

RESEARCH & REGULATORY ANALYSIS

The Protein Is Not the Ingredient

What AlterProtX Reveals About AI, Food Functionality and the Missing Bridge from Proteome Prediction to Industrial Validation

José M. López, *Head Research · INTABIOTECH*

Abstract

A paper published on 8 October 2026 introduces **AlterProtX**, an artificial-intelligence framework that compares animal and non-animal protein sources at proteome scale using predicted thermal stability, intermolecular-interaction descriptors and distribution-based representations. Its importance lies in shifting discovery beyond one protein at a time. Yet a predicted proteome is not a commercial ingredient, and molecular similarity does not establish gelation, emulsification, nutritional equivalence, allergen safety or regulatory permission. This article examines that translation gap against existing food-protein research, European novel-food and allergen rules, evolving British market-access arrangements and the United States’ August 2026 proposal to require GRAS notifications. It argues that the next useful innovation is an evidence-linked workflow connecting source biology, extracted protein composition, processing history, functional testing, nutrition, safety and jurisdiction-specific market access. A proposed Proteome-to-Process Evidence Graph and prospective counterfactual validation programme could make AI-guided protein discovery more reproducible, industrially relevant and legally defensible without mistaking computational prioritisation for a finished food technology.

KEYWORDS: *AI protein discovery; AlterProtX; proteome-scale screening; plant proteins; techno-functionality; novel foods; allergenicity; food regulation*

CATEGORY	Biotechnology · Functional Ingredients · AI & Food · Regulatory Affairs
SEO SLUG	ai-proteome-alternative-proteins-functionality-regulation
META DESCRIPTION	AI can rank alternative protein sources, but can it predict real food performance? A 2026 study exposes the gap between proteomes, processing and food law.
INTERNAL TAGS	Alternative Proteins · AI & Food · Protein Functionality · Novel Foods · Food Processing · Regulatory Affairs · Allergenicity · Emerging Research

1 | The signal from 8 October 2026

On 8 October 2026, Eunyeong Lee and Hee Yang published “**An AI-enabled proteome-scale framework for identifying sustainable protein alternatives for future food systems**” in *npj Science of Food*. Their system, **AlterProtX**, combines a prediction of protein thermal stability with six intermolecular-interaction attributes and converts those molecular descriptors into distribution-based representations of proteomes. The stated purpose is to compare animal and non-animal protein sources, interpret functional similarities and differences, and help prioritise candidates while considering allergenic potential and nutritional adequacy. The paper was accepted on 2 September and appeared online on 8 October; it is presented by the publisher as an early, citable version that may receive editorial amendments. [1]

The significant advance is conceptual. Food formulators rarely replace one purified molecule with another. They replace functional ingredient systems as mixtures of proteins, residual polysaccharides, lipids, minerals, antinutrients and processing-generated structures. Screening the protein repertoire of a biological source rather than treating each protein in isolation is therefore a reasonable direction for computational food science.

Yet this paper also provides an opportunity to examine a more difficult question. **When does a promising computational match become evidence of an industrially useful ingredient?** The answer cannot be “when the proteome looks similar.” It requires a chain of experimental and regulatory evidence that current food-protein databases only partially support.

2 | What AlterProtX does—and what its publication does not establish

According to the published abstract, **AlterProtX** predicts a processing-relevant molecular property, thermal stability, and integrates that prediction with six attributes associated with intermolecular interactions. The resulting proteome-level distributions permit comparative analysis across biological sources. Allergenicity and nutritional adequacy are incorporated into early-stage candidate prioritisation. [1]

Those statements should not be inflated. The publicly accessible abstract does not establish that **AlterProtX** has reproduced the rheology, sensory properties, shelf life or industrial processability of a finished dairy or meat analogue. It does not establish that an identified source is approved for sale, that its proteins are non-allergenic, or that its sustainability performance has been demonstrated by life-cycle assessment.

Established result: the authors report an AI-enabled screening and interpretative framework.

Reasonable interpretation: such a framework could reduce the number of biological sources that must enter expensive wet-laboratory screening.

Unproven leap: a high similarity score is a validated prediction of finished-product equivalence. That last proposition requires prospective, application-specific evidence.

3 | The crucial distinction: genome, proteome, ingredient and food

Four objects are frequently conflated in discussions of AI protein discovery.

Genome or predicted proteome. A sequence database indicates which proteins an organism may encode. It does not by itself reveal which proteins are expressed in the harvested material, at what abundance, or in which biochemical state.

