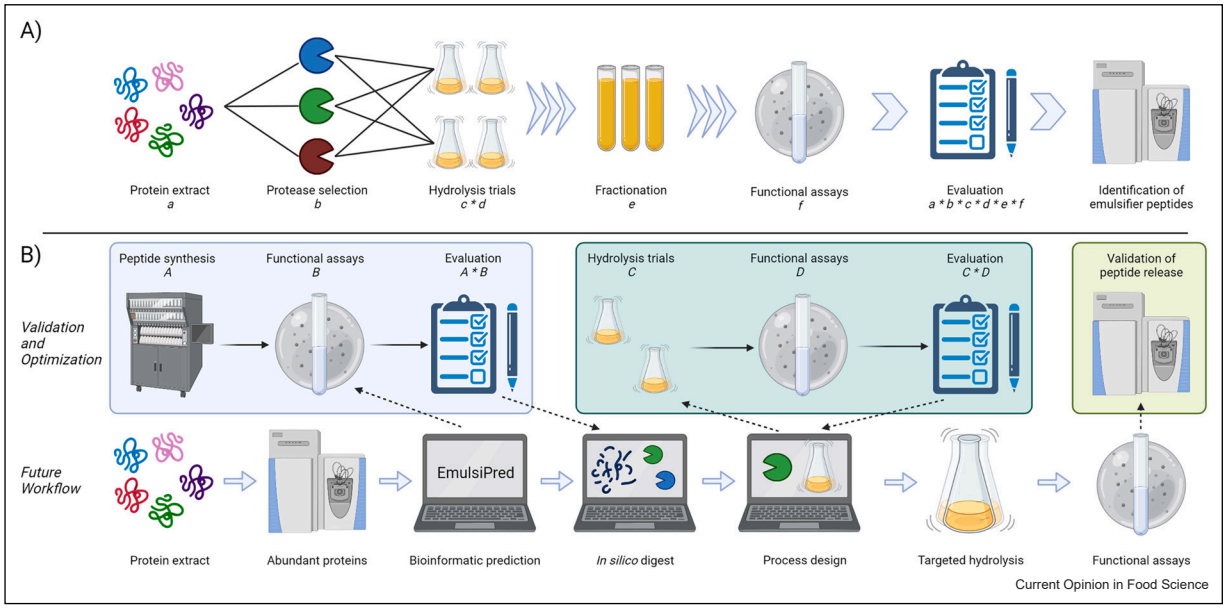
Introduction

Many traditional foods are colloidal dispersions, which are present either in the form of oil-in-water emulsions (i.e. milk, mayonnaise, sauces, and dressings), including complex oil-in-water emulsions (i.e. ice cream, sausages), or in the form of water-in-oil emulsion (i.e. butter, margarine) [1]. Furthermore, wet and dried emulsions are commonly used systems to control the delivery and bioavailability of bioactive ingredients (e.g. omega-3 fatty acids, vitamins, carotenoids, curcuminoids, and others) present in fortified foodstuffs [2]. Emulsions, consisting of at least two immiscible liquid phases and with droplet size ranging from 10 nm to 100 μm, are thermodynamically unstable systems, which achieve kinetic stability by the presence of an interfacial layer of emulsifier(s) surrounding the droplets of the dispersed phase(s) [3].

Emulsifiers locate at the interfacial area formed during emulsification, decreasing interfacial tension and forming a physical barrier between the immiscible phases that provide stability against droplet coalescence [1,3]. Protein-based emulsifiers are excellent stabilizers of emulsions by providing steric hindrance and/or electrostatic repulsions [4]. Indeed, there is a clear trend in the food industry to use emulsifiers derived from sustainable proteins (e.g. plants, microbial, seaweed/microalgae, and insect) in order to replace currently used synthetic surfactants (e.g. sucrose and sorbitan esters) and animal proteins (e.g. dairy proteins) [5–7]. Nevertheless, sustainable proteins are usually more complex (e.g. a mixture of proteins), as well as less soluble and surface active compared with surfactants and animal proteins and thus show limitations to their direct use as emulsifying agents [8–10]. On the other hand, peptides with superior surface properties can be released from these alternative proteins. Peptides have smaller size and higher exposure of both hydrophobic and polar groups, which might result in higher aqueous solubility, diffusivity, and amphiphilic potential compared with the parent proteins. The latter enhances the adsorption of peptides at the interface. Moreover, upon adsorption at the interface, peptides may change their secondary structure. Together, these properties allow the stabilization of a large interfacial area [11–14].

This review presents the recent advances on the production of emulsifying peptides from sustainable proteins. Particularly, it discusses the advantages of using proteomics and bioinformatics for the identification and

Figure 1



Schematic representation of the traditional top-down/trial-and-error and the new bottom-up/bioinformatic/data-centric approach for production of emulsifying hydrolysates and identification of peptide emulsifiers. Created with Biorender.com. **(A)** In the traditional approach, one (or more) protein extracts and/or raw materials (a) are hydrolyzed by a selection of proteases (b) using different process parameters (e.g. temperature, pH, E/S ratio, time, and degree of hydrolysis) at different levels (c * d). Following potential fractionation of hydrolysates (e), hydrolysates are investigated by a number of functional assays (f), resulting in a large number of wet-lab experiments (a * b * c * d * e * f) to be evaluated. The best hydrolysates/fractions may then be investigated (potentially following further fractionation) by LC–MS/MS to identify peptides that may be responsible for the observed emulsification. **(B)** In the bottom-up approach, one (or more) protein extracts and/or raw materials are characterized by quantitative LC–MS/MS proteomics to identify abundant proteins. The sequences of these proteins are then analyzed computationally (EmulsiPred) to predict probable emulsifying peptides embedded in the proteins. A number of peptides (A) are then synthesized and investigated by a number of functional assays (B) to evaluate their emulsifying properties (A * B). Based on performance and protein abundance, a number of peptides is selected as primary targets and their release potential by a range of proteases is investigated by *in silico* proteolysis. Based on this, hydrolysis process can be designed. The process may then be optimized through a range of hydrolysis trials (C) followed by investigation with functional assays (D) to evaluate the properties of the hydrolysates (C * D), leading to a targeted and scalable hydrolysis. Ultimately, release of targeted peptides may be validated by LC–MS/MS peptidomic analysis. In the future, a substantial number of wet-lab experiments ('Validation and optimization') may be bypassed, allowing for an approach ('Future workflow') relying only on quantitative proteomics and bioinformatic analysis to facilitate design of the targeted hydrolysis.

targeted release of peptides with high emulsifying activity. Furthermore, new insights on the relation between interfacial properties of peptides and emulsifying activity, as well as impact on the physicochemical stability of wet and dried emulsions, will be provided.

