

# RaRaMa — Protein Functionality & Thermal Dossier

September 19, 2026 · 22:43 | Tier 3 · Full dossier | source: `bench_agent_run_3338a79ded.json` | 3 proteins | prepared for: `bench_agent`

## Bottom line — the decision in one line per protein

- 1. **beta-lactoglobulin** — LIPOCALIN fold (PROVISIONAL confidence); predicted  $T_d \approx 73\text{ }^{\circ}\text{C}$ ; functional profile in the table below · 3 bitter-prone fragment(s) flagged (rank screen, see the stack).
- 2. **soy glycinin 11S** — CUPIN fold (HIGH confidence); predicted  $T_d \approx 94\text{ }^{\circ}\text{C}$ ; functional profile in the table below · 6 bitter-prone fragment(s) flagged (rank screen, see the stack).
- 3. **alpha-lactalbumin** — ALPHA\_LAC fold (PROVISIONAL confidence); predicted  $T_d \approx 64\text{ }^{\circ}\text{C}$ ; functional profile in the table below · 4 bitter-prone fragment(s) flagged (rank screen, see the stack).

## Protein register — every input, numbered & structurally identified

| # | Protein            | Len (aa) | Fold      | Confidence  |
|---|--------------------|----------|-----------|-------------|
| 1 | beta-lactoglobulin | 162      | LIPOCALIN | PROVISIONAL |
| 2 | soy glycinin 11S   | 476      | CUPIN     | HIGH        |
| 3 | alpha-lactalbumin  | 123      | ALPHA_LAC | PROVISIONAL |

## Virtual-instrument method conditions — the settings behind every number (audit-grade)

| Virtual instrument  | Method conditions (operator controls · defaults where blank)                                                                                     | Reported output                                                |
|---------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------|
| TA DSC (Discovery)  | heating rate 10 °C/min · N <sub>2</sub> purge 50 mL/min · Tzero AI hermetic pan · ~5 mg                                                          | Td/Tm °C · ΔH · ΔCp · cooperativity                            |
| TA rheometer (ARES) | 1% strain · 0.1–100 rad/s frequency sweep · 14% w/v · heat-set gel: heat 20→95 °C (gel point at T <sub>gel</sub> ≈T <sub>d</sub> ), cool to cure | G'/G'' Pa · gel point °C · K'/n'                               |
| TA DMA / TGA        | moisture as supplied% · ramp to die temp · N <sub>2</sub> 20 °C/min                                                                              | Tg °C · E'/E'' · decomp onset °C                               |
| Wyatt SEC-MALS      | calibrated SEC column · dn/dc 0.185 mL/g · 90° + LS detectors                                                                                    | absolute Mw kDa · oligomer · %agg                              |
| Wyatt DynaPro DLS   | back-scatter · 25 °C · aqueous · cumulants fit (ISO 22412)                                                                                       | Rh nm · PDI · monodispersity                                   |
| SAXS (BioSAXS)      | q 0.1–3.0 nm <sup>−1</sup> · Guinier + Kratky · buffer-subtracted                                                                                | Rg nm · Dmax · fold state                                      |
| UV-PDA (ACQUITY eλ) | 190–800 nm · 1.2 nm optical res · 10 mm flow cell · 80 Hz · stray 0.07%                                                                          | ε280 · A280 · λmax · linearity                                 |
| LC-MS (ESI, SYNAPT) | BEH C18 1.7 μm · 5→95% B gradient · 0.5 mL/min · ESI+ · cone 30 V · N2 drift gas                                                                 | intact mass Da · collision cross-section (shape) · tryptic map |
| Nutrition (DIAAS)   | FAO/WHO 2013 reference AA pattern · structural SID digestibility                                                                                 | DIAAS · limiting AA · SID                                      |

Every virtual instrument computes from the amino-acid sequence + the panel's line conditions; the method row records the operator controls that shaped each output (scan rate → kinetic Td shift, flow-cell pathlength → absorbance, drift gas → CCS scale). Blank controls fall back to the standard method defaults shown.

Functional readout — every protein, sorted to your tier (the goal: items 2-7)

| Protein            | T <sub>d</sub> ± band °C | tier | pI  | Gel kPa | LGC % | Emu ls (S AI) | WH C | OHC  | ΔH   | TGA °C | Tg° C | DIA AS |
|--------------------|--------------------------|------|-----|---------|-------|---------------|------|------|------|--------|-------|--------|
| beta-lactoglobulin | 73 ±5                    | PROV | 4.6 | 1.8     | 6.8   | 0.84          | 1.25 | 0.49 | 15.8 | 249    | 68    | 0.83   |
| soy glycinin 11S   | 94 ±9                    | HIGH | 4.9 | 3.1     | 9.7   | 0.82          | 1.73 | 0.37 | 15.8 | 249    | 80    | 0.92   |
| alpha-lactalbumin  | 64 ±5                    | PROV | 4.6 | 5.7     | 7.9   | 0.68          | 1.54 | 0.40 | 15.8 | 249    | 83    | 0.86   |

Tier 3 full dossier: + emulsification surface-activity index (SAI, correlates with EAI m<sup>2</sup>/g per Kato-Nakai), water/oil holding, unfold enthalpy ΔH (DSC), decomposition onset (TGA), glass transition T<sub>g</sub> (DMA shelf-life/caking, at the supplied moisture) & predicted DIAAS (composition × structural digestibility — a screen, not a label claim; PDCAAS stays official). **T<sub>d</sub> is reported as its disclosed ±band** — the value is deterministic; the ± is the fold-family's characterised uncertainty (so a lower-confidence fold is NOT dressed up as a false-precision point). **tier** = HIGH (calibrated cupin core, ±4 °C) · PROVISIONAL (other recognised fold, wider band) · a novel or disordered chain is **refused** (—), never guessed.

Allergen family screen — sequence-homology pre-bench flag (the EFSA-endorsed first step)

| Protein            | Allergen family        | Cross-reactive with                                                                                                                                                                                                                                                                  | Heat stability                                              |
|--------------------|------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------|
| beta-lactoglobulin | milk whey lipocalin    | within the lipocalin super-family                                                                                                                                                                                                                                                    | heat-stable — largely retained through processing           |
| soy glycinin 11S   | cupin 11S legumin      | other legume/tree-nut 7S/11S seed-storage globulins by sequence homology (soy, tree-nut, peanut families) — a homology FLAG, not confirmed IgE; actual cross-reactivity is partner-specific and IgE-tested (e.g. some 11S legumins cross-react with cashew/pistachio but not peanut) | very heat-stable — survives cooking; allergenicity retained |
| alpha-lactalbumin  | milk alpha lactalbumin | within the alpha-lactalbumin super-family                                                                                                                                                                                                                                            | heat-labile — cooking often reduces allergenicity           |

Sequence-homology allergen FAMILY screen — the pre-bench first step EFSA runs AHEAD of the wet-lab IgE panel. This is a family / cross-reactivity FLAG read from the sequence, not a clinical diagnosis; the certified IgE / oral-food-challenge panel stays official. Heat stability indicates whether cooking is likely to reduce the risk — heat-stable families (cupin storage globulins, 2S albumins) retain allergenicity through processing.

Results & mechanism — read like the methods of a paper, generated per protein

Each paragraph is assembled from this protein's own computed values; the clauses that appear, and the numbers in them, differ for every input.

beta-lactoglobulin

| Key readout            | Value     |
|------------------------|-----------|
| Unfolding temp Td      | 73 ± 5 °C |
| Radius of gyration Rg  | 1.72 nm   |
| Hydrodynamic radius Rh | 2.22 nm   |
| Rg / Rh                | 0.78      |
| Unfolding enthalpy ΔH  | 15.8 J/g  |
| Isoelectric point pI   | 4.56      |
| Gel modulus G'         | 1.76 kPa  |
| Emulsification (EAI)   | 0.84      |
| Water-holding (WHC)    | 1.25      |

|                   |      |
|-------------------|------|
| Oil-holding (OHC) | 0.49 |
|-------------------|------|

The assembled particle has a predicted radius of gyration  $R_g = 1.7 \text{ nm}$  (SAXS observable) and a hydrodynamic radius  $R_h = 2.2 \text{ nm}$  predicted from  $R_g$  for a compact globule ( $R_g/R_h = \sqrt{(3/5)} \approx 0.78$  by geometry — a modelled DLS estimate, not a second measurement; confirm independently by DLS), placing it in a small, stable size range. Aggregation propensity (size-driven): **LOW** (score 0.00) — estimated from the predicted particle size, deterministically and with no molecular-dynamics step. A pre-instrument indicator that complements DLS/SEC-MALS. Predicted denaturation temperature  $T_d = 73 \pm 5 \text{ }^\circ\text{C}$  (recognised LIPOCALIN fold, reported as a band — the value is deterministic, the  $\pm$  is the fold-family's disclosed uncertainty) — thermolabile for a whey lipocalin. Unfolding enthalpy  $\Delta H \approx 15.8 \text{ J/g}$  with  $\Delta C_p \approx 9.4 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$  (the folded→unfolded baseline heat-capacity step, set by the buried-apolar surface). This is the fold-class enthalpy expectation (proteins of the same fold share it), not an independent per-protein calorimetric measurement; peak sharpness is set by the cooperativity  $\kappa = \Delta H_{\text{vH}}/\Delta H_{\text{cal}}$  (shown on the DSC thermogram), not by  $\Delta C_p$ . It forms **a weak gel** (predicted  $G' \approx 1.8 \text{ kPa}$  at 14% w/v; least-gelling concentration  $\approx 6.8\%$  w/v). Isoelectric point **pl  $\approx 4.6$** : minimum solubility (and the precipitation pH for purification) sits here, with solubility recovering toward the acid and alkaline extremes. This is the sequence-computed **isoionic** point; a measured IEP can differ by a few tenths of a pH unit (buried-carboxyl offset, bound-ion screening, surface-residue accessibility), so use it to **rank the precipitation-pH window, not as an absolute** — confirm against a bench IEP before an isoelectric precipitation. Interfacial and hydration behaviour: surface-activity index 0.84 (emulsification/foaming); water-holding 1.3 / oil-holding 0.5. **Enabled applications** — read from this computed functional profile (open use-discovery, not a substitution list): Clear high-protein beverage; Cold-set / low-temperature-processed product.

**soy glycinin 11S**

| Key readout            | Value     |
|------------------------|-----------|
| Unfolding temp Td      | 94 ± 9 °C |
| Radius of gyration Rg  | 4.48 nm   |
| Hydrodynamic radius Rh | 5.78 nm   |
| Rg / Rh                | 0.78      |
| Unfolding enthalpy ΔH  | 15.8 J/g  |
| Isoelectric point pl   | 4.93      |
| Gel modulus G'         | 3.13 kPa  |
| Emulsification (EAI)   | 0.82      |
| Water-holding (WHC)    | 1.73      |
| Oil-holding (OHC)      | 0.37      |

The assembled particle has a predicted radius of gyration  $R_g = 4.5 \text{ nm}$  (SAXS observable) and a hydrodynamic radius  $R_h = 5.8 \text{ nm}$  predicted from  $R_g$  for a compact globule ( $R_g/R_h = \sqrt{(3/5)} \approx 0.78$  by geometry — a modelled DLS estimate, not a second measurement; confirm independently by DLS), placing it in a large size range (aggregation-prone on heating). Aggregation propensity (size-driven): **HIGH** (score 0.89) — estimated from the predicted particle size, deterministically and with no molecular-dynamics step. A pre-instrument indicator that complements DLS/SEC-MALS. Predicted denaturation temperature  $T_d = 94 \pm 9 \text{ }^\circ\text{C}$  (recognised CUPIN fold, reported as a band — the value is deterministic, the  $\pm$  is the fold-family's disclosed uncertainty) — thermostable for a plant storage globulin. Unfolding enthalpy  $\Delta H \approx 15.8 \text{ J/g}$  with  $\Delta C_p \approx 27.6 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$  (the folded→unfolded baseline heat-capacity step, set by the buried-apolar surface). This is the fold-class enthalpy expectation (proteins of the same fold share it), not an independent per-protein calorimetric measurement; peak sharpness is set by the cooperativity  $\kappa = \Delta H_{\text{vH}}/\Delta H_{\text{cal}}$  (shown on the DSC thermogram), not by  $\Delta C_p$ . It forms **a moderate gel** (predicted  $G' \approx 3.1 \text{ kPa}$  at 14% w/v; least-gelling concentration  $\approx 9.7\%$  w/v). Isoelectric point **pl  $\approx 4.9$** : minimum solubility (and the precipitation pH for purification) sits here, with solubility recovering toward the acid and alkaline extremes. This is the sequence-computed **isoionic** point; a measured IEP can differ by a few tenths of a pH unit (buried-carboxyl offset, bound-ion screening, surface-residue accessibility), so use it to **rank the precipitation-pH window, not as an absolute** — confirm against a bench IEP before an isoelectric precipitation. Interfacial and hydration behaviour: surface-activity index 0.82 (emulsification/foaming); water-holding 1.7 / oil-holding 0.4. **Enabled applications** — read from this computed functional profile (open use-discovery, not a substitution list): Structural bakery / extruded product.

**alpha-lactalbumin**

| Key readout            | Value     |
|------------------------|-----------|
| Unfolding temp Td      | 64 ± 5 °C |
| Radius of gyration Rg  | 1.57 nm   |
| Hydrodynamic radius Rh | 2.03 nm   |
| Rg / Rh                | 0.78      |
| Unfolding enthalpy ΔH  | 15.8 J/g  |

|                      |          |
|----------------------|----------|
| Isoelectric point pI | 4.59     |
| Gel modulus G'       | 5.74 kPa |
| Emulsification (EAI) | 0.68     |
| Water-holding (WHC)  | 1.54     |
| Oil-holding (OHC)    | 0.4      |

The assembled particle has a predicted radius of gyration  $R_g = 1.6 \text{ nm}$  (SAXS observable) and a hydrodynamic radius  $R_h = 2.0 \text{ nm}$  predicted from  $R_g$  for a compact globule ( $R_g/R_h = \sqrt{(3/5)} \approx 0.78$  by geometry — a modelled DLS estimate, not a second measurement; confirm independently by DLS), placing it in a small, stable size range. Aggregation propensity (size-driven): **LOW** (score 0.00) — estimated from the predicted particle size, deterministically and with no molecular-dynamics step. A pre-instrument indicator that complements DLS/SEC-MALS. Predicted denaturation temperature  $T_d = 64 \pm 5 \text{ }^\circ\text{C}$  (recognised ALPHA\_LAC fold, reported as a band — the value is deterministic, the  $\pm$  is the fold-family's disclosed uncertainty) — thermolabile for a whey Ca-binding protein. Unfolding enthalpy  $\Delta H \approx 15.8 \text{ J/g}$  with  $\Delta C_p \approx 7.1 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$  (the folded→unfolded baseline heat-capacity step, set by the buried-apolar surface). This is the fold-class enthalpy expectation (proteins of the same fold share it), not an independent per-protein calorimetric measurement; peak sharpness is set by the cooperativity  $\kappa = \Delta H_{\text{vh}}/\Delta H_{\text{cal}}$  (shown on the DSC thermogram), not by  $\Delta C_p$ . It forms a **moderate gel** (predicted  $G' \approx 5.7 \text{ kPa}$  at 14% w/v; least-gelling concentration  $\approx 7.9\% \text{ w/v}$ ). Isoelectric point **pI  $\approx 4.6$** : minimum solubility (and the precipitation pH for purification) sits here, with solubility recovering toward the acid and alkaline extremes. This is the sequence-computed **isoionic** point; a measured IEP can differ by a few tenths of a pH unit (buried-carboxyl offset, bound-ion screening, surface-residue accessibility), so use it to **rank the precipitation-pH window, not as an absolute** — confirm against a bench IEP before an isoelectric precipitation. Interfacial and hydration behaviour: surface-activity index 0.68 (emulsification/foaming); water-holding 1.5 / oil-holding 0.4. **Enabled applications** — read from this computed functional profile (open use-discovery, not a substitution list): Clear high-protein beverage.