Measured source proteome. Proteomic analysis of a tissue, seed or microbial biomass estimates the proteins actually present under defined biological and analytical conditions. Expression varies with genotype, maturity, cultivation conditions, stress, tissue and storage.

Commercial ingredient. Isolation, fractionation, concentration, washing, precipitation, drying and milling change protein abundance, folding, aggregation and interactions with non-protein material. A pea protein isolate is not a neutral miniature of the whole pea proteome.

Finished food. Protein functionality is expressed within a formulation and a processing history. It depends on pH, ionic strength, water activity, oil fraction, shear, heat, cooling, enzymes, carbohydrates and storage.

The key translation error is to treat a predicted source proteome as if it were the measured composition of the ingredient that will actually enter the customer's factory. An AI model can be correct about molecular descriptors yet wrong about industrial performance because it was asked to predict the wrong object.

4 | Why the proteome-scale idea is scientifically attractive

Food functionality is frequently emergent as it arises from interactions among multiple components rather than a single protein sequence. Casein systems, whey-protein systems, egg-white mixtures and legume isolates demonstrate that different protein populations can collectively generate interfacial films, elastic networks or viscoelastic structures.

A distribution-based representation can preserve information that a simple average conceals. Two protein sources may have identical mean predicted thermal stability but very different distributions: one dominated by moderately stable proteins, another combining highly labile and highly stable populations. The consequences for sequential denaturation, aggregation and gel-network formation could differ substantially.

This is a **mechanistically plausible rationale**, not proof that a particular distribution uniquely predicts food functionality. Proteins also interact with one another and with non-protein ingredients. The effects of relative abundance, co-extraction, particle structure and process history may be more important than a source-wide average. The scientific opportunity is to connect distributional descriptors to experimentally measured outcomes rather than to stop at visual similarity between computational fingerprints.

5 | Thermal stability is relevant, but it is not a gelation certificate

Thermal stability influences the temperature at which a protein unfolds or undergoes structural transitions. In food systems, those transitions can expose hydrophobic patches and reactive groups, encourage aggregation, alter solubility and permit formation of new networks, but the direction of the functional effect depends on the environment.

A protein that unfolds readily may be useful for heat-set gelation in one matrix and unacceptable for a clear acidic beverage in another. A protein that remains folded at high temperature may retain solubility, yet fail to produce the interfacial film or cohesive gel required by the target product. The presence of disulfide chemistry, the kinetics of aggregation, the concentration threshold for network formation and the availability of water all matter.

Consequently, predicted thermal stability should be treated as a **candidate-selection descriptor**, not a validated substitute for differential scanning calorimetry, rheology, solubility measurement or product-specific processing trials. The value of AlterProtX will depend on whether its descriptors help predict outcomes that industrial developers actually need to control. [1,2]

6 | The hidden variable: the process creates part of the ingredient

Consider two isolates prepared from the same plant species. One is obtained through mild aqueous fractionation; the other undergoes alkaline extraction, acid precipitation, repeated washing and high-temperature spray drying. Their amino-acid sequences can be essentially the same while their solubility, particle size, denaturation state, interfacial behaviour and gelling capacity differ.

This creates a modelling problem known as **domain shift**. A prediction trained on molecular or source-level descriptors may not transfer to ingredients manufactured through different extraction routes. In high-moisture extrusion, the gap widens because shear, residence time, moisture and cooling geometry create anisotropic structures not encoded in the original proteome.

A July 2026 review of AI-assisted plant-protein meat analogues explicitly places protein selection inside a broader design system involving extraction, rheology, fibrillation, nutrition, sensory performance and sustainability. A 2025 experimental comparison of machine-learning and conventional approaches to plant-protein extrusion shows that data-driven optimisation was already being studied before AlterProtX. The new paper should therefore be read as an extension of an active research field, not the invention of AI-guided alternative-protein design. [2,3]

7 | Functional similarity is application-specific

The expression “animal-protein alternative” is too broad to function as an experimental specification. A candidate suitable for a stable oil-in-water emulsion may be unsuitable for a high-moisture meat analogue. A protein that generates excellent foam overrun may collapse during refrigerated storage. A gelling system may meet instrumental hardness targets yet lack elasticity, cohesiveness or desirable mouthfeel.