Approaches to identify and obtain emulsifying peptides

The traditional, and still commonly used, top-down approach to obtain emulsifier protein hydrolysates/peptides consists of a trial-and-error process where enzymatic hydrolysis is performed by several proteases, either added individually or in combination, to obtain a range of protein hydrolysates with different degrees of hydrolysis. Then, the hydrolysate could be used directly, or fractionated before the emulsifying activity is assessed [15–20]. In rare cases, the emulsifying peptides are then identified from the most active fraction [21] (Figure 1A). Recently, a bottom-up approach combining proteomics and bioinformatics has been developed, which saves time and costs of screening analyses [22–24] (Figure 1B). Proteomics is used to quantitatively determine protein composition in the raw material and subsequently, emulsifying peptides are bioinformatically predicted by analyzing the sequences of the most abundant proteins. Following functional validation of individual peptides, a targeted enzymatic hydrolysis process is designed to release the most abundant and potent emulsifying peptides from the raw material. The advances in the different stages of this bottom-up approach as well as future perspectives are further discussed in the sections below.

Proteomic approaches for identifying abundant proteins

Identification of the abundant proteins is essential to maximize potential downstream yields and ensuring that targeting specific peptides represents a viable business case. To this end, application of bottom-up proteomics by liquid chromatography–tandem mass spectrometry (LC–MS/MS) can provide both qualitative and

quantitative insights on the protein composition for an unprocessed protein source or a protein isolate/extract [25]. A fundamental prerequisite for performing a meaningful proteomics analysis is the availability of a suitable reference protein database. UniProt, an open-source knowledgebase, features > 189 million protein sequence records and > 292 000 proteomes (i.e. the complete set of proteins believed to be expressed by an organism), from where reference protein databases can be directly downloaded [26]. UniProt proteomes were recently used for quantitative analysis of a methanotrophic ferment [27] as well as side streams from the potato [23] and codfish [28,29] industry, as the basis for prediction of embedded emulsifying peptides.

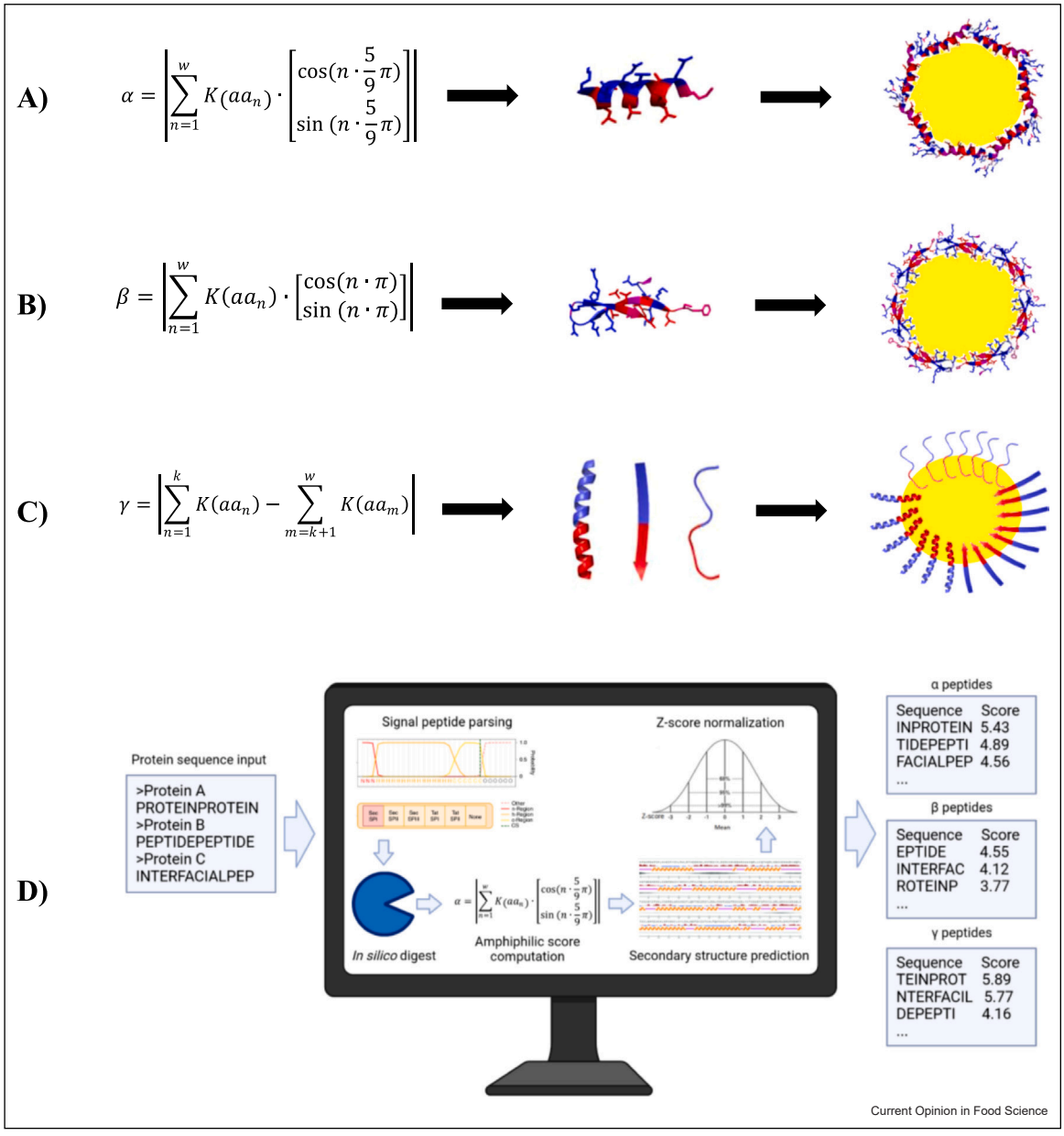
Nonetheless, the UniProt database is still far from complete to describe the plethora of organisms on earth. When dealing with samples where no proteome is available for database construction, it is therefore required to construct a custom database [30]. This can be achieved by imputation from related organisms within the same taxonomic family or order, under the assumption that close phylogenetic relation implies a large degree of protein sequence conservation. Such an approach was recently applied to quantitatively characterize a seaweed protein extract, used as the basis for prediction of embedded emulsifying peptides [31]. Alternatively, it may be possible to access draft genomic or even better transcriptomic data, which can be bioinformatically processed, assembled, and translated into protein sequences. Advances in genomic and transcriptomic analysis, improving throughput and quality while reducing costs substantially, represent an excellent base for obtaining reference protein databases. While using, for example, *de novo* transcriptome assembly may introduce challenges related to annotation of, for example, protein-coding regions and *in vivo* protein function, it can ultimately provide the protein sequences required for constructing a database [30]. The latter approach was recently applied to quantitatively characterize the protein composition of a seaweed protein extract [32], and used for subsequent prediction of embedded emulsifying peptides [27].