## Protein bitter-peptide stack — where it turns bitter, and the sequence causing it

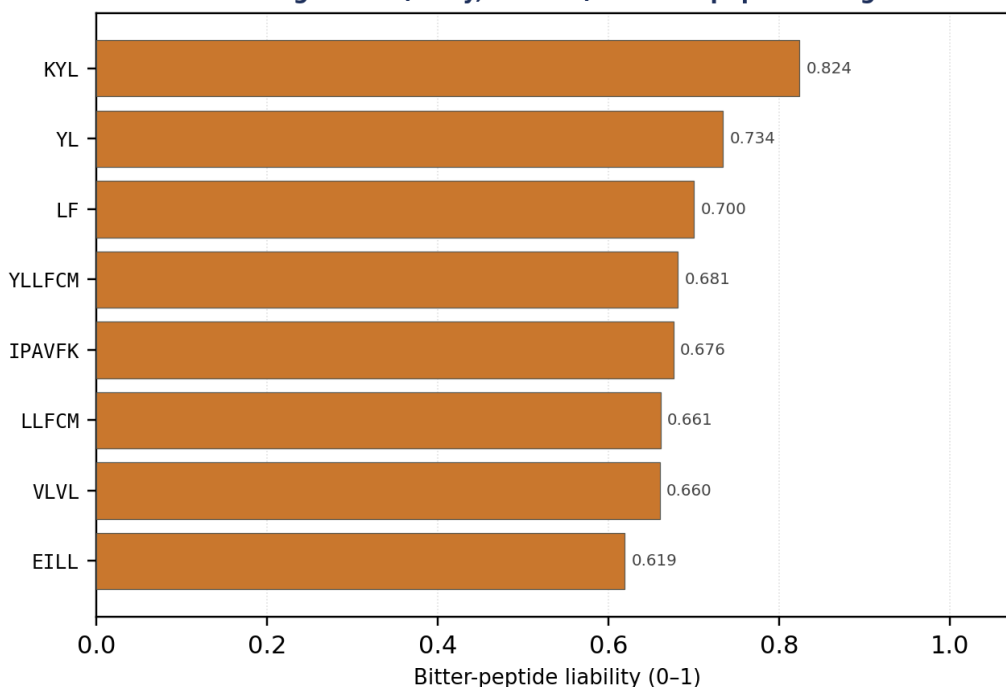
From the protein sequence alone — no samples: the in-silico digest, the bitter peptides ranked by liability, and the degree-of-hydrolysis taste arc showing at what stage of processing / digestion the bitterness emerges. **How to read this stack.** This is a RANK-ORDER screen, not a bitter / not-bitter verdict. A large share of any tryptic digest carries the hydrophobic C-terminal motifs the bitter receptors read, so the screen deliberately casts a wide net and over-flags in absolute terms (a trained panel confirms far fewer than are listed). Use it to compare regions and rank sequences against each other, and to see the degree-of-hydrolysis stage where bitterness is most likely to appear. Confirm any flagged fraction on a panel before you act on it — the value here is knowing where to look first, not a count to take literally.

**beta-lactoglobulin (whey, Bos d 5)** — 162 aa · 235 digest peptides · 122 bitter (51.9% of scanned) · 3 strong culprit(s)

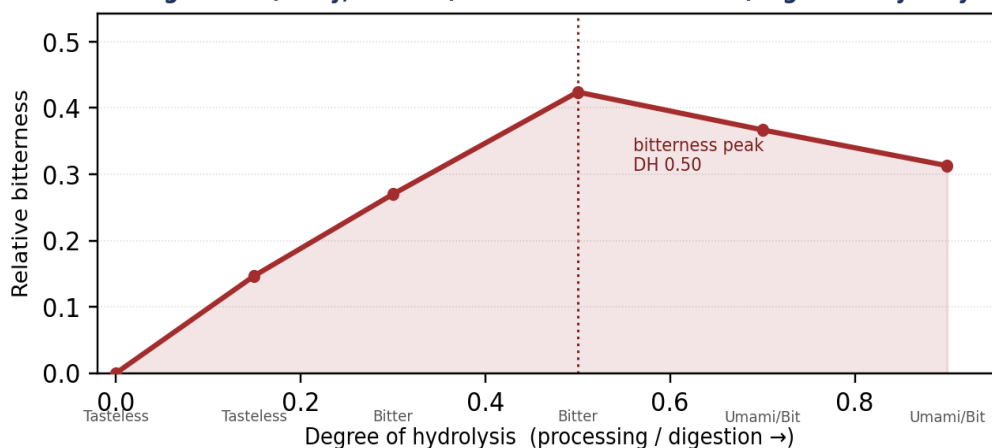
| # | Bitter peptide | Liability (0-1) |
|---|----------------|-----------------|
| 1 | KYL            | 0.824           |
| 2 | YL             | 0.734           |
| 3 | LF             | 0.700           |
| 4 | YLLFCM         | 0.681           |
| 5 | IPAVFK         | 0.676           |
| 6 | LLFCM          | 0.661           |
| 7 | VLVL           | 0.660           |
| 8 | EILL           | 0.619           |

Each peptide above is scored from its sequence alone — no measured taste value is supplied — so the screen ranks fragments that have never been tasted before. The 0–1 liability is an intrinsic per-molecule bitter potency (detection-threshold basis), independent of how much of the peptide is present; a low-abundance fragment can therefore rank above a more abundant one. **Debitter map — strip vs. protect.** The bitter stack above, read against the ACE / DPP-4 inhibitory peptides found in the SAME digest, so removing bitterness never silently discards the bioactivity the hydrolysate is sold on. **STRIP** — bitter, no bioactive motif (remove freely): YLLFCM IPAVFK LLFCM VLVL EILL **PROTECT** — bitter AND ACE / DPP-4 inhibitory (mask or encapsulate, do NOT strip; keeps the bioactivity): KYL YL LF **Metabolic / GLP-1 (incretin).** DPP-4 degrades active GLP-1, so the DPP-4-inhibitory peptides in this digest **prolong** the incretin (a glycaemic / satiety benefit): VL YLL LL GL. GLP-1 **secretagogue** motifs (PepT1 / CaSR di- and tripeptides with an N-terminal aromatic or branched residue): FCM FDK LF LL LLF LQK · glutamine 5.6% (a documented secretagogue). structure-derived motif screen: DPP-4-inhibitory peptides prolong active GLP-1 (incretin, glycaemic); PepT1/CaSR di-/tripeptides with an N-terminal aromatic/branched residue trigger L-cell GLP-1 secretion; glutamine is a documented secretagogue. RELATIVE motif potential only - absolute glycaemic index, in-vivo satiety and plasma hormone levels require the physical food.

## beta-lactoglobulin (whey, Bos d 5) — bitter-peptide danger list



## beta-lactoglobulin (whey, Bos d 5) — when it turns bitter (degree-of-hydrolysis arc)



**Degree-of-hydrolysis taste arc:** DH 0.00 **Tasteless** → DH 0.15 **Tasteless** → DH 0.30 **Bitter** → DH 0.50 **Bitter** → DH 0.70 **Umami/Bitter** → DH 0.90 **Umami/Bitter**

**Predicted denaturation temperature:** 73 °C — recognised LIPOCALIN fold — provisional Td (own physics, not cupin) (separation: isoelectric precipitation)

**Load-bearing coupling & thermal headroom. 29 buried intra-chain contact pair(s)** (10 aromatic + 18 basic residues) that stiffen the fold — the structural grip behind its heat stability (counted from the sequence; a 3-D structure confirms the buried contacts); predicted **failure mode: HYDRATION COLLAPSE** — weakest link is the hydration shell.

**soy glycinin 11S (plant isolate)** — 476 aa · 432 digest peptides · 109 bitter (25.2% of scanned) · 6 strong culprit(s)

| # | Bitter peptide | Liability (0-1) |
|---|----------------|-----------------|
| 1 | FYL            | 0.999           |
| 2 | FY             | 0.890           |
| 3 | LKY            | 0.801           |
| 4 | WMY            | 0.779           |
| 5 | IY             | 0.745           |
| 6 | YL             | 0.734           |
| 7 | KFL            | 0.696           |

| # | Bitter peptide | Liability (0-1) |
|---|----------------|-----------------|
| 8 | <b>RFY</b>     | 0.696           |

Each peptide above is scored from its sequence alone — no measured taste value is supplied — so the screen ranks fragments that have never been tasted before. The 0–1 liability is an intrinsic per-molecule bitter potency (detection-threshold basis), independent of how much of the peptide is present; a low-abundance fragment can therefore rank above a more abundant one.

**Debitter map — strip vs. protect.** The bitter stack above, read against the ACE / DPP-4 inhibitory peptides found in the SAME digest, so removing bitterness never silently discards the bioactivity the hydrolysate is sold on.

**STRIP** — bitter, no bioactive motif (remove freely): FYL LKY WMY KFL RFY

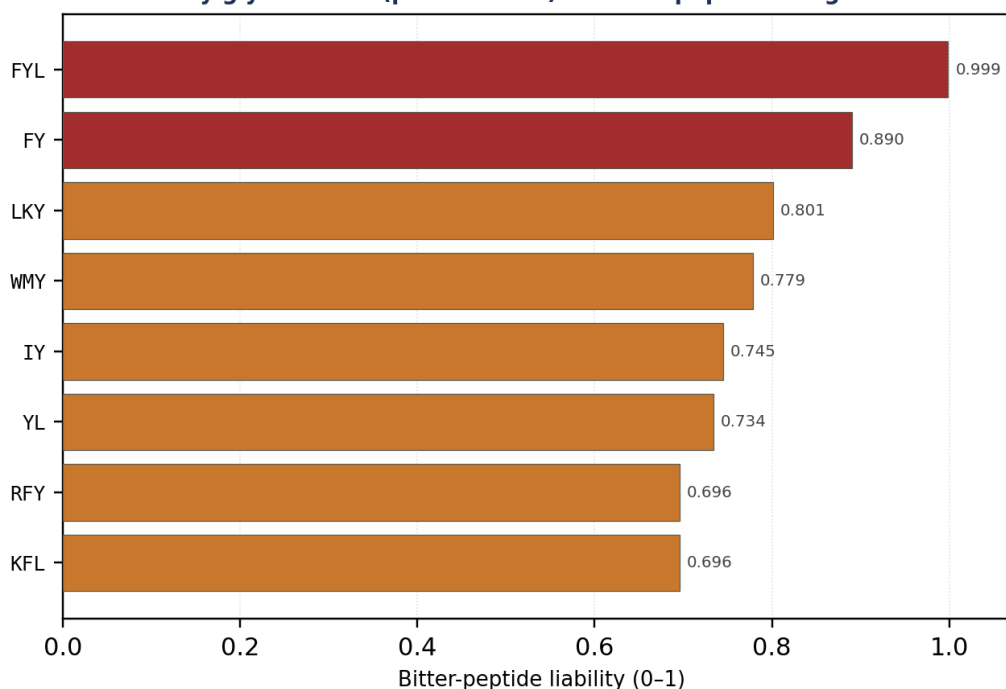
**PROTECT** — bitter AND ACE / DPP-4 inhibitory (mask or encapsulate, do NOT strip; keeps the bioactivity): FY IY YL

**Metabolic / GLP-1 (incretin).** DPP-4 degrades active GLP-1, so the DPP-4-inhibitory peptides in this digest **prolong** the incretin (a glycaemic / satiety benefit): IY. GLP-1 **secretagogue** motifs (PepT1 / CaSR di- and tripeptides with an N-terminal aromatic or branched residue): FL FY FYL IY LK LKY · glutamine 10.1% (a documented secretagogue).

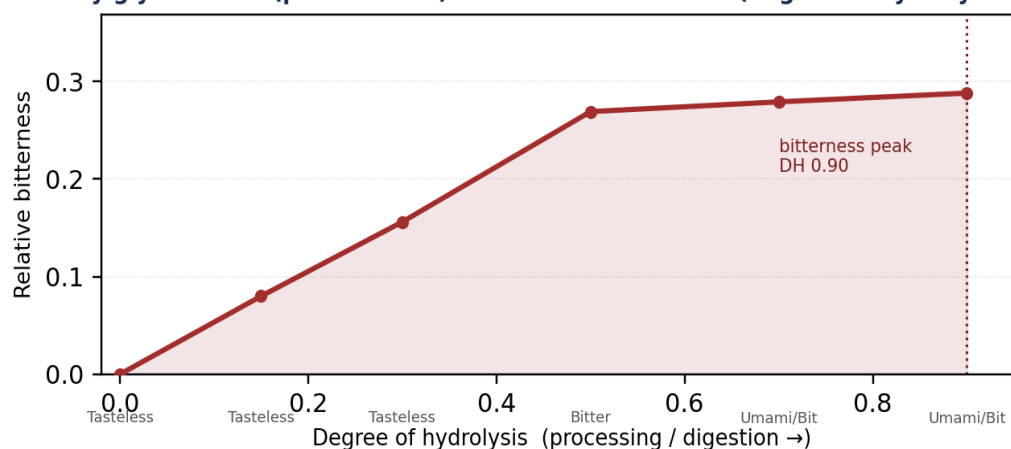
structure-derived motif screen: DPP-4-inhibitory peptides prolong active GLP-1 (incretin, glycaemic); PepT1/CaSR di-/tripeptides with an N-terminal aromatic/branched residue trigger L-cell GLP-1 secretion; glutamine is a documented secretagogue.

RELATIVE motif potential only - absolute glycaemic index, in-vivo satiety and plasma hormone levels require the physical food.

#### soy glycinin 11S (plant isolate) — bitter-peptide danger list



#### soy glycinin 11S (plant isolate) — when it turns bitter (degree-of-hydrolysis arc)



**Degree-of-hydrolysis taste arc:** DH 0.00 **Tasteless** → DH 0.15 **Tasteless** → DH 0.30 **Tasteless** → DH 0.50 **Bitter** → DH 0.70 **Umami/Bitter** → DH 0.90 **Umami/Bitter**

**Predicted denaturation temperature:** 94 °C — denatures within the extrusion window ( $\leq$  die temp) (separation: isoelectric precipitation)

**Process conditions (as submitted):** S-S/protomer 1.0

**Load-bearing coupling & thermal headroom. 72 buried cross-subunit contact pair(s)** (35 aromatic + 51 basic residues) that hold the assembly together — the structural grip behind its heat stability (counted from the sequence; a 3-D structure confirms the buried contacts); predicted **failure mode: HYDRATION COLLAPSE** — weakest link is the hydration shell.

**alpha-lactalbumin (whey, Ca-holo)** — 123 aa · 166 digest peptides · 71 bitter (42.8% of scanned) · 4 strong culprit(s)

| # | Bitter peptide | Liability (0-1) |
|---|----------------|-----------------|
| 1 | <b>PEW</b>     | 0.787           |
| 2 | <b>KIL</b>     | 0.767           |
| 3 | <b>IW</b>      | 0.752           |
| 4 | <b>WL</b>      | 0.701           |
| 5 | <b>IWCK</b>    | 0.665           |
| 6 | <b>LDQWL</b>   | 0.549           |
| 7 | <b>WLAHK</b>   | 0.533           |
| 8 | <b>ILDK</b>    | 0.492           |

Each peptide above is scored from its sequence alone — no measured taste value is supplied — so the screen ranks fragments that have never been tasted before. The 0-1 liability is an intrinsic per-molecule bitter potency (detection-threshold basis), independent of how much of the peptide is present; a low-abundance fragment can therefore rank above a more abundant one.