Industrial validation should define the intended function before selecting a protein source. For an emulsion, useful endpoints include droplet-size distribution, interfacial adsorption, creaming stability and response to thermal cycling. For foam, overrun, drainage, bubble-size distribution and half-life matter. For a gel, storage modulus, loss modulus, gel point, fracture stress and water-holding capacity are relevant. For an extrudate, anisotropy, tensile behaviour, cooking loss and sensory texture become decisive.

These outcomes are not interchangeable. **An alternative protein is only an alternative relative to a specified function, matrix and processing window.** This is the first principle of a defensible AI-assisted screening brief.

8 | Molecular resemblance can conceal nutritional non-equivalence

A functional replacement does not automatically provide an equivalent nutritional contribution. A plant protein may perform well as an emulsifier while supplying a less favourable indispensable amino-acid pattern than the animal protein it replaces. Conversely, a nutritionally attractive isolate may lack the functional behaviour required by the product.

The FAO’s 2013 expert consultation recommended the Digestible Indispensable Amino Acid Score (DIAAS) as a scientifically stronger approach to dietary protein-quality assessment than relying exclusively on total protein or older digestibility-corrected scores. Digestible indispensable amino acids, not crude protein alone, are relevant to nutritional adequacy. [4]

An AI screen may use sequence-derived amino-acid composition as an early indicator. But sequence information does not measure ileal digestibility, antinutrient effects, thermal damage, lysine availability or the consequences of the finished matrix. Those require appropriate analytical and biological evidence. A “protein-equivalent” product may therefore be technologically similar but nutritionally different; the distinction is particularly important for children, older adults and products positioned as major protein sources.

9 | Allergenicity cannot be dismissed by a low model score

Allergen assessment is not a binary classification that can be delegated to a single algorithm. Sequence comparison and bioinformatics can help identify potential cross-reactivity with known allergens, but absence of a known match does not demonstrate absence of de novo sensitisation potential.

EFSA’s updated 2024 novel-food guidance distinguishes proteins from known allergenic sources, proteins from allergenic foods outside mandatory EU labelling categories, and novel proteins with uncertain allergenic

potential. Its tiered approach can require source history, phylogenetic assessment, sequence comparison, immunological testing with sera from allergic individuals and, where justified, clinical evidence. EFSA explicitly recognises the absence of validated predictive methods for de novo sensitisation. [5]

For candidate prioritisation, a computational allergenicity flag is valuable as a **screening and escalation tool**. It is not a safety clearance. A food ingredient's actual allergen profile can also differ from its source proteome because fractionation enriches or removes particular proteins. The analytical object must be the final commercial ingredient and, when appropriate, the processed food.

10 | A sustainable protein must be assessed as a system

The word “sustainable” in a paper title should not be interpreted as a demonstrated environmental advantage for every candidate identified. Resource intensity depends on agricultural yield, land use, fertiliser, irrigation, processing energy, solvent recovery, concentration yield, co-product allocation, transport and food loss. A novel source can appear promising at the organism level while requiring energy-intensive extraction or generating a large waste stream.

An appropriate comparative life-cycle assessment must define a functional unit. Comparing one kilogram of wet biomass with one kilogram of purified protein is not meaningful. Comparing one kilogram of protein without considering digestibility or technological functionality can also mislead. In some applications, a lower-protein ingredient that achieves the same gel strength at a lower inclusion rate could have a different environmental profile from an isolate with superior amino-acid composition.

Expert interpretation: future AI screening should include environmental constraints only where data quality, system boundaries and functional equivalence are explicit. An environmental score is not evidence of sustainability unless its underlying inventory and assumptions are disclosed.

11 | The 2026 literature and patent record: this is an emerging synthesis, not a blank field

The novelty boundary deserves care. Computational protein design, machine learning for food formulation and high-moisture extrusion optimisation are established research directions. A comprehensive July 2026 review described AI as an integrative layer connecting protein source, extraction, texture, nutrition, processing and sustainability. A 2025 study directly compared machine-learning and conventional experimental optimisation of plant-protein extrusion. A separate 2026 review surveys computational approaches to alternative proteins. [2,3,6]

Public patent literature also shows that alternative-protein functionality is already an active intellectual-property domain. International application WO2025027133A1, published in 2025, concerns methods for improving gelation properties of alternative proteins and includes pea-protein-related subject matter. Its existence does not mean that it anticipates AlterProtX or the research agenda proposed here; it shows that a broad claim to “optimising alternative-protein gelation” would not be credible as a newly discovered field. [7]

The defensible novelty claim for AlterProtX is the authors' reported **particular multiscale integration and proteome-level representation**, subject to detailed examination of the full paper and prior art. The research proposals in this article are described as hypotheses and development architectures, not as patented inventions or established standards.