Another approach for identifying abundant proteins is quantitative visualization of the proteomic composition by electrophoresis (sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)) followed by excision of intense bands and subsequent in-gel digestion and proteomic analysis [33]. While insight on sample complexity and quantitative composition is lost with this approach, it may be useful for *de novo* protein sequencing, when very little a priori knowledge (e.g. suitable reference protein database) exists [34]. Ultimately, this may allow to imputatively identify a specific protein from related organisms through open and/or metaproteomic analysis [34]. Electrophoresis may also be used for indirect protein fractionation, where the sample lane is excised in larger pieces. While this should be done with caution and with risk of loss/bias due to additional handling, the approach was successfully applied for protein extracts from, for example, potato [23], bacterial biomass [27], and seaweeds [32,35].

Bioinformatics to predict emulsifying peptides

Peptides are by nature complex, and the combinatorial space increases exponentially with length, n, as 20^n, only taking into account the twenty natural amino acids (AAs). Not only do AAs differ in side-chain functional groups and thereby physicochemical properties, but also in their propensities to form secondary structures, adding an additional layer of complexity in the understanding of peptide functionality. Several studies on the design of *de novo* emulsifying peptides, which include experimental validation, reported that soluble peptides with an amphiphilic α-helix or β-sheet structure exhibit emulsifying properties [36–38]. Previously, Eisenberg et al. [39] determined the amphiphilicity of a helix by calculating the mean helical hydrophobic moment and indicated that this principle can be generalized to amphiphilic peptides in any conformation. Bearing this in mind, we developed the open source, Python-based predictor EmulsiPred (https://github.com/MarcatiliLab/EmulsiPred) [22,23], which allows identification of potential emulsifying peptides from the full-length sequence of proteins. EmulsiPred determines an amphiphilic vector for all embedded peptides in each of three potential conformations at the interface: i) α-helix, showing facial amphiphilicity (i.e. helix with hydrophobic and hydrophilic faces), ii) β-strand, also exhibiting facial amphiphilicity (i.e. hydrophobic and hydrophilic AA side chains oriented alternatively above and below the plane of the β-strand), or iii) γ-peptides, having axial amphiphilicity (e.g. peptides with any secondary structure, including unordered peptides, that have a hydrophobic N-terminal domain and a hydrophilic C-terminal domain or vice versa) (Figure 2a–c). It is worth noting that, for α and β scores, the predictor considers only peptides capable of adopting the specific structure at the interface by determining this property through protein-level secondary structure propensity prediction with NETSURFP-2.0 [40]. Moreover, EmulsiPred parses signal peptides to only consider the mature form of submitted proteins using SignalP [41] and additionally features a z-score normalization to avoid bias in the amphiphilic score imposed by peptide length (Figure 2d). EmulsiPred has already been used to successfully predict emulsifying peptides from different sustainable protein sources such as potato, microbial ferments, and seaweed [22,23,27,31].

Figure 2



Equations used by EmulsiPred to calculate amphiphilic scores of peptides and the EmulsiPred computational workflow. Determination of amphiphilicity and emulsifying mode of action for **(a)** α-helix peptides (facial amphiphilicity), **(b)** β-strand peptides (facial amphiphilicity), and **(c)** γ-peptides (axial amphiphilicity), comprising α-helix, β-strand or unordered structure. **(d)** Schematic representation of the EmulsiPred workflow, where a protein sequence input (.fasta format) is initially parsed for predicted signal peptides and then subjected to *in silico* digestions (all possible peptides within a defined length range), before calculating the amphiphilic score in all modes (α, β, and γ). If peptides are unlikely to adopt a specific conformation based on secondary structure prediction, peptides are rejected (score = 0). Last, scores are normalized using Z-score (at all peptide lengths) before output files are generated containing lists of predicted α, β, and γ peptides (above a scoring threshold). Created with Biorender.com. α-helix and β-strand peptides adsorb horizontally, whereas γ-peptides adsorb perpendicularly at the oil/water interface. Hydrophobic regions (red) and hydrophilic regions (blue) project to the oil and water phases, respectively. *K(aa$_n$)* is the Kyte–Doolittle score of *aa$_n$*, *n* and *m* represent the position of the AA in the peptide sequence, *w* is the length of the peptide, and *k* is the position of the AA that separates the hydrophobic and hydrophilic parts of the peptide (γ peptides only).
(a), (b) and (c) (adapted from [23]).

Validation of emulsifying activity: relation between interfacial properties and functionality

After potential emulsifying peptides have been identified by EmulsiPred, validation of their emulsifying activity is possible. García-Moreno et al. [22,23] and Yesiltas et al. [27] evaluated the emulsifying properties of ~70 synthetic peptides (7–30 AAs), which were embedded in abundant proteins from potato, seaweed, or microbial biomass and predicted by EmulsiPred to exhibit emulsifying activity. In agreement with Enser et al. [36], we found that predicted α-helical peptides required a certain length (> 18 AAs) to efficiently adsorb at the oil/water interface, reduce interfacial tension, and form physically stable emulsions [23,27]. On the other hand, and similarly to other studies [37,42], middle-length (13–16 AAs) β-strand peptides, with lower tendency to form aggregates and thus with higher solubility and diffusivity, were generally found to exhibit improved emulsifying activity [23,27]. Indeed, the predicted α-helix or β-strand conformations, when adsorbed at the oil/water interface, were generally confirmed by synchrotron radiation circular dichroism (SRCD) for selected potato peptides (Figure 3a,b) [12]. Although γ-peptides were predicted to orient perpendicularly at the oil/water interface, SRCD results indicated that these peptides adopt a more well-defined interfacial conformation. For instance, we identified three γ-peptides embedded in protein side streams from potato starch production (γ-1 and its variants γ-75 and γ-76), which adopted a predominantly α-helical conformation, whereas another peptide from the same raw material (γ-36) adopted predominantly β-strand conformation, while others adopted predominantly unordered structure, but with a substantial content of β-strand (γ-38 and γ-40) (Figure 3b). These findings denote that parallel adsorption of γ-peptides at the interface appears favored through the binding between hydrophobic patches of peptides and the oil phase [12]. Moreover, as these peptides showed strong emulsifying properties, it further indicates that a combination of axial and facial amphiphilicity may indeed be beneficial to produce physically stable emulsions.