**Debitter map — strip vs. protect.** The bitter stack above, read against the ACE / DPP-4 inhibitory peptides found in the SAME digest, so removing bitterness never silently discards the bioactivity the hydrolysate is sold on.

**STRIP** — bitter, no bioactive motif (remove freely): KIL IWCK LDQWL WLAHK ILDK

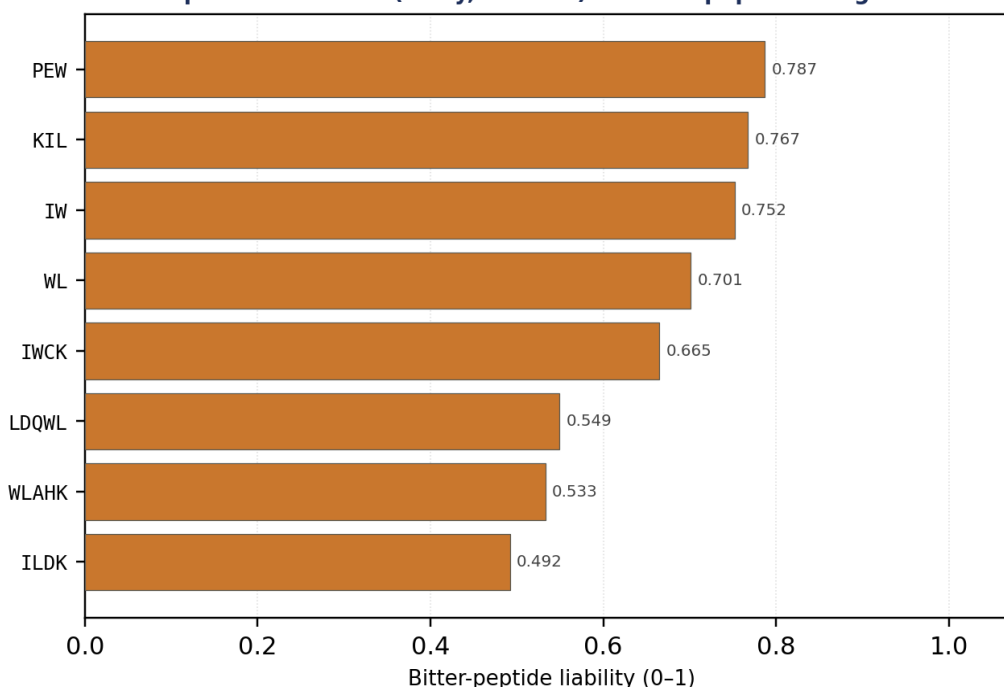
**PROTECT** — bitter AND ACE / DPP-4 inhibitory (mask or encapsulate, do NOT strip; keeps the bioactivity): PEW IW WL

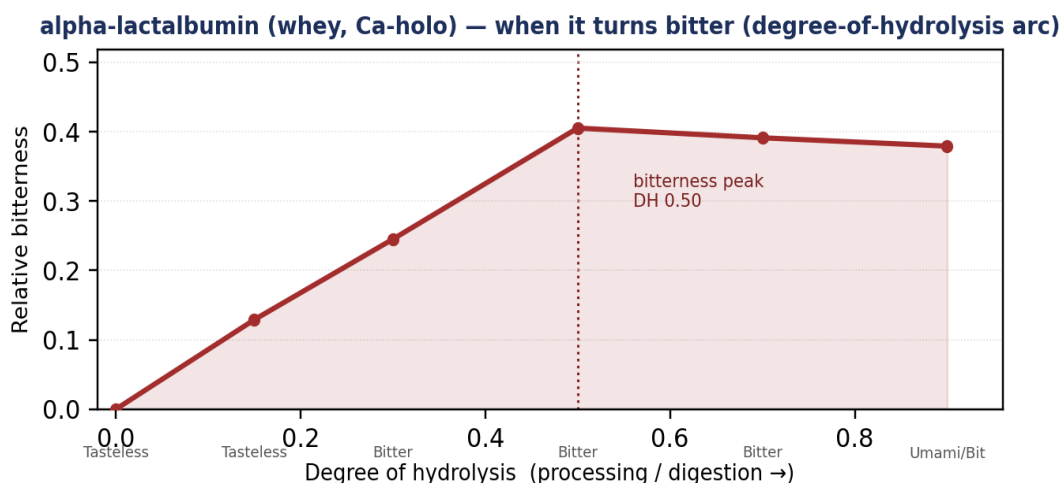
**Metabolic / GLP-1 (incretin).** DPP-4 degrades active GLP-1, so the DPP-4-inhibitory peptides in this digest **prolong** the incretin (a glycaemic / satiety benefit): IW WL IL GL. GLP-1 **secretagogue** motifs (PepT1 / CaSR di- and tripeptides with an N-terminal aromatic or branched residue): FL IL IW WL · glutamine 4.9% (a documented secretagogue).

structure-derived motif screen: DPP-4-inhibitory peptides prolong active GLP-1 (incretin, glycaemic); PepT1/CaSR di-/tripeptides with an N-terminal aromatic/branched residue trigger L-cell GLP-1 secretion; glutamine is a documented secretagogue.

RELATIVE motif potential only - absolute glycaemic index, in-vivo satiety and plasma hormone levels require the physical food.

#### alpha-lactalbumin (whey, Ca-holo) — bitter-peptide danger list



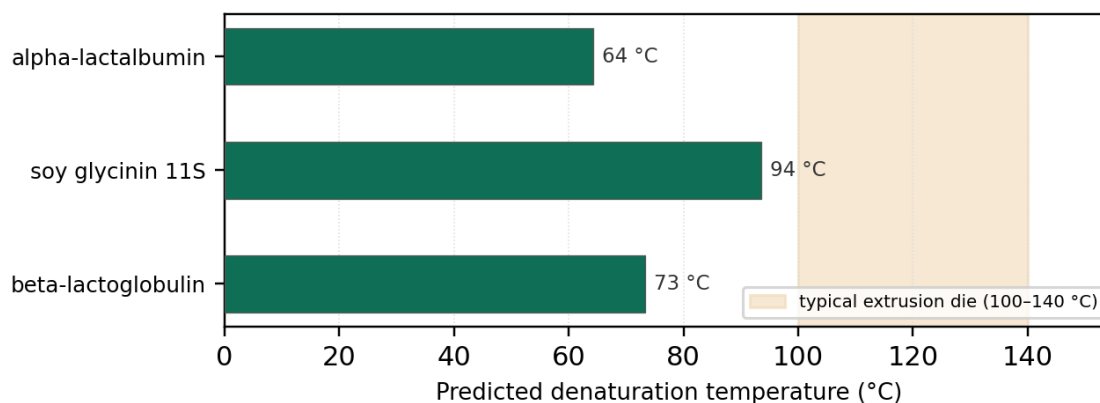


**Degree-of-hydrolysis taste arc:** DH 0.00 **Tasteless** → DH 0.15 **Tasteless** → DH 0.30 **Bitter** → DH 0.50 **Bitter** → DH 0.70 **Bitter** → DH 0.90 **Umami/Bitter**

**Predicted denaturation temperature:** 64 °C — recognised ALPHA\_LAC fold — provisional T<sub>d</sub> (own physics, not cupin) (separation: isoelectric precipitation)

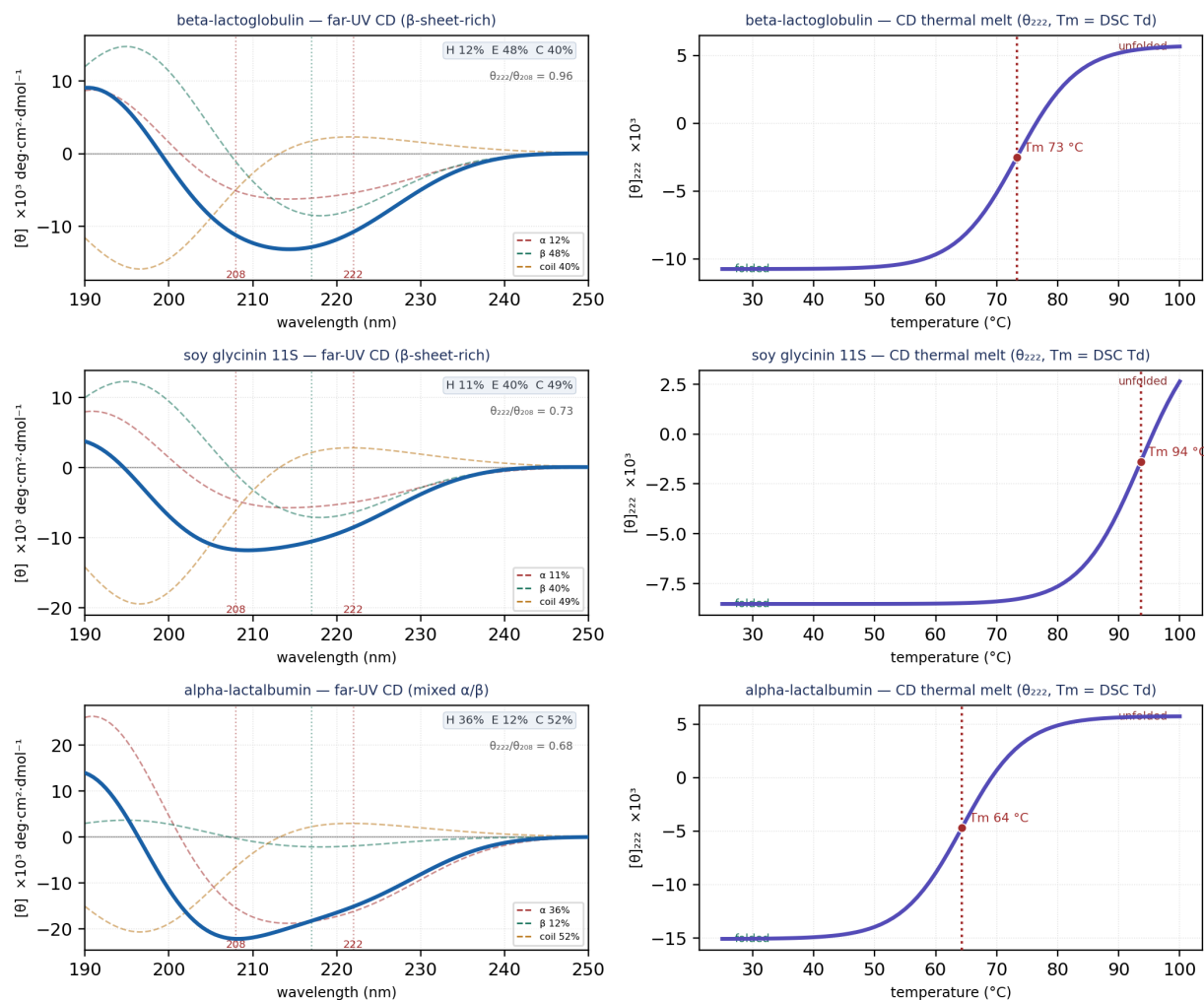
**Load-bearing coupling & thermal headroom. 28 buried intra-chain contact pair(s)** (12 aromatic + 13 basic residues) that stiffen the fold — the structural grip behind its heat stability (counted from the sequence; a 3-D structure confirms the buried contacts); predicted **failure mode: HYDRATION COLLAPSE** — weakest link is the hydration shell.

#### Protein denaturation vs the extrusion window

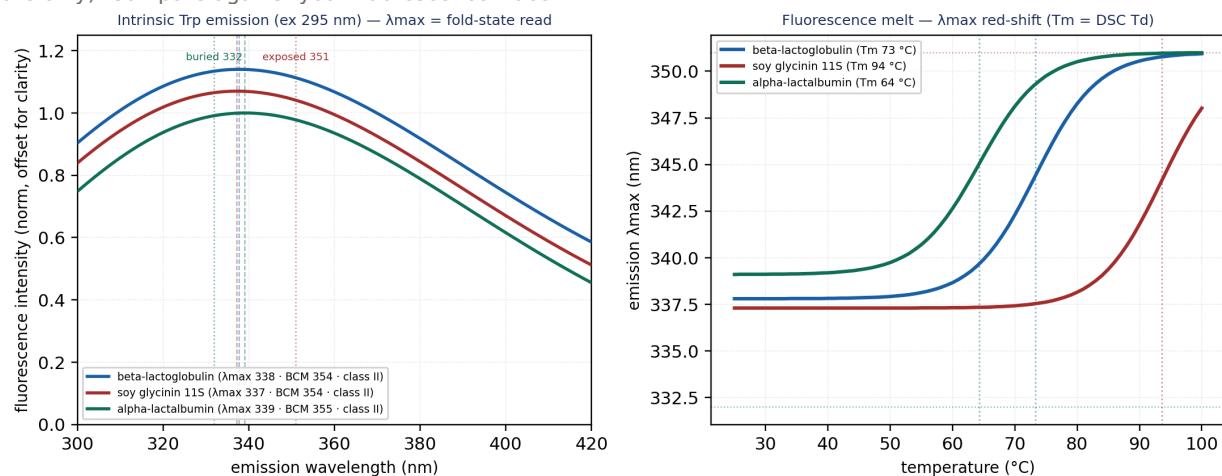


Denaturation temperatures are structure-derived estimates, blind-correlated to  $\pm 3\text{--}5$  °C across the storage-globulin core ( $\sim 84\text{--}90$  °C). The model saturates at the high end and COMPRESSES genuinely hyperstable globulins (phaseolin, lupin  $\alpha$ -conglutin, some faba legumin — really  $\sim 95\text{--}105$  °C) DOWN into the  $84\text{--}90$  °C band, where they can read  $\sim 15$  °C low. So a predicted T<sub>d</sub> in that band may be a true core value OR a compressed hyperstable outlier — confirm any hyperstable-class candidate by DSC before setting a thermal-processing (extrusion / UHT / retort) target.