12 | Research & Innovation Opportunity: a Proteome-to-Process Evidence Graph

Unresolved problem

There is no universally adopted, industrially validated information architecture linking computational protein-source descriptors to the actual extracted ingredient, its processing history, its behaviour in a specified food

matrix, nutritional quality, allergen evidence and market-access status. The literature identifies the underlying challenge, but an interoperable end-to-end evidence standard is not established. [1,2]

Why it matters

AI screening can select a biologically interesting source that fails during extraction, has poor solubility in the customer's matrix, generates an unacceptable flavour, lacks adequate nutritional quality or cannot lawfully enter the target market. The cost of this mismatch is not just unsuccessful experimentation; it is misallocated R&D effort and premature commercial claims.

What is already known

Protein-sequence models, thermal-stability prediction, proteomic analysis, laboratory functional testing, DIAAS-oriented nutritional assessment, allergen screening, novel-food dossiers and digital laboratory information systems already exist. None of those elements should be presented as newly invented here. [1,2,4,5]

Proposed hypothesis

A **Proteome-to-Process Evidence Graph (PPEG)** could improve candidate ranking and transferability by explicitly linking each prediction to the material and conditions for which it was made. The graph would distinguish *predicted proteome* from *measured protein fraction* and *finished formulation* rather than allowing these objects to collapse into a single "protein source" record.

Minimum architecture

Taxon/cultivar → genome/proteome version → predicted molecular descriptors → measured expression/abundance → extraction and drying history → ingredient specification → formulation and process window → rheology/interfacial function → nutritional and allergen evidence → legal status by jurisdiction → industrial trial.

Each relationship would carry provenance, measurement method, uncertainty, batch identifier and date. A prediction would be linked to the model version and training domain; an experimental result to its protocol and material batch. The graph should retain negative results, because failed candidates define the boundaries of model usefulness.

Validation pathway

Stage 1: select a limited, well-characterised group of animal and non-animal protein systems with different processing histories. Stage 2: establish standard assays and measure protein abundance, solubility, thermal transitions, emulsification, gelation and relevant nutritional variables. Stage 3: compare candidate rankings using sequence-only, proteome-distribution and process-aware models. Stage 4: test external lots, independent laboratories and unseen protein sources. Stage 5: evaluate calibration, false-positive candidate selection, performance after scale-up and time saved per validated industrial lead.

Regulatory constraints and industrial relevance

The graph would support dossier preparation and supplier qualification, not replace EFSA assessment, food-law classification or clinical allergen evidence. Commercial data governance and intellectual-property rights would need defined access controls. If validated, the system could reduce blind formulation trials and make ingredient selection more reproducible across food categories. **This is a proposed R&D architecture, not an existing regulatory standard.**

13 | A second opportunity: the counterfactual formulation challenge

A frequent weakness of AI studies is that they demonstrate internal correlation but not useful decision improvement. A stronger prospective test would ask whether a model can choose a non-animal ingredient that performs as well as a deliberately difficult animal-protein reference under an unfamiliar processing condition.

The experimental challenge could define three target systems: a heat-set gel, an oil-in-water emulsion and an aerated foam. For each, an animal-protein control would be characterised across a specified pH–salt–temperature design space. Candidate sources would then be ranked without access to the withheld performance measurements. Researchers would formulate the top-ranked candidates and an equal number of randomly or conventionally selected controls.

The primary endpoint would not be prediction accuracy on a historical dataset. It would be the **number of independent industrially relevant formulations meeting prespecified acceptance criteria per laboratory trial**. Secondary endpoints could include nutritional adequacy, allergen-screening burden, process robustness and environmental data completeness.

This is a proposed validation experiment. It does not imply that AlterProtX has already passed it. Its value would be to test whether proteome-scale intelligence creates genuine experimental efficiency beyond familiar heuristics and simpler machine-learning baselines.

14 | Model validation: prevent leakage, overfitting and false precision

Proteins from related species and homologous families are not statistically independent. Randomly splitting sequences into training and test sets can produce optimistic results when close homologues appear in both. Proteome comparisons can inherit the same problem at source level: different cultivars or closely related taxa may share much of their protein repertoire.