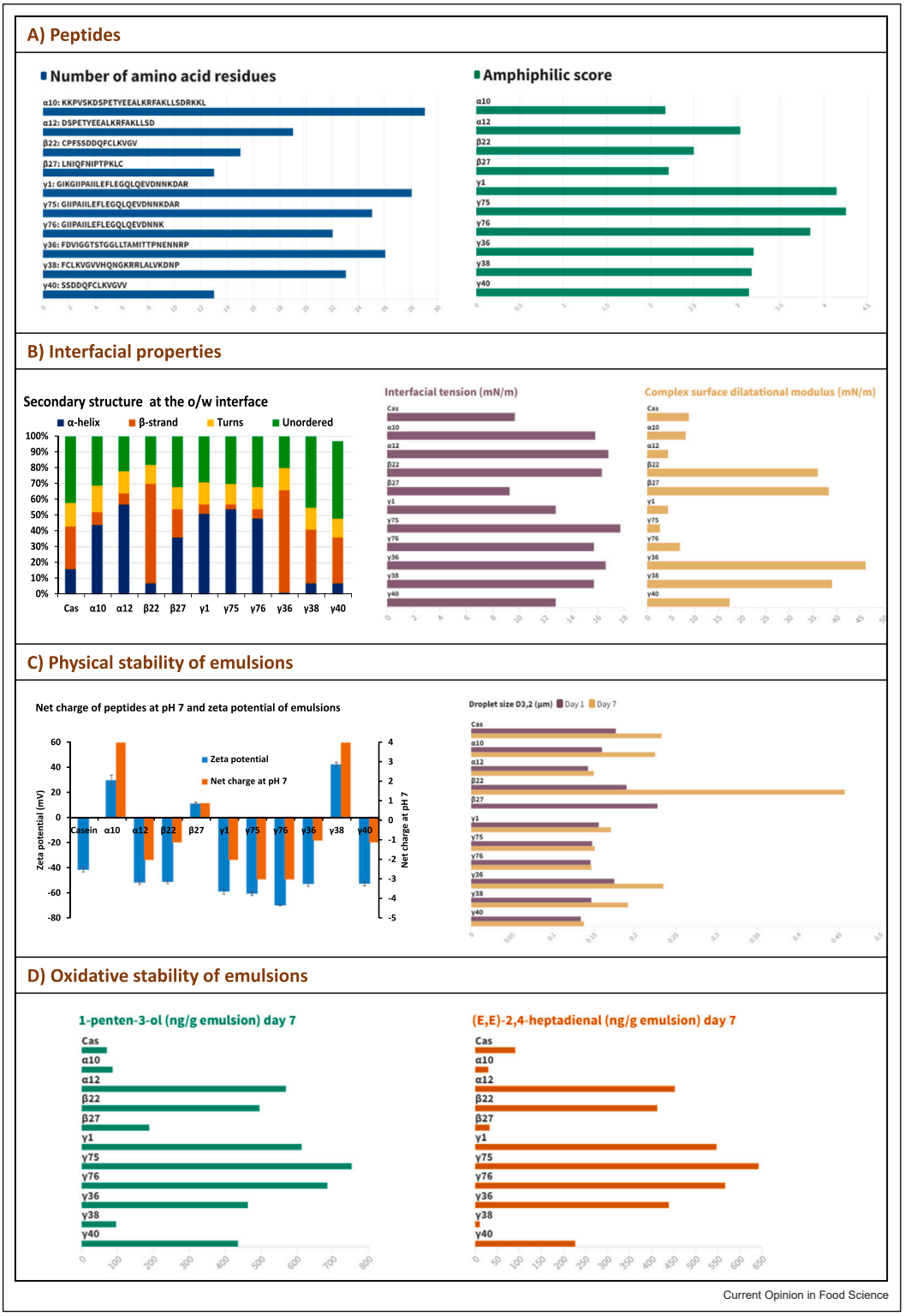
Potato peptides with a predominantly α-helical conformation at the oil–water interface (α-10, α-12, γ-1, γ-75, and γ-76) result in less stiff, weak, and more stretchable interfaces (e.g. low complex surface dilatational modulus), providing physical stabilization to emulsions mainly through steric hindrance and electrostatic repulsions (Figure 3b,c). Contrarily, potato peptides comprising mainly β-strand structure at the oil–water interface (β-22, γ-36, and γ-38) show significant in-plane attractive interactions at the interface, leading to stiff, solid-like viscoelastic peptide layers (Figure 3b) [12]. The latter might suggest further rearrangement and interpeptide interaction with potential formation of intermolecular β-sheets or β-sheet stacking [5]. Undoubtedly, together with physical stability, the oxidative stability of oil-in-water emulsions stabilized with peptides is of utmost importance. Curiously, recent results [12] indicated that positively charged peptides, which repel cationic metal from the interface, conferred superior oxidative stability to emulsions (e.g. where oxidation is favored by cationic metal ions) when compared with negatively charged peptides (Figure 3c,d). Nevertheless, peptide surface charge does not determine the oxidative stability of fish oil-loaded microcapsules (e.g. dried emulsions) where peptides are used as emulsifiers. This finding is attributed to the fact that oxidative stability of microcapsules is mainly determined by oxygen permeability (e.g. due to the reduced catalytic effect of metal ions at low water activity) [43].

Altogether, these recent studies prove the feasibility of using proteomics and bioinformatics to identify the sequence of potent emulsifying peptides and allowed us to gain new insights into the relation between interfacial properties of peptides and their functionality.