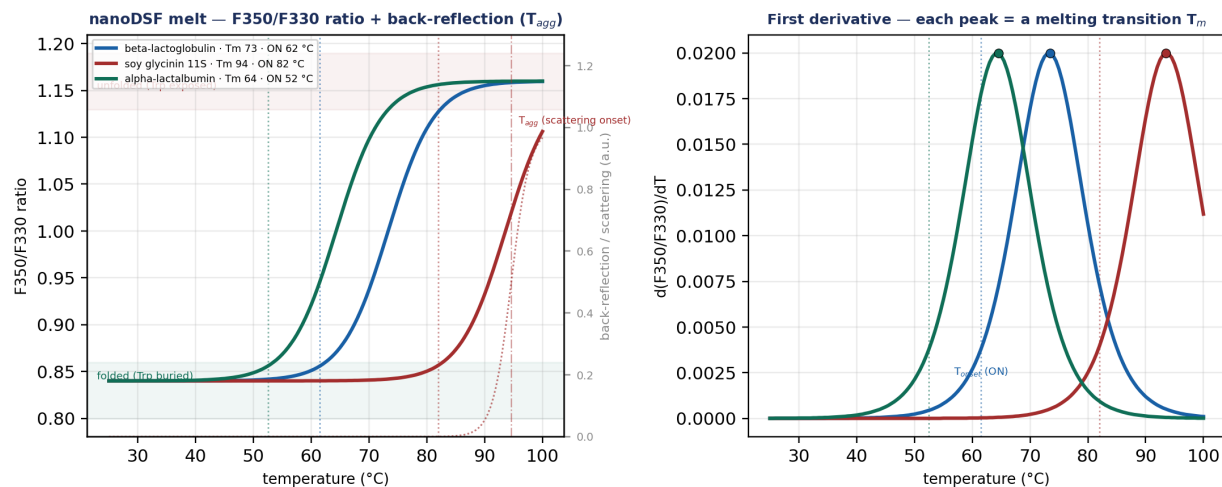
Predicted far-UV circular dichroism (CD) — the fold-state read (is it folded, and did the process unfold it?). For a RECOGNISED fold the %-secondary-structure is the fold-family canonical signature (the fold DEFINES its secondary structure), so within one fold family the per-protein distinction lives in the MELT (T<sub>d</sub>), not the spectrum shape; for a fold with a cooperative aqueous melt the CD thermal-melt midpoint tracks the DSC T<sub>d</sub> (a prolamin, gluten or intrinsically-disordered chain has no clean aqueous melt — the spectrum then reports the fold state only). A structure-derived estimate to compare against your CD trace, not a substitute for it.



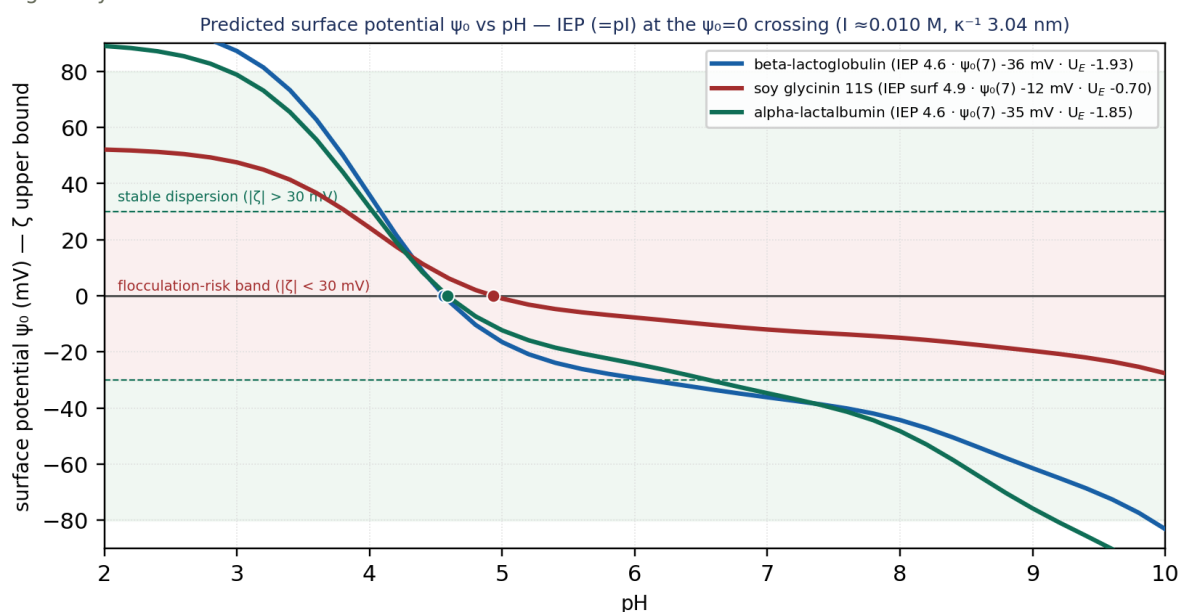
Predicted intrinsic (tryptophan) fluorescence — the tertiary-structure read: a blue emission maximum means the Trp are buried in a native core, a red-shift means they have become solvent-exposed (unfolded). Where the fold has a cooperative aqueous melt, its thermal midpoint tracks the CD and DSC estimates (a prolamin / disordered chain has none — the emission reports the fold state only). Compare against your fluorescence trace.



Predicted nanoDSF (F350/F330) thermal shift — the modern label-free way labs read T<sub>m</sub>: the intrinsic-fluorescence ratio rises as the core unfolds, and each peak of its first derivative is a melting temperature (the same midpoint as the DSC/CD estimates). A single-domain protein shows one peak; WHERE a bilobal cofactor protein is present, it melts in two transitions — an early cofactor-free lobe (S1) and a late cofactor-bound lobe (S2) — so the S1/S2 areas report the fractional cofactor saturation a single DSC scan cannot (storage globulins are cofactor-free → one transition). Compare against your Prometheus trace.

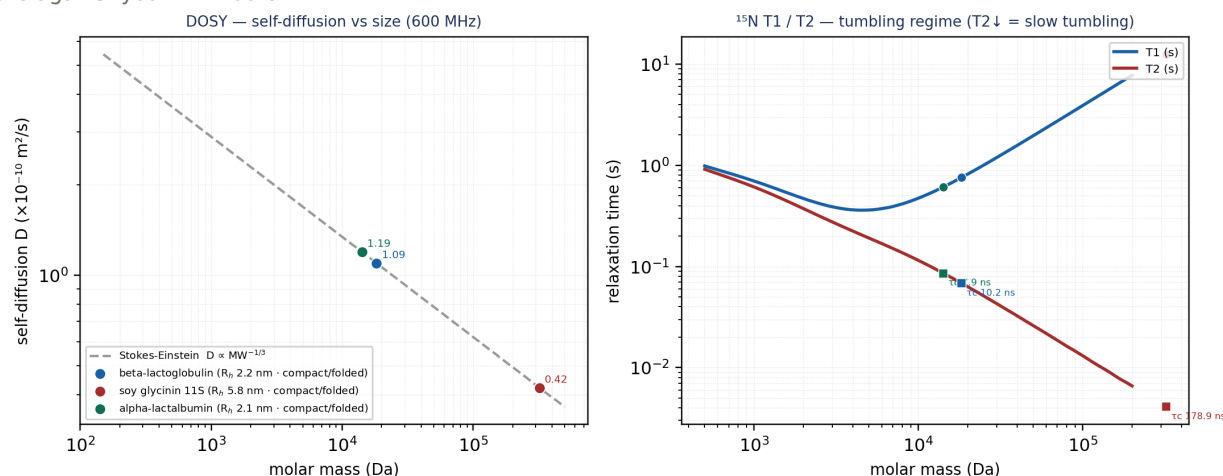


Predicted zeta potential vs pH — the colloidal-stability map: the isoelectric point ( $\zeta = 0$ ) is the pH of least stability (flocculation/precipitation), and  $|\zeta| > 30$  mV is the stable-dispersion window. Structure-derived from the charge profile; compare against your Zetasizer trace.



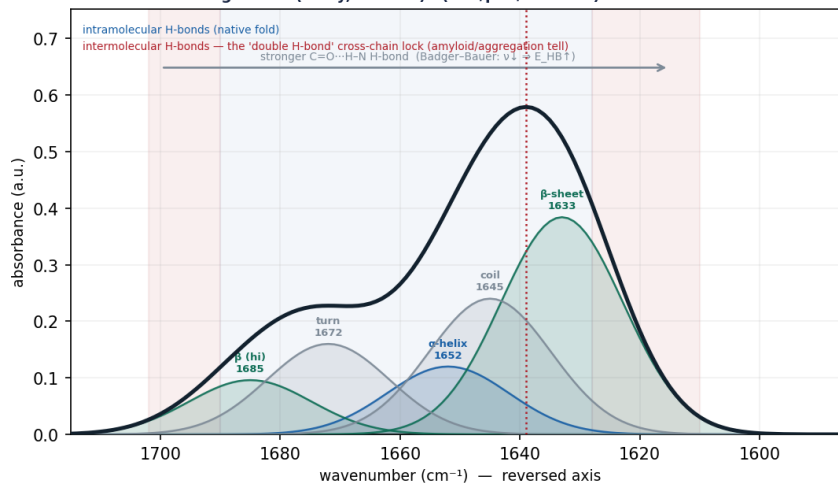
$\psi_0$  = Debye-Hückel surface potential (upper bound on shear-plane  $\zeta$ ); electrophoretic mobility  $U_E$  via Henry  $U_E = 2\epsilon\epsilon_0\zeta f(ka)/3\eta$ . Henry  $f(ka) = 1.02$  (Ohshima);  $ka = 0.68$  — near-Hückel (small particle / low salt)

Predicted NMR dossier (600 MHz) — self-diffusion (DOSY) vs size and the  $^{15}\text{N}$  T1/T2 tumbling regime, structure-derived from the hydrodynamic radius. Complements DLS/SEC-MALS on aggregation and reports the fold-tumbling timescale; an estimate to compare against your NMR data.



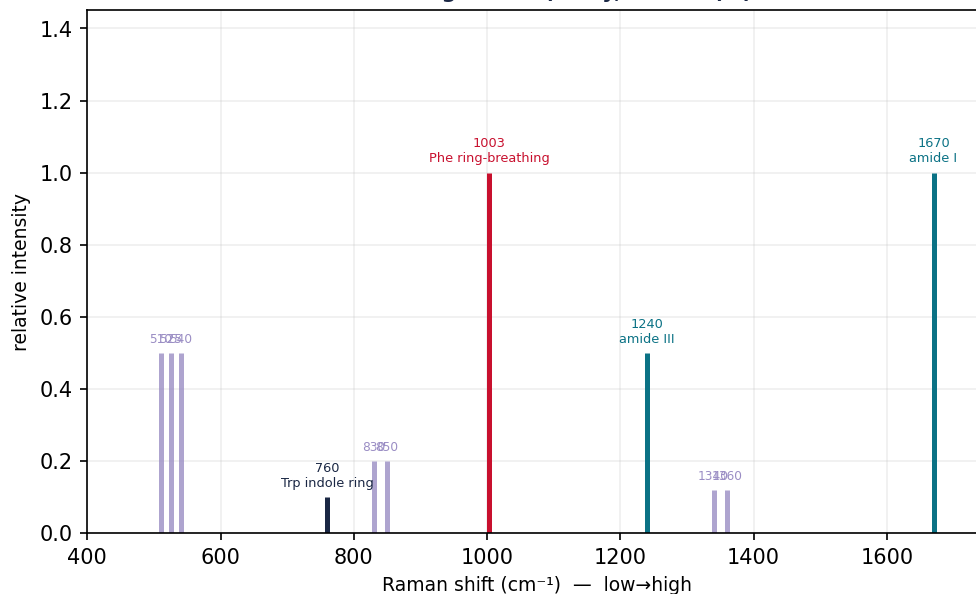
Predicted FTIR amide-I band — a vibrational read of the fold: the amide-I C=O stretch sits at a wavenumber set by the backbone H-bonding, so its position inverts to secondary structure ( $\alpha$ -helix  $\sim 1652$ ,  $\beta$ -sheet  $\sim 1633$ , intermolecular- $\beta$  aggregate  $\sim 1620$   $\text{cm}^{-1}$ ). Computed from the same fold fractions as the CD and XRD panels — three instruments (far-UV ellipticity · X-ray d-spacing · IR amide-I), one structure, cross-consistent by construction. Reversed wavenumber axis per convention; compare against your ATR-FTIR trace.

FTIR amide-I deconvolution — beta-lactoglobulin (whey, Bos d 5) ( $\alpha 12/\beta 48/\text{coil} 40$  %) · intermolecular H-bond <1% (no aggregation tell)



Predicted Raman marker-band POSITIONS from the amino-acid composition — the phenylalanine 1003  $\text{cm}^{-1}$  internal standard, the disulfide S-S rotamer triplet (510/525/540), and the tyrosine (850/830) and tryptophan (1360/1340) Fermi doublets. The burial RATIOS ( $I_{850}/I_{830}$ ,  $I_{1360}/I_{1340}$ ) and the dominant S-S rotamer are 3-D environment reads and are NOT computed from sequence. Compare positions against your Raman spectrum.

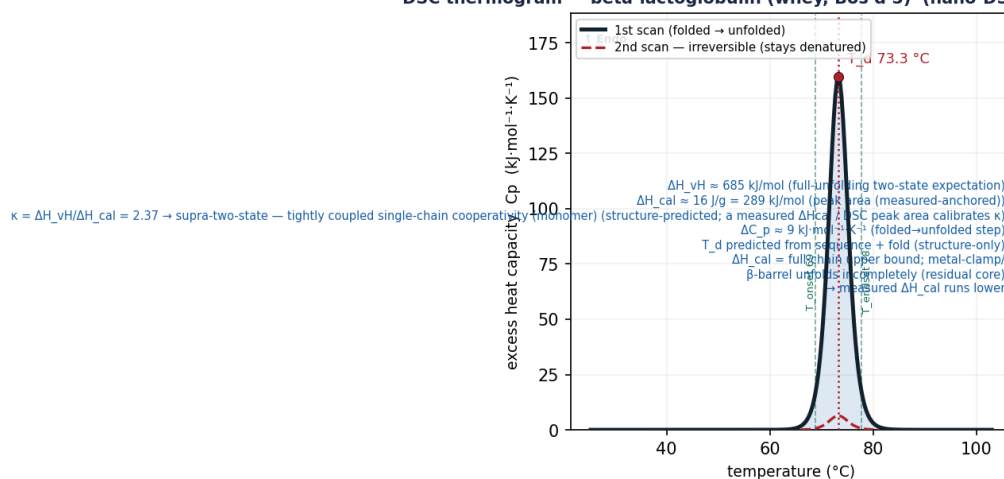
Raman marker bands — beta-lactoglobulin (whey, Bos d 5) (Phe 1003 = internal std)



S-S rotamer triplet 510/525/540 · Tyr 850/830 · Trp 1360/1340 shown as POSITIONS (violet); the burial RATIOS 1850/1830, 1360/1340 and the dominant S-S rotamer are 3-D geometry reads — not computed from sequence.