A rigorous validation programme should therefore include phylogeny-aware or source-held-out tests, out-of-distribution evaluation and transparent baselines. It should report uncertainty and calibration, not just a correlation coefficient. Where functional labels come from different laboratories, the model should record assay protocols, pH, ionic strength, protein concentration, temperature ramp and extraction route.

The important question is whether a prediction remains useful when the **source, supplier, production batch, extraction process or equipment** changes. A model that ranks known candidates correctly but fails on unseen commercial lots may be academically interesting and industrially weak. This is an expert methodological recommendation based on general machine-learning principles and the data-fragmentation problems identified in current food-protein reviews. [2,3,6]

15 | EU and Spain: AI does not create an exemption from food law

For the European Union, including Spain, the regulatory starting point is the **actual ingredient and its manufacturing route**, not the algorithm that identified it. Regulation (EU) 2015/2283 establishes the novel-food framework. Whether a protein ingredient is novel depends on its status, significant history of consumption before 15 May 1997 where relevant, production process and the categories defined in the Regulation. The Union list of authorised novel foods is established by Implementing Regulation (EU) 2017/2470, as amended. [8,9]

It would be incorrect to claim that every new plant-protein isolate automatically requires novel-food authorisation. It would be equally incorrect to assume that every plant protein is non-novel because the plant itself has long been eaten. A materially different fraction, new source or novel production process can change the assessment. The European Commission's Novel Food Status Catalogue is useful but non-exhaustive and should not replace product-specific legal analysis. [8,9,10]

EFSA's 2024 scientific guidance requires evidence on identity, production, composition, specifications, proposed uses and intakes, nutritional information, toxicology and allergenicity as appropriate. The applicant must demonstrate safety under intended conditions of use; computational similarity is supportive information, not an authorisation. [5]

If genetic modification is involved in creating or producing the ingredient, Regulation (EC) No 1829/2003 and associated rules may become relevant depending on the precise relationship between the GMO, the process and the marketed food. **Using AI to select a conventional protein source does not, by itself, make the ingredient a GMO.** The legal analysis must follow the manufacturing facts, not marketing terminology. [11]

16 | EU labelling and nutrition claims: separate function from nutritional representation

Regulation (EU) No 1169/2011 governs food information and requires declaration of listed allergens in relevant circumstances. Its Annex II includes, among others, cereals containing gluten, peanuts, soybeans, milk and eggs, subject to defined exceptions. Novel proteins outside those categories may still pose allergenic risks; the absence of a mandatory Annex II source label is not evidence of safety. [12]

Regulation (EC) No 1924/2006 provides a separate framework for nutrition and health claims. In its Annex, "source of protein" generally requires at least 12% of the food's energy to be provided by protein; "high protein" requires at least 20%. Those are legal conditions for specified nutrition claims, not proof of digestibility, amino-acid adequacy or equivalence to animal protein. [13]

A formulation that successfully replaces egg foaming function cannot automatically claim egg-like nutrition. Likewise, a "high protein" claim does not mean that a food has been shown to provide the same indispensable amino-acid availability as a comparator. Food technologists should maintain separate files for **techno-functionality, nutrition, allergenicity** and **legal claim substantiation**.

17 | Great Britain and Northern Ireland: present divergence, prospective realignment

As of 9 October 2026, Great Britain has a domestic authorisation route for novel foods under its assimilated regulatory framework, while Northern Ireland continues to apply relevant EU novel-food law under the existing arrangements. The UK Food Standards Agency maintains a register of novel-food authorisations for Great Britain and publishes an Innovative Food Guidance Hub for emerging technologies. [14,15]

There is also an important forward-looking development. UK government guidance updated on 30 July 2026 states that the government is negotiating a sanitary and phytosanitary (SPS) agreement with the EU and **intends** the new arrangements to take effect in mid-2027. The guidance anticipates that EU authorisations for regulated products, including novel foods, would replace Great British authorisations within the agreed scope. But the government expressly states that negotiations and implementation details are not yet final. [16]

For an alternative-protein R&D programme, the correct response is scenario planning, not premature assertion that alignment has already occurred. A candidate intended for EU, GB and Northern Ireland distribution should be tracked separately under current law, with a clearly marked future-alignment scenario. This is particularly important where a product relies on a GB-specific authorisation or proprietary-data protection that may not survive the proposed transition.