Targeted enzymatic hydrolysis

The release of predicted and verified emulsifying peptides from sustainable proteins will ultimately require specific hydrolysis at both ends of the identified segment. Controlled enzymatic release of the emulsifying peptides requires the availability of proteases with the desired specificity and devoid of activities leading to excessive degradation. For a range of different peptide bioactivities, this can be accomplished by using toolboxes such as those provided in, for example, BIOPEP-UWM [44], however, there are currently no integrated workflows to accommodate prediction and *in silico* proteolysis for functional properties such as emulsifying activity. In such scenarios, *de novo* prediction by EmulsiPred and *in silico* proteolysis must be combined by the user (Figure 4). A straightforward approach is to use tools based on known substrate sequence specificity of commercially available enzymes. Tools such as Peptide Cutter (available at https://web.expasy.org/peptide_cutter) can be used for this purpose. We recently demonstrated how such a workflow allowed for designing a targeted hydrolysis strategy [24]. By application of the high-specificity protease trypsin, selected based on *in silico* digestion analysis, we produced a potato protein hydrolysate, which contained targeted peptides. Moreover, the hydrolysate exhibited superior interfacial properties compared with hydrolysates produced with various industrial proteases to different degrees of hydrolysis. In another recent study on hydrolysates from a seaweed protein extract, it was shown that there was correlation between the hydrolysate-emulsifying properties and the proportion of peptides predicted to be potential emulsifiers (results not published). However,

Figure 3

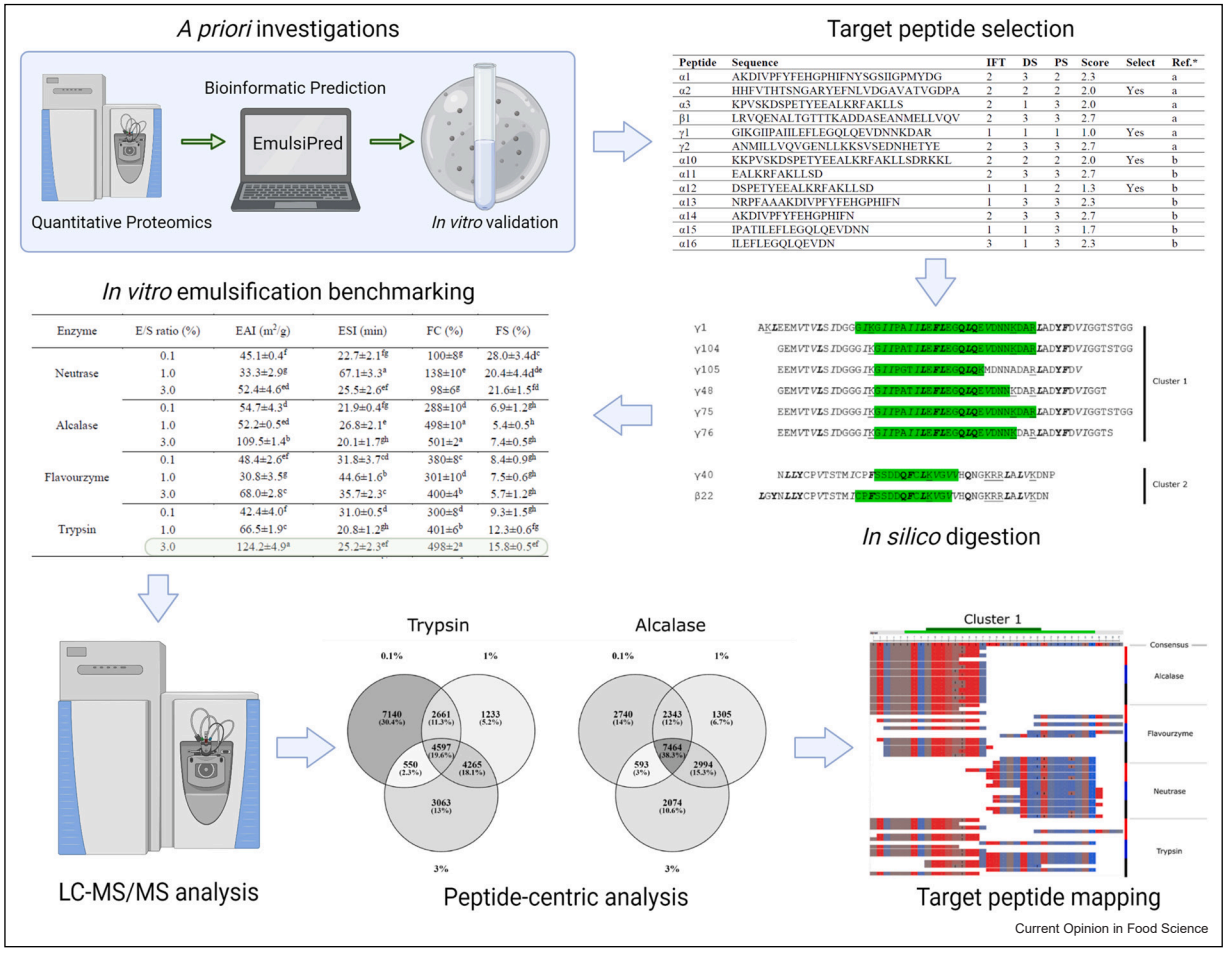


Current Opinion in Food Science

Examples of emulsifying peptides, their interfacial properties and the physical and oxidative stabilization of oil-in-water emulsions. (a) Selected potato-emulsifying peptides identified by proteomics and bioinformatics, **(b)** interfacial properties of the selected peptides, **(c)** physical stability, and **(d)** oxidative stability of 5% fish oil-in-water emulsions (pH 7) stabilized with the peptides. Data were obtained from [12]. [1] Amphiphilic score was obtained by the bioinformatics tool EmulsiPred (https://github.com/MarcatiliLab/EmulsiPred). [2] Secondary structure at the oil/water interface was determined by SRCD. [3] Interfacial tension values for potato peptide solutions of 0.1 g/L (pH 7) at the tricaprylin oil–water interface after 3 h. [4] Complex surface dilatational modulus for tricaprylin oil–water interfaces stabilized by potato peptides (0.1 g/L, pH 7, 20 °C). Frequency: 0.02 Hz. Amplitude of deformation: 5%. [5] Net charge was calculated by using peptide property calculator from INNOVAGEN (Innovagen AB, Lund, Sweden). [6] Droplet size of emulsions after 1 and 7 days of storage in the dark at 20 °C. [7] Content of secondary volatile oxidation products 1-penten-3-ol and (*E,E*)-2,4-heptadienal after 7 days of storage in the dark at 20 °C. Lipid oxidation was accelerated by adding Fe^{2+}.