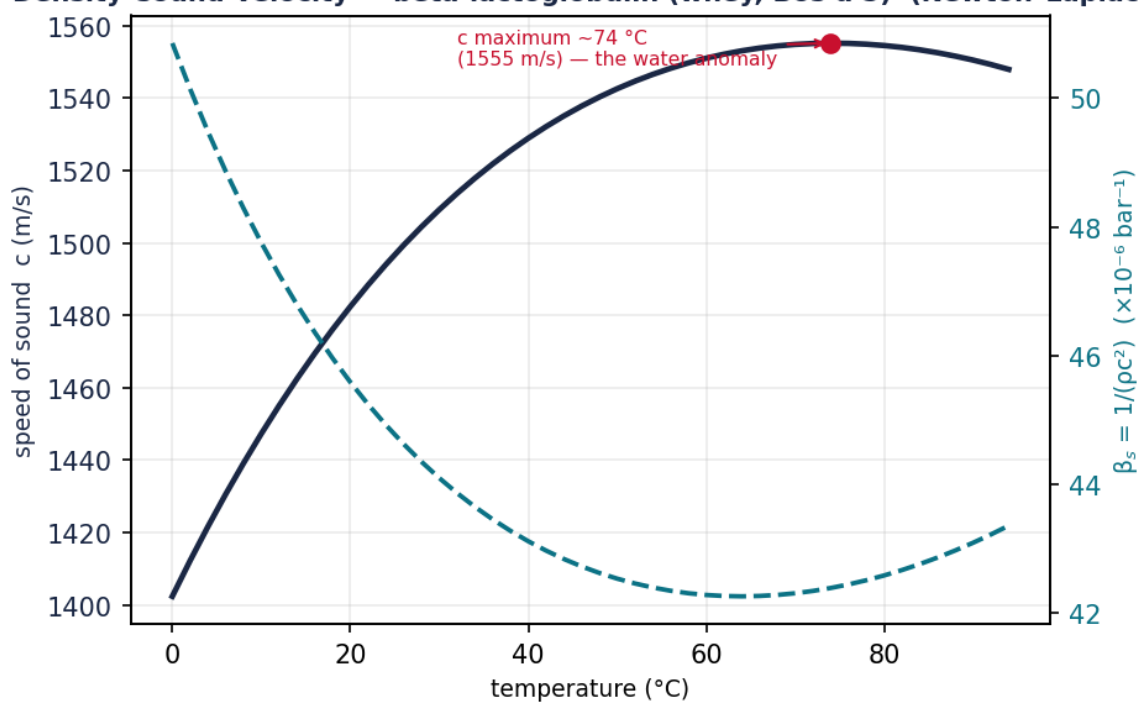
Predicted DSC thermogram (nano-DSC / MicroCal-class) — the excess-heat-capacity  $C_p$  unfolding endotherm at the **dilute-buffer** denaturation temperature  $T_d$  (the value a dilute DSC / CD / fluorescence run share; the functional table reports  $T_d$  at your **submitted moisture**: at a dilute-buffer / high-moisture input the two coincide, while a dry or concentrated input reads a few  $^{\circ}\text{C}$  lower via plasticisation — they differ only by that moisture term, never by disagreement), the cooperativity  $\kappa = \Delta H_{\text{vH}}/\Delta H_{\text{cal}}$  (the full-unfolding van't Hoff expectation against the measured calorimetric enthalpy —  $\kappa \approx 1$  two-state,  $\kappa > 1$  oligomer dissociation), and the 2nd re-scan that reports reversibility (small disulfide-clamped folds refold; large multi-domain/glyco folds stay denatured). A structure-derived estimate — compare against your DSC  $C_p$  trace.

## DSC thermogram — beta-lactoglobulin (whey, Bos d 5) (nano-DSC / MicroCal-class)



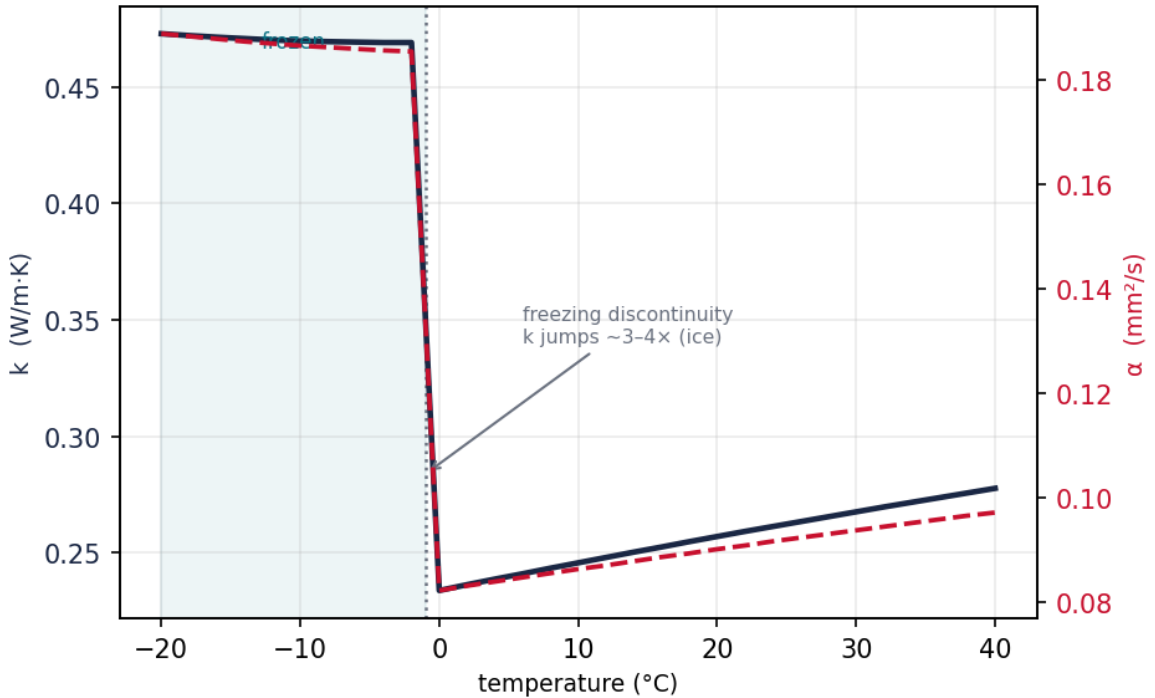
Density-sound-velocity (DSA-5000-class) — the speed-of-sound  $c(T)$  signature with its  $\sim 74^{\circ}\text{C}$  water-anomaly maximum and adiabatic compressibility  $\beta_s = 1/(\rho c^2)$ , the hydration-structure read the protein's computed partial specific volume  $\bar{v}$  and solution density (reported above) are measured against on an oscillating-U-tube densitometer.

## Density-sound-velocity — beta-lactoglobulin (whey, Bos d 5) (Newton-Laplace)



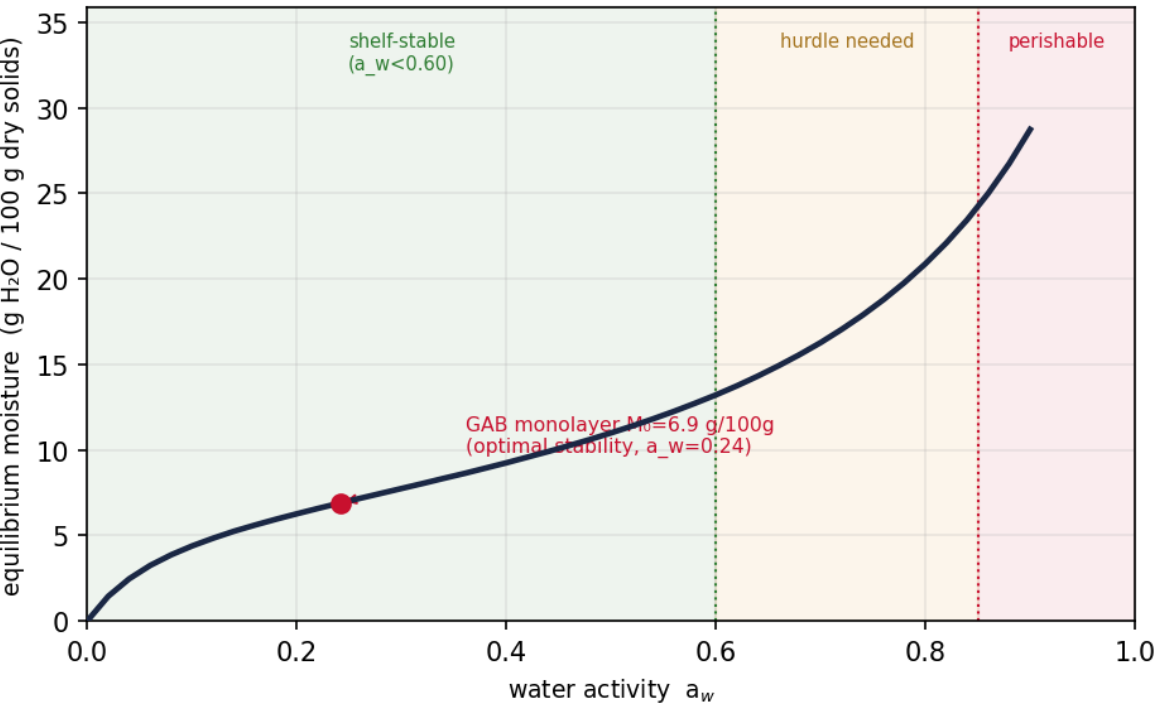
Thermal transport (Choi-Okos,  $\alpha = k/\rho c_p$ ) — thermal conductivity  $k$  and diffusivity  $\alpha$  vs temperature computed from a standard-reference protein composition (10% moisture assumed — supply moisture\_pct for your ingredient), with the freezing discontinuity where  $k$  jumps  $\sim 3\text{--}4\times$  as water $\rightarrow$ ice. The process-heat read for drying, pasteurisation and freezing operations.

Thermal transport — beta-lactoglobulin (whey, Bos d 5) (Choi-Okos,  $\alpha=k/\rho c_p$ )



Moisture-sorption isotherm (GAB) — the powder shelf-life reference: equilibrium moisture vs water activity with the GAB monolayer (optimal-stability point) and the microbial hurdle bands ( $a_w < 0.60$  shelf-stable /  $0.60\text{--}0.85$  hurdle needed /  $> 0.85$  perishable) the dried-powder glass-transition and caking read (above) sits on.

Moisture-sorption isotherm — beta-lactoglobulin (whey, Bos d 5) (GAB)



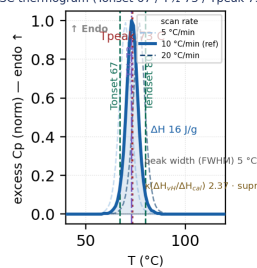
Paired-instrument diagnostics — the cross-machine reads a multi-instrument lab derives by hand, computed ONCE from the one structure so they are cross-consistent by construction.

| Diagnostic               | Cross-instrument read                                 |
|--------------------------|-------------------------------------------------------|
| Oligomer                 | $n \approx 1.0$ (native 18281.0 / subunit 18281.0 Da) |
| 2° structure (CD + FTIR) | $\alpha$ 12.0% / $\beta$ 48.0%                        |

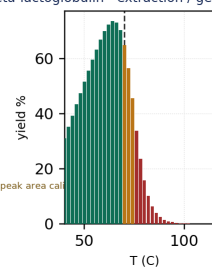
|                        |                                                                           |
|------------------------|---------------------------------------------------------------------------|
| Unfolding Td           | 73.3 °C (CD-melt = DSC; nanoDSF agrees, scan-rate shift $\approx 9.4$ °C) |
| Safe processing window | 68–249 °C (DMA Tg $\rightarrow$ TGA onset)                                |

Predicted thermal profile (DSC-style) & extraction/gelation window — per protein, from the structure-derived  $T_m$ . An estimate to compare against your DSC trace, not a substitute for it.

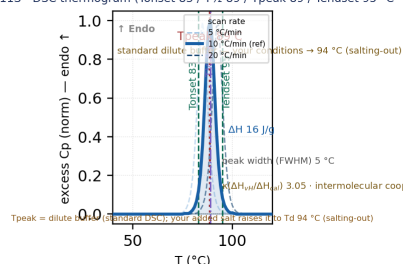
beta-lactoglobulin - DSC thermogram (Tonset 67 /  $T_{1/2}$  73 / Tpeak 73 / Tendset 80 °C · FWHM 5)



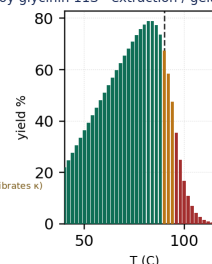
beta-lactoglobulin - extraction / gelation window



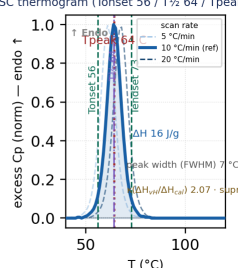
soy glycinin 11S - DSC thermogram (Tonset 83 /  $T_{1/2}$  89 / Tpeak 89 / Tendset 95 °C · FWHM 5)



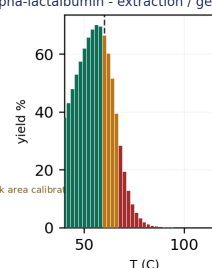
soy glycinin 11S - extraction / gelation window



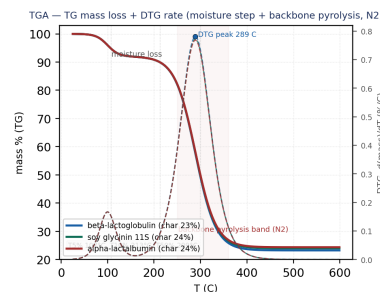
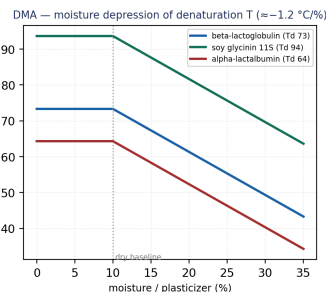
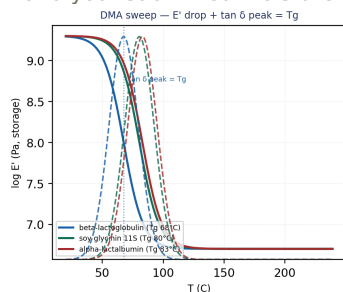
alpha-lactalbumin - DSC thermogram (Tonset 56 /  $T_{1/2}$  64 / Tpeak 64 / Tendset 73 °C · FWHM 7)



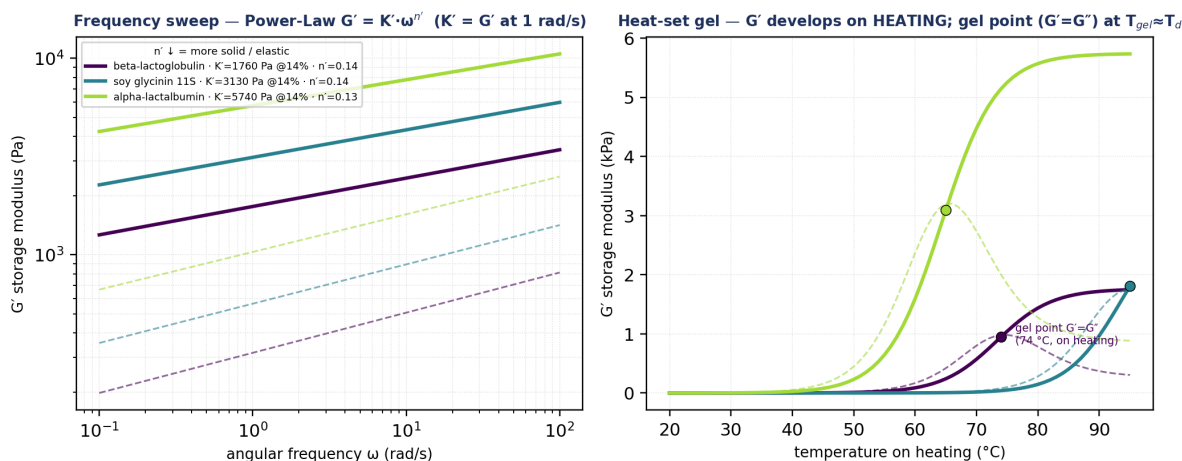
alpha-lactalbumin - extraction / gelation window



Predicted DMA (moisture plasticization) & TGA (mass-loss) — structure-derived estimates, consistent with the same  $T_d$ . The TGA decomposition onset is set by peptide-backbone pyrolysis (the weakest bond), so it sits near  $\sim 250$  °C for essentially all proteins — a shared backbone constant, not a per-protein discriminator; the DMA glass transition, by contrast, moves with composition and your submitted moisture.