18 | United States: GRAS is changing, but the August proposal is not yet law

The United States does not operate an EU-style horizontal novel-food authorisation system. Under the Federal Food, Drug, and Cosmetic Act, the legality of an intentionally added substance depends on its intended use and whether it is an authorised food additive, generally recognised as safe (GRAS) for that use, or otherwise excluded from the food-additive definition. FDA explains that GRAS status requires generally available evidence and

recognition of safety among qualified experts. Under the present framework, FDA's GRAS notification programme is voluntary. [17]

A significant 2026 regulatory signal could change the reporting environment. On 10 August 2026 FDA announced a proposal, published in the *Federal Register* on 11 August, to **require submission of GRAS notices** for human- and animal-food substances claimed to be GRAS under their intended conditions of use. The proposal is open for comments until 9 December 2026. It is a proposed rule, not an enacted requirement as of this article's date. [18]

For AI-discovered protein ingredients, the lesson is that **a model-generated safety score cannot establish GRAS status**. Identity, manufacturing process, specifications, exposure and a generally recognised safety evidence base remain central. The prospective notification reform increases the strategic value of early, well-organised safety dossiers, but should not be presented as already binding.

19 | The regulatory decision tree must follow the material, not the marketing category

A useful pre-commercial decision tree begins with five questions.

- 1. What is the source?** Conventional edible plant, unusual plant, algae, fungus, bacterium, animal-derived material or recombinant production organism?
- 2. What is the ingredient?** Whole flour, concentrate, isolate, hydrolysate, purified protein, mixed biomass or fermentation-derived recombinant protein?
- 3. How was it produced?** Conventional fractionation, enzymatic modification, new extraction process, fermentation, genetically modified organism or another route?
- 4. What is the intended use?** Food category, inclusion rate, vulnerable population, replacement claim and expected exposure?
- 5. Where will it be sold?** EU/Spain, Great Britain, Northern Ireland, United States or another jurisdiction?

These questions determine which regulatory assessments are relevant. "Alternative protein" is a commercial category, not a uniform legal status. A plant-derived isolate, a microbial biomass protein and a purified recombinant milk protein may require materially different evidence packages even if all three perform the same emulsifying function.

20 | The industrial challenge: replace the function without losing control

Dairy alternatives and acidic beverages

A candidate may have promising molecular descriptors but precipitate near its isoelectric point, interact unfavourably with minerals or generate sediment during shelf life. Screening must measure solubility and particle stability at the actual beverage pH, ionic strength, heat treatment and storage conditions. Emulsification alone is not enough.

Meat analogues and textured proteins

High-moisture extrusion requires a coupled response to moisture, shear, temperature and cooling. A source with good gelation in a beaker may fail to form fibrous anisotropy in an extruder. Model outputs should be linked to rheology, protein denaturation, microstructure and process conditions. [2,3]

Egg-function replacements

Foaming, emulsification, heat-setting and binding are distinct functions. Replacing one does not guarantee the others. A bakery application should be evaluated in the complete batter and after baking, not only through isolated protein foam tests.

Nutritional products

A candidate selected for high protein content must also meet digestibility, amino-acid, sensory and contaminant specifications. Claims concerning health benefits or suitability for special populations require separate evidence and legal review.

21 | A practical development sequence for ingredient suppliers

The most useful operational consequence is to reorder the R&D pipeline. Begin with a **product-function specification**, not with a fashionable source. Define a quantitative acceptance window against a commercial benchmark and document the target matrix, process and shelf-life requirements.

Next, use sequence and proteome tools to prioritise candidate sources. Then verify the actual ingredient: compositional analysis, protein fraction distribution, batch variation, extraction history and relevant impurities. Conduct rapid bench tests before complex pilot trials. Only candidates meeting predefined technical thresholds should progress to process validation, nutritional and allergen assessment, jurisdiction-specific legal classification and industrial scale-up.

A failed candidate should not disappear from the database. Negative results can reveal that the model overlooked a specific pH interaction, a drying effect, an antinutrient, a mineral sensitivity or an equipment limitation. A disciplined feedback loop converts those failures into improved future screening.

Economic modelling belongs after technical feasibility and a credible industrial process have been established. Premature cost claims can create false precision around an ingredient whose yield, functionality and legal status remain uncertain.

22 | Where this matters for INTABIOTECH

For INTABIOTECH, the strongest connection is to its protein-ingredient and formulation-development domains, particularly textured plant-protein systems and functional protein applications. The opportunity is not to claim that existing INTABIOTECH ingredients have been validated by AlterProtX or that an AI score confers superior performance. No such evidence is presented here.