these studies also highlight caveats related to application of broad- specificity proteases, where hydrolysis products are challenging to predict. This substantiates the need for further developments into understanding the hydrolysis dynamics but also for more and diverse proteases of high specificity, to fully capitalize on the targeted hydrolysis approach. Nevertheless, we consider these examples highly encouraging for continuing down this path toward a workflow fully devoid of *in vitro* peptide-level validation but only relying on MS-based

Figure 4



Application of the bottom-up strategy for production of a potato protein hydrolysate with superior emulsifying and foaming properties, as presented by [24]. Based on a priori quantitative proteomics followed by bioinformatic prediction and *in vitro* validation of emulsifying peptides, a shortlist of target peptides was created. The (abundant) proteins of origin (including isoforms) for the selected peptides were subjected to *in silico* digestions by multiple proteases, and based on this, trypsin was selected as the best match for releasing peptides. In a benchmarking trial, hydrolysis was performed using multiple proteases to different degrees of hydrolysis and the emulsifying and foaming properties of the hydrolysates investigated. While trypsin showed to indeed produce a hydrolysate with superior properties, a subsequent deep (> 10 000 peptides identified in each hydrolysate) LC–MS/MS peptidomic analysis also revealed that overall, only trypsin was able to release the target peptides, which were linked with improved functionality. (adapted from [24] and created with Biorender.com).

proteomics and bioinformatic analysis (Figure 1B 'Future workflow').

Another option is fermentation by food-grade microorganisms to release desired peptides from the protein matrix. Lactic acid bacteria (LAB) are widely used in food fermentation and their proteolytic properties are usually due to cell wall-associated extracellular proteinases (CEP). The CEP enzymes from LAB have been reviewed by Ji et al. [45] and a recent review focused on microbial proteases with specificity for plant proteins [46]. Many LAB species are food grade and can be used directly as starter cultures for production of fermentates, whereas CEP enzymes are not usually used as a source of industrial enzymes due to the anchoring to the cell wall. The cleavage specificity of CEP enzymes from LAB seems not to be determined by the AA sequence around the cleavage site, but rather to arise from recognition of the substrate protein and a size selection of the resulting peptide products [47]. As peptide emulsifiers have certain requirements for length [23,27,37,42], and because specific proteins within a complex extract/isolate may be a source of more potent emulsifying peptides than others [23,27,31], CEP enzymes have properties suitable for the production of peptides with the desired properties in a selective manner. Enzymes with desired substrate selectivity can probably be found in nature or alternatively be evolved through protein engineering [46,48]. We will thus expect this type of enzymes to give rise to a new class of industrial proteases to produce sustainable food ingredients.

Conclusions and perspectives

Peptides produced by enzymatic hydrolysis are an emerging class of natural and sustainable emulsifiers. Although the conventional trial-and-error top-down approach is still dominant to produce emulsifying peptides, recent advancements have been made in a fundamentally orthogonal bottom-up strategy, facilitated by MS-based quantitative proteomics and bioinformatic functional prediction. Following identification of abundant proteins, full-length protein sequences can be processed bioinformatically to predict embedded peptides with functional properties, such as emulsification. Based on predicted, and potentially validated peptide-level functionality, designing a hydrolysis strategy based on *in silico* proteolysis has shown promise as a viable approach for improving bulk functional properties of hydrolysates.

Nevertheless, there are still challenges for applying this approach to create an economically viable production. First, increasing the proportion of targeted/functional peptides and thereby reducing the release of untargeted peptides by utilizing/developing proteases with high specificity would simplify the potential downstream processing (e.g. separation, purification, and stabilization). Second, increasing the overall yield of the process, by aiming at the production of additional functional peptides (e.g. antioxidants) from the protein-rich residue left after the obtaining of emulsifying targeted peptides, would add to the overall process economy. Third, the lack of analytical methods and bioinformatic workflows for accurate, quantitative evaluation of peptide composition makes the overall process evaluation challenging. Last, the effect of untargeted peptides on the functionality of targeted peptides is unexplored. A better understanding of such dynamics would add great value in the overall process evaluation and provide additional input for determining if additional downstream processing of the complex hydrolysate is required and/or desirable. Nonetheless, although the approach is still in its infancy and substantial advancements in multiple fields are needed to fully evaluate and capitalize on the concept, we foresee that this type of approach will gain significant headway in the coming years. Moreover, improved understanding of which physicochemical and structural attributes constitute strong peptide emulsifiers may facilitate developments of more advanced bioinformatic predictors. We believe that recent progress in peptide emulsifiers and the interdisciplinary actions presented here illustrate an undergoing paradigm shift within food science. The inclusion of molecular characterization, big data, and bioinformatics will continue to increase within the field. Through this, a transition from a conventional top-down way of thinking toward a bottom-up approach, where tailored functionality can be obtained through predictive, data-centric process design, will pave the way for significant advances in the quest for a more sustainable future.

CRediT authorship contribution statement

All authors: Conceptualization, Writing – original draft, Writing – review & editing.

Data Availability

Data will be made available on request.

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.

Acknowledgements

We are grateful for the financial support from Innovation Fund Denmark (Grant nr: 7045-00021B, PROVIDE project).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Bai L, Huan S, Rojas OJ, McClements DJ: **Recent innovations in emulsion science and technology for food applications**. *J Agric Food Chem* 2021, **69**:8944-8963.

2. Mozuraityte R, Kotsoni E, Rustad T: **Low-fat (< 50%) oil-in-water emulsions**. *Omega-3 Deliv Syst Prod Phys Charact Oxidative Stab* 2021,241-254, https://doi.org/10.1016/B978-0-12-821391-9.00010-7

3. Carrera Sánchez C, Rodríguez Patino JM: **Contribution of the engineering of tailored interfaces to the formulation of novel food colloids**. *Food Hydrocoll* 2021, **119**:106838.

4. Berton-Carabin CC, Sagis L, Schroën K: **Formation, structure, and functionality of interfacial layers in food emulsions**. *Annu Rev Food Sci Technol* 2018, **9**:551-587, https://doi.org/10.1146/annurev-food-030117

5. Yang J, Sagis LMC: **Interfacial behavior of plant proteins — novel sources and extraction methods**. *Curr Opin Colloid Interface Sci* 2021, **56**:101499.