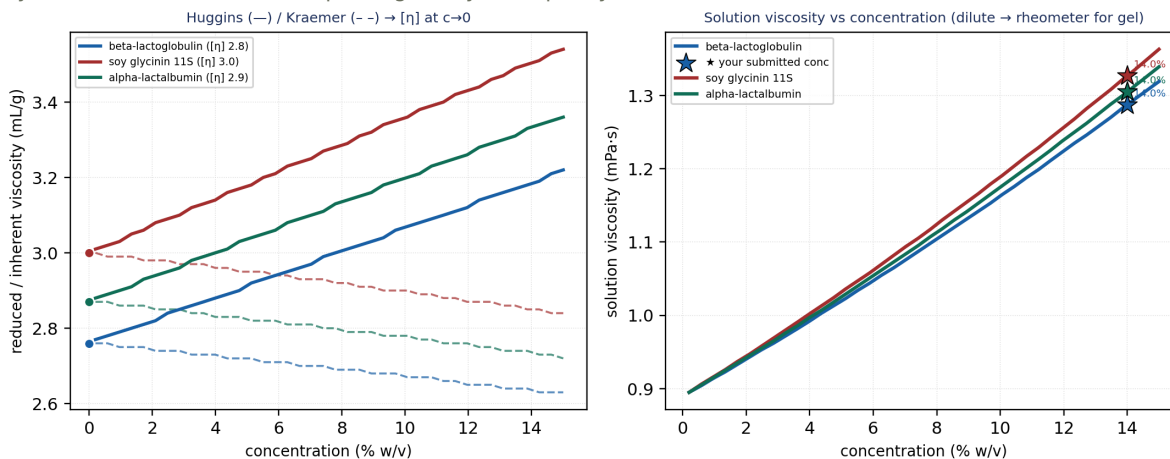


## Predicted rheology — structure-derived $G'$ frequency sweep (Power-Law $K' \cdot \omega^{n'}$ ) & heat-set gel point (on heating, $T_{gel} \approx T_d$ )

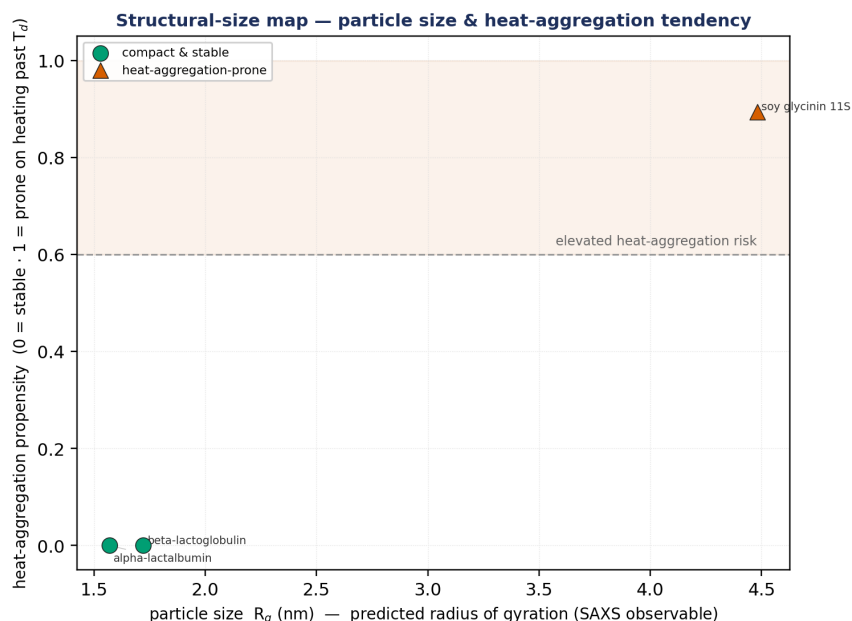


Predicted small-amplitude oscillatory  $G'$  from the structure-derived gel modulus +  $T_d$  (heat-set globulin: the network forms on HEATING at  $T_{gel} \approx T_d$ ; cooling cures it further). Absolute scale is rank-valid — supply a measured  $G'$  to calibrate it.

**Virtual rheometer.** Predicted gel storage modulus  $G'$  across the panel spans 1.8–5.7 kPa at 14% w/v: **alpha-lactalbumin** forms the strongest network, **beta-lactoglobulin** the weakest — a 3.3× spread that ranks the panel for gelling applications. The frequency exponent  $n'$  (lower = more solid-like) is set per protein by its predicted network strength. Predicted solution viscosity — the intrinsic viscosity  $[\eta]$  (Huggins/Kraemer double extrapolation) and the dilute-to-semi-dilute viscosity vs concentration for beverage/RTD processing. Accurate for compact globular proteins; the concentrated/gel regime is covered by the rheometer above. Compare against your capillary/rotational viscometer.



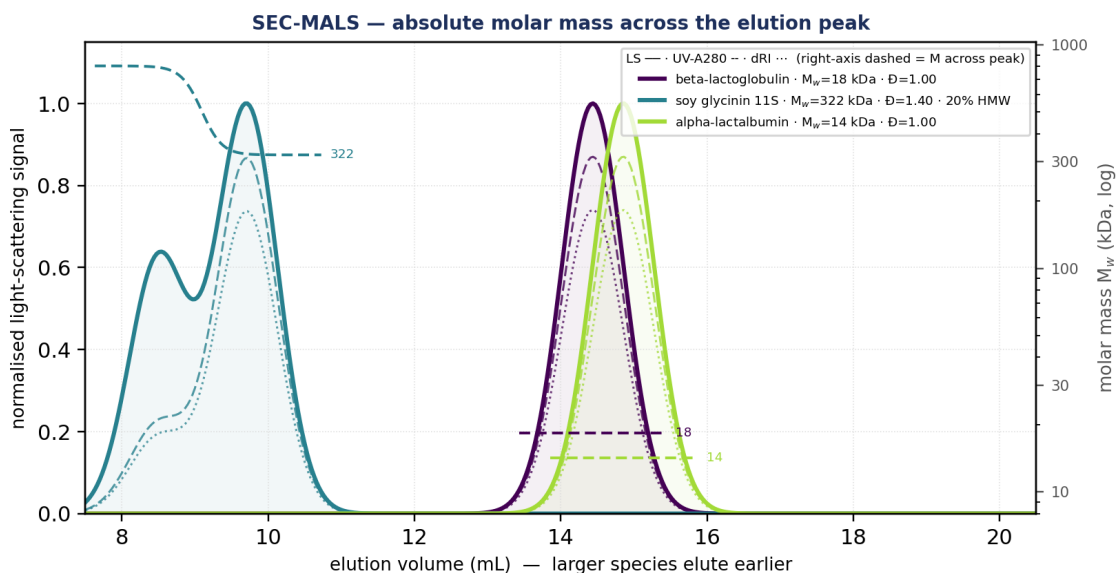
## Structural-size map — particle size and aggregation tendency



Propensity to aggregate IRREVERSIBLY on heating past  $T_d$  (the gelation route) — larger particles are more prone; the scale discriminates (small compact monomers read LOW, not all-high). The NATIVE folded-oligomer size is separately monodisperse on DLS.

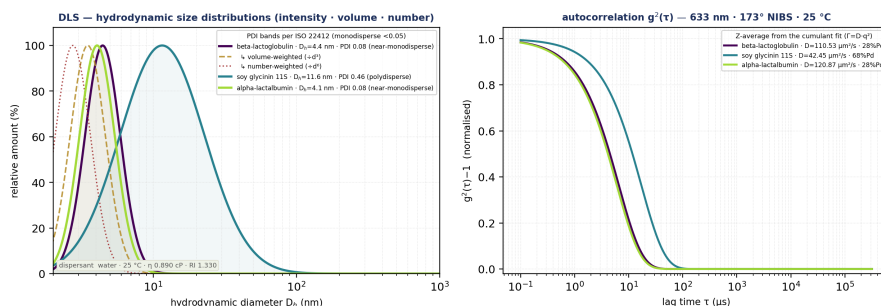
**Structural-size map.** Predicted particle size (radius of gyration  $R_g$ ) ranges 1.6–4.5 nm across this panel; particle size sorts the proteins by their expected heat-aggregation behaviour: **2 compact & stable** (small particles → resist SIZE-DRIVEN heat-aggregation: beta-lactoglobulin, alpha-lactalbumin — a free-thiol whey protein can still form a separate thiol heat-set gel, mapped in the gelation panel); **1 aggregation-prone** (large particles: soy glycinin 11S).  $R_g$  is computed from the sequence/oligomer — a structure-derived estimate to prioritise characterisation, not a substitute for a SAXS/DLS measurement. Larger particles trend toward the irreversible aggregation big globulins show on heating.

## SEC-MALS — absolute molar mass & oligomeric state (computed from sequence + oligomer)



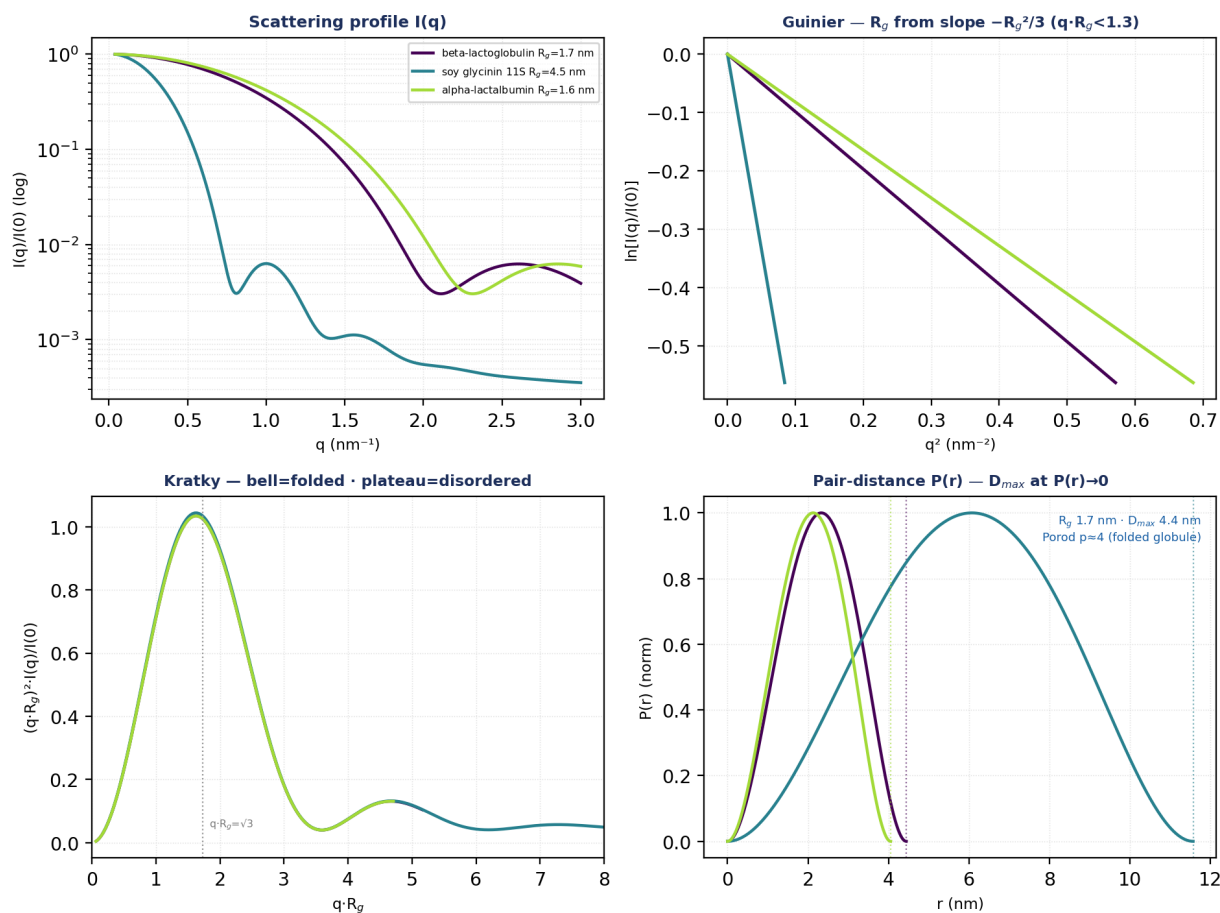
Three co-eluting detectors (LS · UV A280 · dRI);  $M_w$  structure-only.  $\bar{D} = M_w/M_n + \%HMW$  = the PREDICTED heat-aggregation propensity, the SAME structure prediction the DLS panel reads. The NATIVE state is monodisperse ( $\bar{D}=1.00$ , flat M); the %HMW shoulder is what forms on heating past  $T_d$  (the gelation route) — consistent with the DLS native read, not a native impurity. A measured trace refines it.

## DLS — hydrodynamic size distribution ( $R_h$ , monodispersity)



What a DLS run measures is a rate, not a size: the scattered-light autocorrelation decays as the particles diffuse, the instrument fits that decay to the translational self-diffusion coefficient  $D$ , then inverts the Stokes-Einstein relation ( $R_h = k_B T / 6\pi\eta D$ ) to the hydrodynamic radius — the size is read out of a dynamical measurement at the length scale set by the 633 nm laser and back-scatter geometry. Predicted self-diffusion for your proteins (25 °C, water): beta-lactoglobulin  $D=111 \mu\text{m}^2/\text{s}$  · soy glycinin 11S  $D=42 \mu\text{m}^2/\text{s}$  · alpha-lactalbumin  $D=121 \mu\text{m}^2/\text{s}$ . This is the NATIVE-state read: a single narrow peak = monodisperse folded oligomer. On heating past  $T_d$ , the size-driven mode aggregates irreversibly (the gelation route, mapped separately) → a second large-size population which, being intensity-weighted  $\propto R_h^6$ , dominates the reported Z-average even at trace levels.

## SAXS — scattering profile & Kratky fold state ( $R_g$ , compact vs disordered)



Structure-derived:  $R_g$  (Guinier),  $D_{\text{max}} = 2 \cdot R_g \cdot \sqrt{5/3}$  sphere estimate (elongated particles read higher), Porod exponent 4 = smooth folded sphere / 2 = disordered coil. Compare against your ATSAS/SASBDB analysis.

## XRD / WAXS — fold order & amorphous halos (Bragg d-spacings)

A diffractometer records the angle  $2\theta$  at which the probe X-ray ( $\text{Cu-K}\alpha$ , 1.54 Å) constructively interferes, then inverts Bragg's law ( $n\lambda = 2d \cdot \sin\theta \rightarrow d = n\lambda / 2 \cdot \sin\theta$ ) to the backbone d-spacings the solid shows:

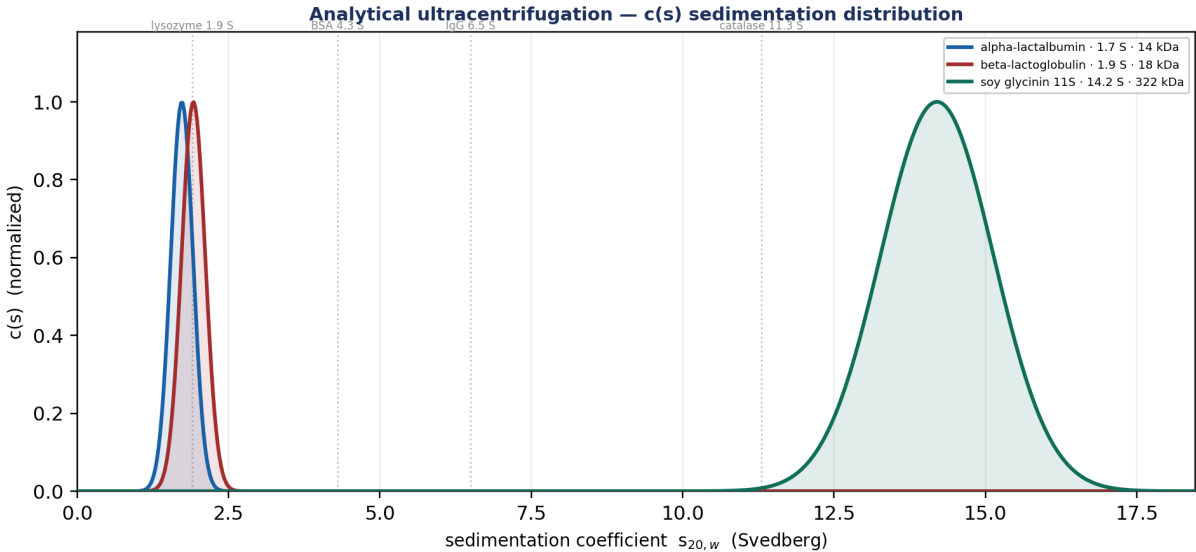
| Backbone feature                  | d-spacing (Å) | 2θ (°) |
|-----------------------------------|---------------|--------|
| β-strand H-bond spacing           | ~4.7          | ≈ 18.9 |
| Inter-sheet / inter-helix packing | ~9–10         | —      |
| Amorphous halo                    | ~4.4          | ≈ 20   |

Predicted from your proteins' secondary structure (the same %helix/%sheet the CD panel reports): beta-lactoglobulin 60% ordered SS, main halo 4.7 Å at 2θ=18.9° (β-strand H-bond spacing) · soy glycinin 11S 51% ordered SS, main halo 4.7 Å at 2θ=18.9° (β-strand H-bond spacing) · alpha-lactalbumin 48% ordered SS, main halo 4.4 Å at 2θ=20.2° (amorphous halo). Note a dried globular protein is X-ray amorphous — these are broad halos, not sharp crystalline reflections; the ordered-SS fraction is a solid-state fold-order read that complements solution SAXS.