A practical internal pilot could start with a small portfolio of commercially relevant protein ingredients -such as pea-protein-based texturisation candidates- and a few tightly defined functions: water retention, emulsion stability, gel strength and extrusion behaviour. Each lot would be linked to its specification, processing history, measured protein profile, functional assays and customer application.

Over time, this could support a **function-first ingredient dossier**: not merely “protein ≥65%”, but a documented relationship between raw material, fractionation, processing conditions, rheological response, nutritional constraints and suitable industrial uses. Such a dossier would improve supplier qualification and make technical recommendations more defensible. It could also support future AI and big-data initiatives without substituting modelling for laboratory evidence.

23 | Data governance and intellectual property are part of the research design

Protein-sequence databases, public scientific papers, supplier technical files and customer formulation data carry different rights and confidentiality conditions. An AI programme must document provenance, licences, model versions, validation datasets and restrictions on use. Customer formulations and proprietary processing windows should not be exposed to third-party systems without appropriate contractual safeguards.

Patent clearance is separate from scientific novelty. The existence of prior patent publications on protein gelation and functional modification means that a commercial R&D programme should conduct a targeted freedom-to-operate review before committing to a particular processing route. A published paper does not confer freedom to

operate, and a patent application does not establish that a claimed invention is valid or infringed by a specific process. [7]

The defensible innovation asset may be a curated, high-quality dataset and a reproducible decision workflow rather than an entirely new protein molecule. That possibility should be evaluated on its own technical and legal merits.

24 | Evidence map: what can be said today

Established evidence. Food proteins exhibit matrix- and process-dependent functionality; alternative-protein AI and extrusion optimisation are active research fields; FAO and EFSA have published frameworks for protein quality and allergen assessment; EU and US food-law obligations depend on ingredient identity, production route and intended use. [2–5,8,12,17]

Newly reported evidence. The 8 October 2026 AlterProtX paper reports a multiscale AI framework combining thermal-stability prediction, six intermolecular-interaction attributes, distribution-based proteome representations, and early-stage consideration of allergenic and nutritional factors. [1]

Expert interpretation. Proteome-level similarity is likely to be more useful when paired with measured ingredient composition and processing data. The principal bottleneck for industrial transfer is the missing evidence chain between source biology and finished-product behaviour.

Hypotheses. A Proteome-to-Process Evidence Graph and a prospective counterfactual formulation challenge could improve candidate prioritisation and model generalisability. Neither is claimed here as an established method, regulatory standard or patentable novelty.

Regulatory status. The EU novel-food framework is in force; current GB authorisation rules remain relevant pending any concluded SPS agreement; FDA’s August 2026 mandatory-GRAS-notification initiative is a proposal, not a final rule. [8,14,16,18]

25 | Conclusions: from prediction to proof

AlterProtX is timely because it asks the right upstream question: can AI compare entire protein repertoires and identify non-animal sources with properties relevant to food technology? Its proteome-scale perspective is more aligned with the complexity of food ingredients than a narrow one-protein substitution model. [1]

The harder downstream question remains open. The proteome is not the ingredient. The ingredient is not the formulation. The formulation is not the process. And successful process performance is not proof of nutritional equivalence, allergen safety, sustainability or market authorisation.

The next step for food biotechnology is therefore not another unqualified claim that AI can “replace animal protein”. It is a **traceable chain of evidence** that begins with computational selection and ends with independently reproducible industrial performance under a lawful intended use.

If AI can shorten that chain without weakening it, its contribution will be substantial. If it merely generates rankings that cannot survive extraction, formulation, scale-up and regulatory scrutiny, it will remain an elegant discovery interface rather than a food-industry technology.

The future of alternative proteins will be decided not by the protein the model predicts, but by the ingredient the factory can consistently make and the food the consumer can safely use.

Related Knowledge Areas of this paper

- Biotechnology — proteome screening, protein characterisation and biological source selection.
- Functional Ingredients — emulsification, gelation, foaming and protein-system performance.

- Processing Technologies — extraction, drying, rheology and high-moisture extrusion.
- Nutrition — indispensable amino acids, digestibility and nutritional equivalence.
- AI & Food — evidence graphs, model validation and data governance.
- Regulatory Affairs — novel-food classification, allergen assessment, GRAS and claims.
- Emerging Research — prospective counterfactual testing and protein-to-process validation.