6. Gould J, Wolf B: **Interfacial and emulsifying properties of mealworm protein at the oil/water interface**. *Food Hydrocoll* 2018, **77**:57-65.

7. Bertsch P, Böcker L, Mathys A, Fischer P: **Proteins from microalgae for the stabilization of fluid interfaces, emulsions, and foams**. *Trends Food Sci Technol* 2021, **108**:326-342.

8. Sagis LM, Yang J: **Protein-stabilized interfaces in multiphase food: comparing structure-function relations of plant-based and animal-based proteins**. *Curr Opin Food Sci* 2022, **43**:53-60.

9. Drusch S, Klost M, Kieserling H: **Current knowledge on the interfacial behaviour limits our understanding of plant protein functionality in emulsions**. *Curr Opin Colloid Interface Sci* 2021, **56**:101503.

10. Nikbakht Nasrabadi M, Sedaghat Doost A, Mezzenga R: **Modification approaches of plant-based proteins to improve their techno-functionality and use in food products**. *Food Hydrocoll* 2021, **118**:106789.

11. Schröder A, Berton-Carabin C, Venema P, Cornacchia L: **Interfacial properties of whey protein and whey protein hydrolysates and their influence on O/W emulsion stability**. *Food Hydrocoll* 2017, **73**:129-140.

12. •• García-Moreno PJ, Yang J, Gregersen S, Jones NC, Berton-Carabin CC, Sagis LMC, Hoffmann SV, Marcatili P, Overgaard MT, Hansen EB, *et al.*: **The structure, viscoelasticity and charge of potato peptides adsorbed at the oil-water interface determine the physicochemical stability of fish oil-in-water emulsions**. *Food Hydrocoll* 2021, **115**:106605.
The authors demonstrate the relationship between secondary structure (using SRCD), interfacial properties of the peptides and emulsion physical and oxidative stabilities.

13. Ricardo F, Pradilla D, Cruz JC, Alvarez O: **Emerging emulsifiers: conceptual basis for the identification and rational design of peptides with surface activity**. *Int J Mol Sci* 2021, **22**:4615.

14. Lin D, Sun L-C, Chen Y-L, Liu G-M, Miao S, Cao M-J: **Peptide/protein hydrolysate and their derivatives: their role as emulsifying agents for enhancement of physical and oxidative stability of emulsions**. *Trends Food Sci Technol* 2022, **129**:11-24.

15. Tamm F, Herbst S, Brodkorb A, Drusch S: **Functional properties of pea protein hydrolysates in emulsions and spray-dried microcapsules**. *Food Hydrocoll* 2016, **58**:204-214.

16. Morales-Medina R, Tamm F, Guadix AM, Guadix EM, Drusch S: **Functional and antioxidant properties of hydrolysates of sardine (S. pilchardus) and horse mackerel (T. mediterraneus) for the microencapsulation of fish oil by spray-drying**. *Food Chem* 2016, **194**:1208-1216.

17. García-Moreno PJ, Guadix A, Guadix EM, Jacobsen C: **Physical and oxidative stability of fish oil-in-water emulsions stabilized with fish protein hydrolysates**. *Food Chem* 2016, **203**:124-135.

18. Liu C, Bhattarai M, Mikkonen KS, Heinonen M: **Effects of enzymatic hydrolysis of fava bean protein isolate by alcalase on the physical and oxidative stability of oil-in-water emulsions**. *J Agric Food Chem* 2019, **67**:6625-6632.

19. Padial-Domínguez M, Espejo-Carpio FJ, Pérez-Gálvez R, Guadix A, Guadix EM: **Optimization of the emulsifying properties of food protein hydrolysates for the production of fish oil-in-water emulsions**. *Foods* 2020, **9**:636.

20. Opazo-Navarrete M, Morales E, Calderón-Chiu C, Calderón-Santoyo M, César Barros-Castillo J, Alfredo Díaz J, Arturo Ragazzo-Sánchez J, Salvador E, Salvador SC: **Structural modification of jackfruit leaf protein concentrate by enzymatic hydrolysis and their effect on the emulsifier properties**. *Colloids Interfaces* 2022, **6**:52.

21. Yu C, Zheng L, Cai Y, Zhao Q, Zhao M: **Desirable characteristics of casein peptides with simultaneously enhanced emulsion forming ability and antioxidative capacity in O/W emulsion**. *Food Hydrocoll* 2022, **131**:107812.

22. García-Moreno PJ, Jacobsen C, Marcatili P, Gregersen S, Overgaard MT, Andersen ML, Sørensen ADM, Hansen EB: **Emulsifying peptides from potato protein predicted by bioinformatics: stabilization of fish oil-in-water emulsions**. *Food Hydrocoll* 2020, **101**:105529.

23. •• García-Moreno PJ, Gregersen S, Nedamani ER, Olsen TH, Marcatili P, Overgaard MT, Andersen ML, Hansen EB, Jacobsen C: **Identification of emulsifier potato peptides by bioinformatics: application to omega-3 delivery emulsions and release from potato industry side streams**. *Sci Rep* 2020, **10**:690.
This is the first comprehensive study reporting a novel bottom-up approach to identify emulsifying peptides from sustainable proteins (e.g. potato) by using bioinformatics and quantitative proteomics.

24. •• Gregersen Echers S, Jafarpour A, Yesiltas B, García-Moreno PJ, Greve-Poulsen M, Hansen DK, Jacobsen C, Overgaard MT, Hansen EB: **Targeted hydrolysis of native potato protein: a novel workflow for obtaining hydrolysates with improved interfacial properties**. *Food Hydrocoll* 2022, **137**:108299, https://doi.org/10.1016/j.foodhyd.2022.108299.
This is an important study presenting a novel workflow for producing a potato protein hydrolysate with improved emulsifying and foaming properties using quantitative proteomics and bioinformatic prediction to facilitate targeted hydrolysis design.

25. Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR: **Protein analysis by shotgun/bottom-up proteomics**. *Chem Rev* 2013, **113**:2343-2394.

26. Bateman A, Martin MJ, Orchard S, Magrane M, Agivetova R, Ahmad S, Alpi E, Bowler-Barnett EH, Britto R, Bursteinas B, *et al.*: **UniProt: the universal protein knowledgebase in 2021**. *Nucleic Acids Res* 2021, **49**:D480-D489.