AUC — c(s) sedimentation distribution: the sedimentation coefficient  $s_{20,w}$  (Svedberg) is structure-derived from the exact molar mass and hydrodynamic radius. Resolves oligomeric state and shape — accurate for compact globulars, ~15–20% high for elongated folds. Reference ruler:

| Standard | $s_{20,w}$ (S) |
|----------|----------------|
| Lysozyme | 1.9            |
| BSA      | 4.3            |
| IgG      | 6.5            |
| Catalase | 11.3           |

Compare against your Beckman Optima c(s) trace.



Density / sound-velocity (DSA) — compressibility & hydration

A density-sound-velocity analyser measures the solution density and the speed of sound  $c$  in one cycle, and inverts the Newton-Laplace relation ( $\beta = 1/\rho c^2$ ) to the adiabatic compressibility — a direct probe of interior packing and hydration. The partial specific volume  $\bar{v}$  is exact from the amino-acid composition; the sign of the apparent specific adiabatic compressibility follows from the fold — a compact globular interior with interior packing voids reads positive (void-dominated), an extended or disordered chain reads negative (hydration-dominated). beta-lactoglobulin  $\bar{v}$ =0.737 cm<sup>3</sup>/g, apparent adiabatic compressibility predicted positive (packing-void-dominated) · soy glycinin 11S  $\bar{v}$ =0.711 cm<sup>3</sup>/g, apparent adiabatic compressibility predicted positive (packing-void-dominated) · alpha-lactalbumin  $\bar{v}$ =0.720 cm<sup>3</sup>/g, apparent adiabatic compressibility predicted positive (packing-void-dominated). A measured sound-velocity increment fixes the magnitude and the hydration number (Gekko-Hasegawa; Kharakoz-Sarvazyan).

Thermal transport — conductivity, diffusivity & heat capacity (processing)

For process design (heating, cooling, drying, freezing): a laser-flash run measures the rear-face half-rise time after a light pulse and inverts Parker's relation ( $\alpha = 0.1388 \cdot L^2/t^{1/2}$ ) to the thermal diffusivity; the conductivity  $k$ , specific heat  $c_p$  and diffusivity  $\alpha$  follow from the composition (Choi-Okos), across moisture:

| Moisture % | $k$ (W/m·K) | $c_p$ (kJ/kg·K) | $\alpha$ (mm <sup>2</sup> /s) |
|------------|-------------|-----------------|-------------------------------|
| 10         | 0.26        | 2.3             | 0.090                         |

|    |      |     |       |
|----|------|-----|-------|
| 50 | 0.44 | 3.1 | 0.124 |
| 80 | 0.55 | 3.7 | 0.139 |

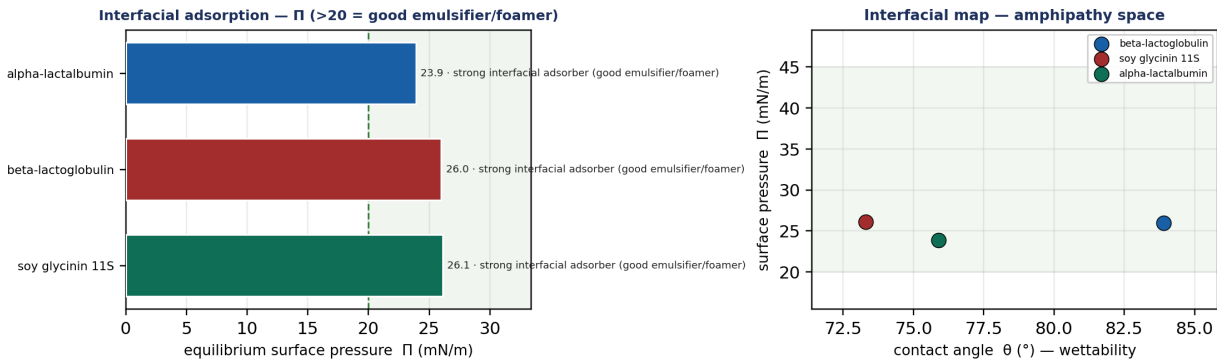
On freezing, the conductivity rises  $\approx 4.0\times$  (ice conducts  $\sim 4\times$  water) and the diffusivity  $\approx 8.1\times$  — why a frozen product cools faster than it thawed. A held-out LFA or hot-wire run refines these.

## Water activity ( $a_w$ ) — shelf-stability & food safety

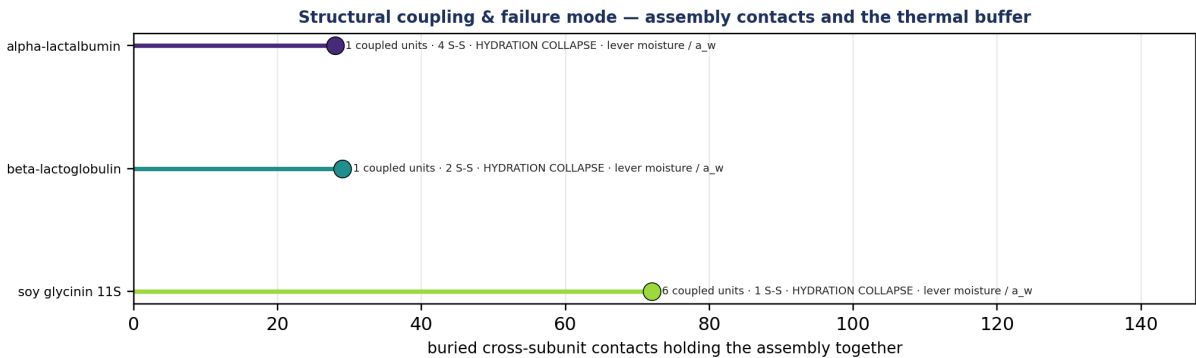
A water-activity meter reads the equilibrium relative humidity  $a_w = p/p_0$  and inverts the moisture-sorption isotherm (GAB) to the water-binding state — the single best predictor of microbial and chemical shelf-life. The monolayer  $M_0$  (the maximum-stability moisture, at the lipid-oxidation minimum) is structure-estimated from the polar-residue content: beta-lactoglobulin  $M_0=0.069$  g/g (stability optimum  $a_w\approx 0.24$ ); at 10% moisture  $a_w\approx 0.45$  · soy glycinin 11S  $M_0=0.075$  g/g (stability optimum  $a_w\approx 0.24$ ); at 10% moisture  $a_w\approx 0.40$  · alpha-lactalbumin  $M_0=0.075$  g/g (stability optimum  $a_w\approx 0.24$ ); at 10% moisture  $a_w\approx 0.40$ .

| Microbial hurdle    | aw threshold            |
|---------------------|-------------------------|
| No microbial growth | < 0.60                  |
| Molds               | < 0.70                  |
| FDA shelf-stable    | $\leq 0.85$             |
| Bacterial pathogens | > 0.90 (S. aureus 0.86) |

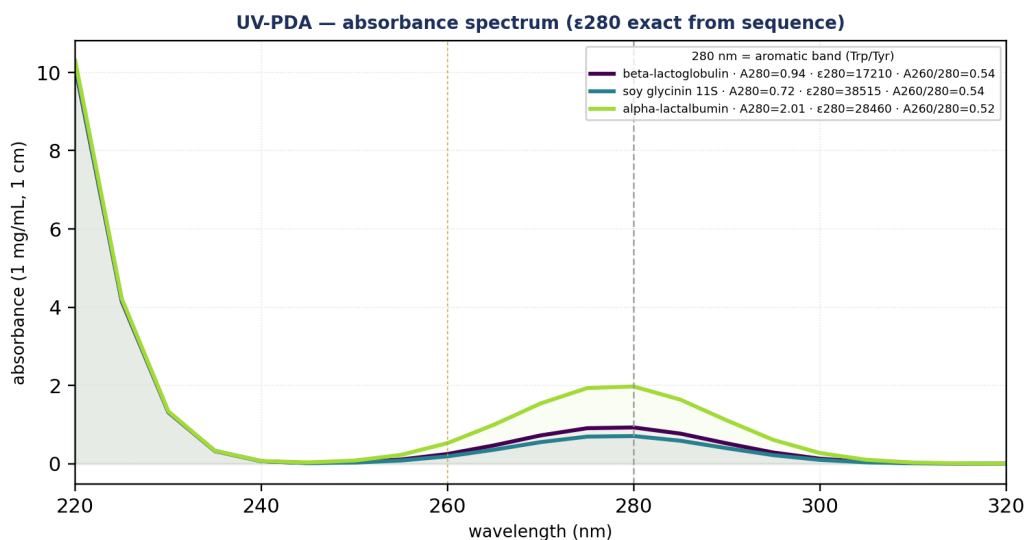
A measured sorption isotherm refines  $M_0$  (ICMSF / Beuchat / FDA thresholds). Predicted interfacial behaviour (pendant-drop / tensiometer) — the equilibrium surface pressure  $\Pi$  each protein develops at the air–water interface ( $\Pi > \sim 20$  mN/m marks a strong adsorber  $\rightarrow$  a good emulsifier/foamer), and the contact-angle (wettability) vs  $\Pi$  map placing each protein in amphipathy space. Structure-derived from the hydrophobic-moment profile; compare against your tensiometer.



Predicted structural coupling & failure mode — the buried cross-subunit contacts that hold each assembly together, with its assembly cooperativity (set by how many subunits and disulfide links share the load) and the mechanism it fails by when pushed past its limit (the subunits pulling apart vs. hydration collapse). A structure-derived stability-mechanism read, not a wet measurement.



## UV-PDA detector — absorbance spectrum & $\epsilon_{280}/A_{280}$ (exact from sequence; converts a UV peak to concentration)

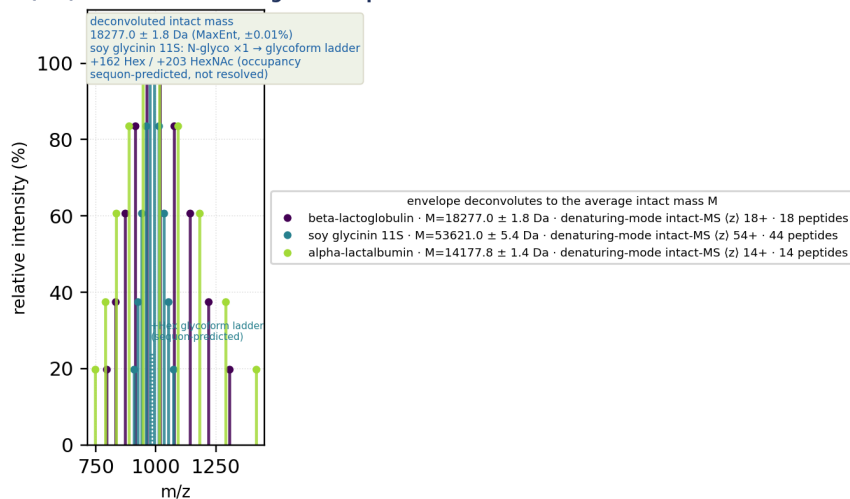


$\epsilon_{280} = 5500 \cdot \text{Trp} + 1490 \cdot \text{Tyr} + 125 \cdot \text{disulfide (Pace)}$  — EXACT from sequence (matches ProtParam: lysozyme 37,970 /  $A_{280}$  2.65).  $A_{280}(0.1\%) = \epsilon_{280}/\text{MW}$  converts the UV peak to concentration; peptide band ~205 nm.

Concentration by Beer-Lambert: a UV detector measures absorbance  $A$  and inverts  $A = \epsilon \cdot c \cdot l$  ( $\epsilon$  the molar extinction,  $l$  the 1 cm path length) to concentration —  $c = A / (\epsilon \cdot l)$ . Because  $\epsilon_{280}$  is exact from the Trp/Tyr/disulfide content (Pace), a measured  $A_{280}$  reads straight to mg/mL for your proteins: beta-lactoglobulin  $\epsilon_{280}=17210 \text{ M}^{-1}\text{cm}^{-1}$ ,  $A_{280}(1 \text{ mg/mL})=0.94$  · soy glycinin 11S  $\epsilon_{280}=38515 \text{ M}^{-1}\text{cm}^{-1}$ ,  $A_{280}(1 \text{ mg/mL})=0.72$  · alpha-lactalbumin  $\epsilon_{280}=28460 \text{ M}^{-1}\text{cm}^{-1}$ ,  $A_{280}(1 \text{ mg/mL})=2.01$ . The law is linear over  $A \approx 0.1$ – $1.5$ ; above that (a concentrated stock) dilute to bring the reading back into range.

## LC-MS (ESI) — intact-mass charge envelope (the exact intact mass, disulfide-corrected) & tryptic peptide-mass fingerprint

### LC-MS (ESI) — intact-mass charge envelope

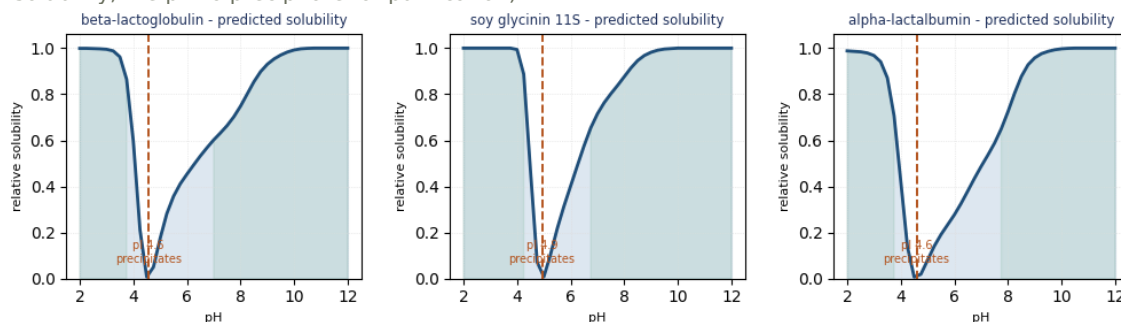


ESI  $[M+nH]^+$ :  $m/z = (M+n \cdot 1.00728)/n$ .  $M$  = AVERAGE intact mass (sequence); the monoisotopic mass + its 0-mDa check are on the stick-spectrum panel. Envelope spacing deconvolutes back to  $M$ ; tryptic map = the peptide-mass fingerprint. Glycoform ladder shows only with an N-glyco sequon (+Hex/+HexNAc, sequon-predicted).