Editorial Scope and Verification Note

Evidence and legal status assessed as of 9 October 2026. The interpretation of **AlterProtX** is limited to the verifiable publication abstract and bibliographic record; no unreported numerical model-performance result, industrial demonstration or authorisation is attributed to the authors. Jurisdiction-specific status, amended legal texts and product-specific authorisations should be rechecked before commercial decisions. This article is an independent critical analysis, not an endorsement of any unvalidated ingredient or algorithm.

References and Primary Sources

- [1] Lee E, Yang H. An AI-enabled proteome-scale framework for identifying sustainable protein alternatives for future food systems. *npj Science of Food*. Published 8 October 2026. DOI: 10.1038/s41538-026-01159-x. [Official source / DOI](#)
- [2] Amiri A, Masztalerz K, Maggo S, Singh A, Korolev V, Pratap-Singh A. Artificial Intelligence-Assisted Design of Plant-Protein Meat Analogues: Integrating Nutrition, Functionality, and Fibrillation-Based Texturization. *Comprehensive Reviews in Food Science and Food Safety*. 2026;25:e70548. DOI: 10.1111/1541-4337.70548. [Official source / DOI](#)
- [3] Jiang Y, Noor Azlee NIB, Ko WS, Chen K, Lim BG, Nelson AZ. Plant-based protein extrusion optimization: Comparison between machine learning and conventional experimental design. *Current Research in Food Science*. 2025;101157. DOI: 10.1016/j.crfs.2025.101157. [Official source / DOI](#)
- [4] FAO. Dietary protein quality evaluation in human nutrition. Report of an FAO Expert Consultation. FAO Food and Nutrition Paper 92. Rome, 2013. [Official source / DOI](#)
- [5] EFSA. Guidance on the scientific requirements for an application for authorisation of a novel food in the context of Regulation (EU) 2015/2283. *EFSA Journal*. 2024;22:e8961. DOI: 10.2903/j.efsa.2024.8961. [Official source / DOI](#)
- [6] Vyshnav S, Sundararajan V, Sudhakaran SL. Computational methods for alternative proteins. *Advances in Protein Chemistry and Structural Biology*. 2026;154:383–439. DOI: 10.1016/bs.apcsb.2026.08.013. [Official source / DOI](#)
- [7] WIPO/PCT. WO2025027133A1. Method of improving gelation properties of alternative proteins. Published 2025. Public patent disclosure; legal status and claim scope require separate examination. [Official source / DOI](#)
- [8] European Parliament and Council. Regulation (EU) 2015/2283 on novel foods. In force; consolidated legal text. [Official source / DOI](#)
- [9] European Commission. Implementing Regulation (EU) 2017/2470 establishing the Union list of novel foods. Consolidated text available as of August 2026. [Official source / DOI](#)
- [10] European Commission. Novel Food Status Catalogue. Official non-exhaustive status resource. [Official source / DOI](#)
- [11] European Parliament and Council. Regulation (EC) No 1829/2003 on genetically modified food and feed. [Official source / DOI](#)
- [12] European Parliament and Council. Regulation (EU) No 1169/2011 on the provision of food information to consumers; Annex II. [Official source / DOI](#)
- [13] European Parliament and Council. Regulation (EC) No 1924/2006 on nutrition and health claims made on foods; Annex, source of protein and high protein. [Official source / DOI](#)
- [14] Food Standards Agency. Authorised Regulated Food and Feed Products for Great Britain: novel-food authorisations register. Updated March 2026. [Official source / DOI](#)
- [15] Food Standards Agency. Innovative Food Guidance Hub. Updated 30 June 2026. [Official source / DOI](#)
- [16] UK Department for Environment, Food & Rural Affairs. Food manufacturing, processing and standards: preparing your business for a proposed UK–EU SPS agreement. Updated 30 July 2026. [Official source / DOI](#)
- [17] U.S. Food and Drug Administration. Understanding How the FDA Regulates Food Additives and GRAS Ingredients. [Official source / DOI](#)
- [18] U.S. Food and Drug Administration. Substances Generally Recognized as Safe. Proposed rule; Federal Register 91 FR 51834, published 11 August 2026; comment deadline 9 December 2026; Docket FDA-2025-N-3262. [Official source / DOI](#)
- [19] U.S. Food and Drug Administration. Plant-Based Milk and Animal Food Alternatives; 2025 draft labelling guidance context. [Official source / DOI](#)

[20] Sun Y et al. AI proteomics: from protein identification to virtual cells. Nature Methods. 2026;23:1697–1709. DOI: 10.1038/s41592-026-03085-y. [Official source / DOI](#)

© 2026 José M. López · All rights reserved.