27. •• Yesiltas B, Gregersen S, Lægsgaard L, Brinch ML, Olsen TH, Marcatili P, Overgaard MT, Hansen EB, Jacobsen C, García-Moreno PJ: **Emulsifier peptides derived from seaweed, methanotrophic bacteria, and potato proteins identified by quantitative proteomics and bioinformatics**. *Food Chem* 2021, **362**:130217.
This study expands the use of the bottom-up approach combining quantitative proteomics, bioinformatics prediction, and functional validation to identify novel emulsifying peptides from seaweed and methanotrophic bacteria.

28. Jafarpour A, Gregersen S, Gomes RM, Marcatili P, Olsen TH, Jacobsen C, Overgaard MT, Sørensen ADM: **Biofunctionality of enzymatically derived peptides from codfish (Gadus morhua) frame: bulk in vitro properties, quantitative proteomics, and bioinformatic prediction**. *Mar Drugs* 2020, **18**:599.

29. Jafarpour A, Gomes RM, Gregersen S, Sloth JJ, Jacobsen C, Moltke Sørensen AD: **Characterization of cod (Gadus morhua) frame composition and its valorization by enzymatic hydrolysis**. *J Food Compos Anal* 2020, **89**:103469.

30. • Raghavan V, Kraft L, Mesny F, Rigerte L: **A simple guide to de novo transcriptome assembly and annotation**. *Brief Bioinform* 2022, **23**:bbab563.
This study presents an overview of and guide for performing *de novo* transcriptome assembly to facilitate construction of a reference protein database.

31. Gregersen Echers S, Abdul-Khalek N, Mikkelsen RK, Holdt SL, Jacobsen C, Hansen EB, Olsen TH, Sejberg JJP, Overgaard MT: **Is Gigartina a potential source of food protein and functional peptide-based ingredients? Evaluating an industrial, pilot-scale extract by proteomics and bioinformatics**. *Futur Foods* 2022, **6**:100189.

32. • Gregersen S, Pertseva M, Marcatili P, Holdt SL, Jacobsen C, García-Moreno PJ, Hansen EB, Overgaard MT: **Proteomic characterization of pilot scale hot-water extracts from the industrial carrageenan red seaweed Eucheuma denticulatum**. *Algal Res* 2022, **62**:102619.
This work presents a novel combination of transcriptomics, proteomics, and bioinformatics to determine theprotein composition of seaweed extracts.

33. Shevchenko A, Tomas H, Havliš J, Olsen JV, Mann M: **In-gel digestion for mass spectrometric characterization of proteins and proteomes**. *Nat Protoc* 2006, **1**:2856-2860.

34. Medzihradszky KF, Chalkley RJ: **Lessons in de novo peptide sequencing by tandem mass spectrometry**. *Mass Spectrom Rev* 2015, **34**:43-63.

35. Gregersen S, Kongsted ASH, Nielsen RB, Hansen SS, Lau FA, Rasmussen JB, Holdt SL, Jacobsen C: **Enzymatic extraction improves intracellular protein recovery from the industrial carrageenan seaweed Eucheuma denticulatum revealed by quantitative, subcellular protein profiling: a high potential source of functional food ingredients**. *Food Chem X* 2021, **12**:100137.

36. Enser M, Bloomberg GB, Brock C, Clark DC: **De novo design and structure-activity relationships of peptide emulsifiers and foaming agents**. *Int J Biol Macromol* 1990, **12**:118-124.

37. Saito M, Ogasawara M, Chikuni K, Shimizu M: **Synthesis of a peptide emulsifier with an amphiphilic structure**. *Biosci Biotechnol Biochem* 1995, **59**:388-392.

38. Dexter AF, Middelberg APJ: **Peptides as functional surfactants**. *Ind Eng Chem Res* 2008, **47**:6391-6398.

39. Eisenberg D, Weiss RM, Terwilliger TC: **The helical hydrophobic moment: a measure of the amphiphilicity of a helix**. *Nature* 1982, **299**:371-374.

40. Klausen MS, Jespersen MC, Nielsen H, Jensen KK, Jurtz VI, Sønderby CK, Sommer MOA, Winther O, Nielsen M, Petersen B, *et al*.: **NetSurfP-2.0: improved prediction of protein structural features by integrated deep learning**. *Proteins* 2019, **87**:520-527.

41. Nielsen H: **Predicting secretory proteins with SignalP**. *Methods Mol Biol* 2017, **1611**:59-73.

42. Dexter AF: **Interfacial and emulsifying properties of designed β-strand peptides**. *Langmuir* 2010, **26**:17997-18007.

43. Bjoerlie M, Yesiltas B, Garcia-Moreno PJ, Espejo-Carpio FJ, Rahmani-Manglano NE, Guadix EM, Jafarpour A, Hansen EB, Marcatili P, Echers SG, *et al*.: **Bioinformatically predicted emulsifying peptides and potato protein hydrolysate improves the oxidative stability of microencapsulated fish oil**. *bioRxiv* 2022, https://doi.org/10.1101/2022.11.18.517034

44. •• Minkiewicz P, Iwaniak A, Darewicz M: **BIOPEP-UWM virtual—a novel database of food-derived peptides with in silico-predicted biological activity**. *Appl Sci* 2022,7204.
This work outlines recent development of the BIOPEP-UWM database for bioactive peptides and bioinformatic tools for *in silico* analysis and identification.

45. Ji D, Ma J, Xu M, Agyei D: **Cell-envelope proteinases from lactic acid bacteria: biochemical features and biotechnological applications**. *Compr Rev Food Sci Food Saf* 2021, **20**:369-400.

46. Christensen LF, García-Béjar B, Bang-Berthelsen CH, Hansen EB: **Extracellular microbial proteases with specificity for plant proteins in food fermentation**. *Int J Food Microbiol* 2022, **381**:109889.

47. Hansen EB, Marcatili P: **Modeled structure of the cell envelope proteinase of Lactococcus lactis**. *Front Bioeng Biotechnol* 2020, **8**:6139861476.

48. Bell EL, Finnigan W, France SP, Green AP, Hayes MA, Hepworth LJ, Lovelock SL, Niikura H, Osuna S, Romero E, *et al*.: **Biocatalysis**. *Nat Rev Methods Prim* 2021, **1**:46.