## LC-MS/MS fragmentation — b/y ion ladder (CID peptide map; how the molecule fragments)

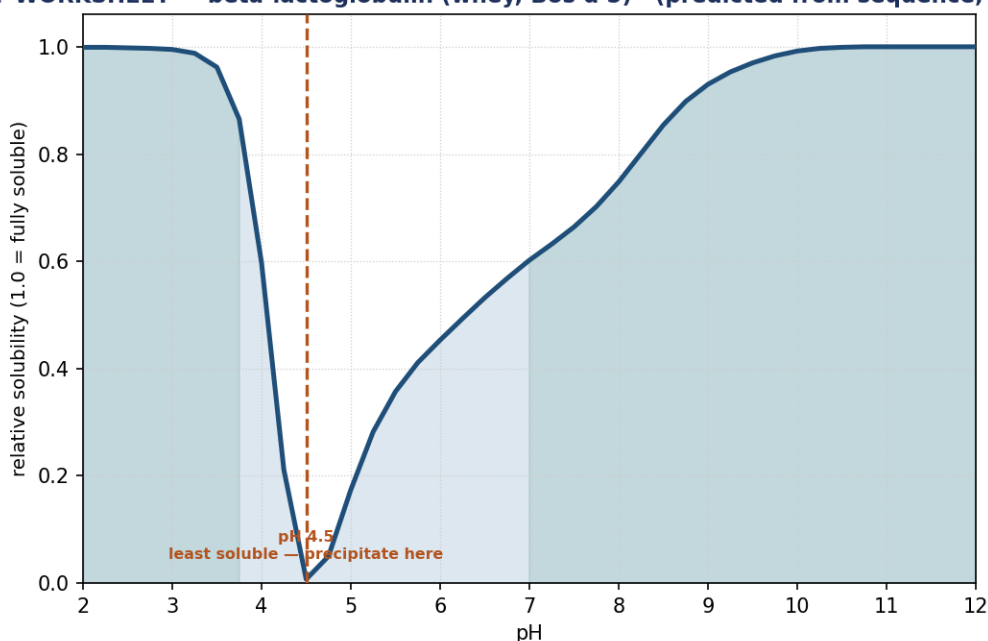
A representative tryptic peptide's collision-induced b/y fragment ladder — the MS/MS a peptide-mapping lab runs to confirm the sequence. Fragment  $m/z$  from the monoisotopic residue masses: beta-lactoglobulin — LIVTQTMK: 7 b-ions / 7 y-ions (e.g. b1 114.0913, b2 227.1754, y1 147.1128, y2 278.1533) · soy glycinin 11S — EQPQNECQIQK: 11 b-ions / 11 y-ions (e.g. b1 130.0499, b2 258.1084, y1 147.1128, y2 275.1714) · alpha-lactalbumin — FLDDDLTDDIMCVK: 13 b-ions / 13 y-ions (e.g. b1 148.0757, b2 261.1598, y1 147.1128, y2 246.1812).

Predicted solubility vs pH & isoelectric point — the functionality bottleneck for plant protein, predicted from sequence (the pH to avoid for solubility, the pH to precipitate for purification)



Solubility worksheet — read the predicted solubility at any pH straight off the sheet (graphical calculation aid; no software needed).

### SOLUBILITY WORKSHEET — beta-lactoglobulin (whey, Bos d 5) (predicted from sequence, no wet test)

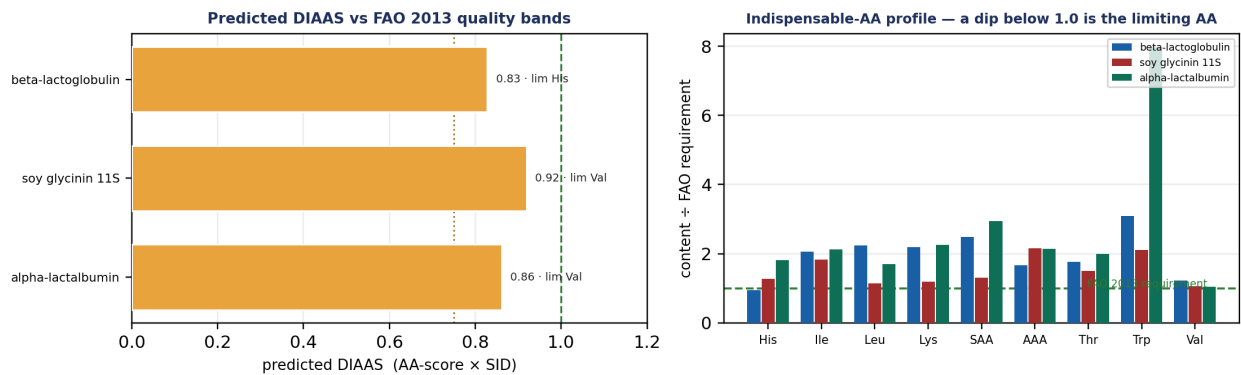


#### READ OFF THE CURVE:

1. At what pH is this protein LEAST soluble? (read the valley → the pI line)
2. Name a pH band where solubility  $\geq 0.6$  (the soluble working window, shaded green).
3. At your process pH (supply pH in the panel to pin it) — read the solubility straight up from that pH.
4. To purify by isoelectric precipitation, which pH drops it out of solution?

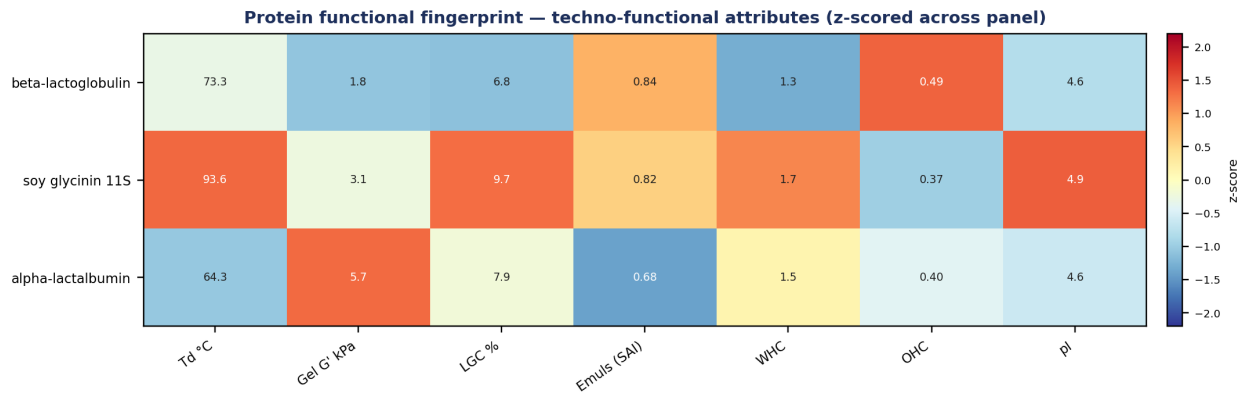
Hans-Made Research · client sheet

Predicted DIAAS (protein nutrition) — the amino-acid-analyzer read: each protein's predicted DIAAS (composition AA-score  $\times$  structure-derived digestibility SID) against the FAO 2013 quality bands, labelled with its limiting amino acid, plus the indispensable-amino-acid profile showing which AA dips below the requirement and caps the score. A composition + structural SCREEN — official DIAAS needs an in-vivo ileal-digestibility assay.



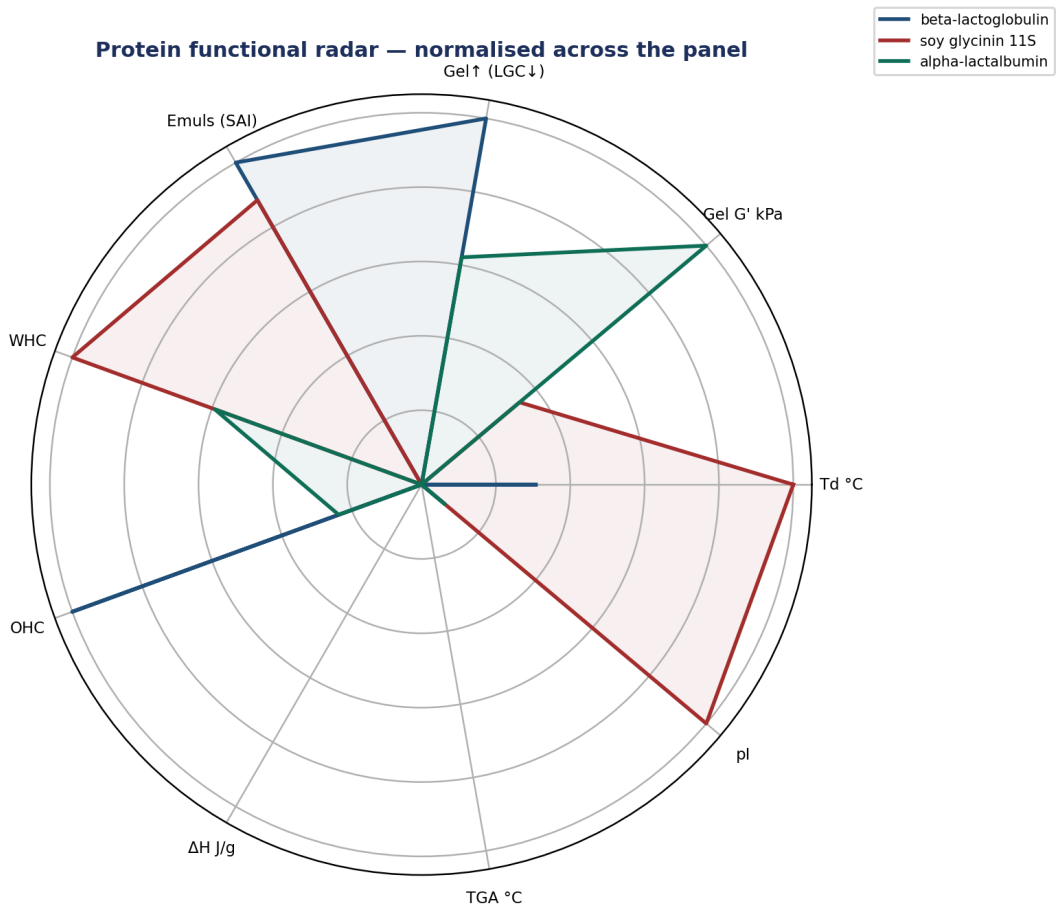
## Functional fingerprint heatmap — every protein, every axis

Each protein across every techno-functional axis ( $T_d$ , gel modulus, LGC, emulsification, water/oil holding,  $\Delta H$ , TGA, pl), z-scored so the cross-panel pattern reads at a glance.



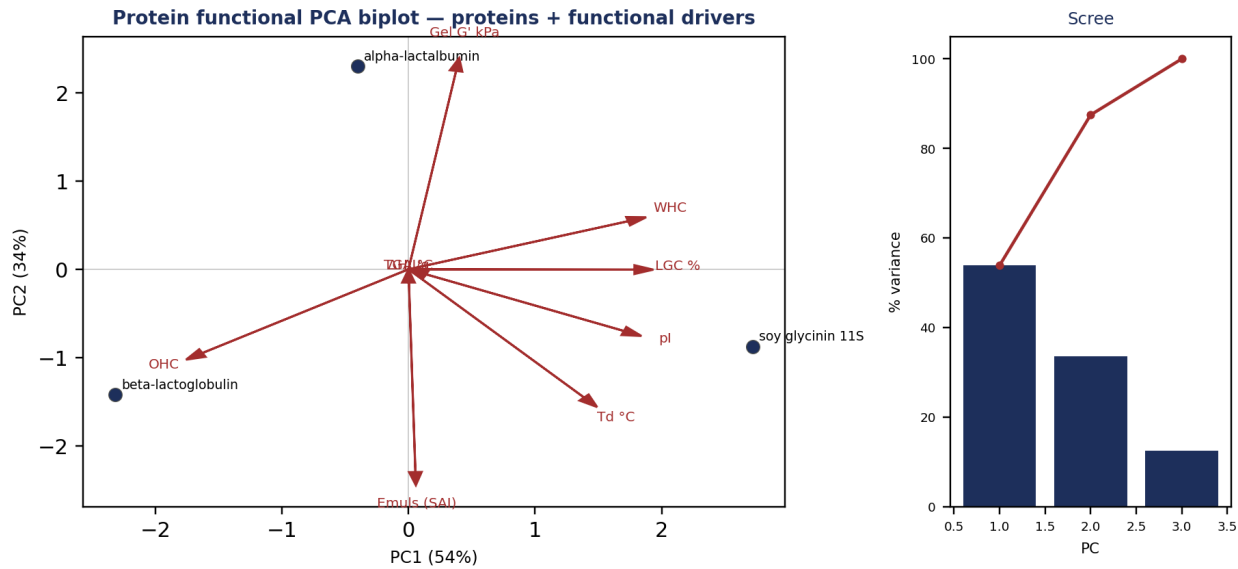
## Functional radar — multi-attribute fingerprints

Each protein's functional fingerprint overlaid (outward = stronger; gelling drawn as inverted LGC).



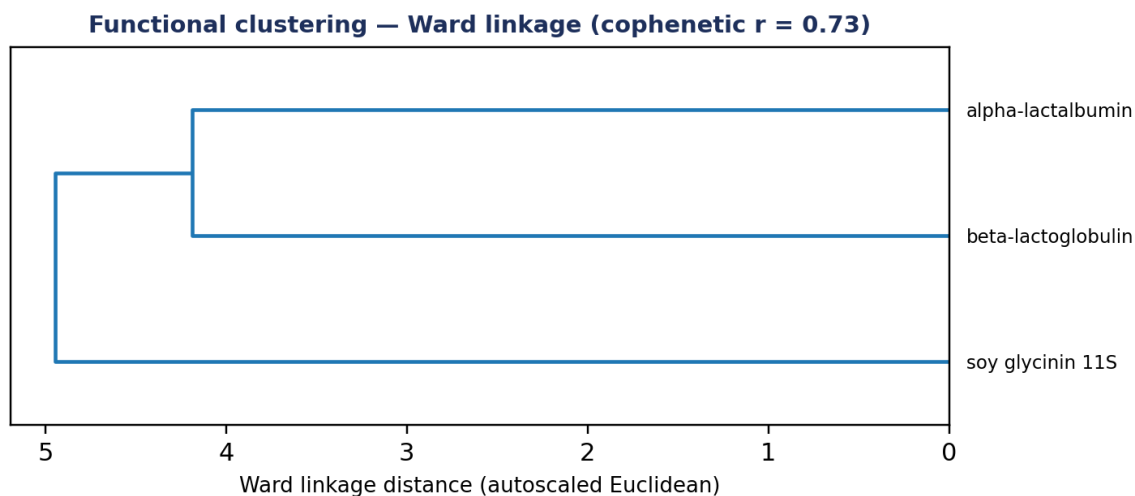
Functional PCA biplot — distinctness map

Which proteins cluster by functionality and which axes drive the spread — the novelty / distinctness map for non-obviousness.



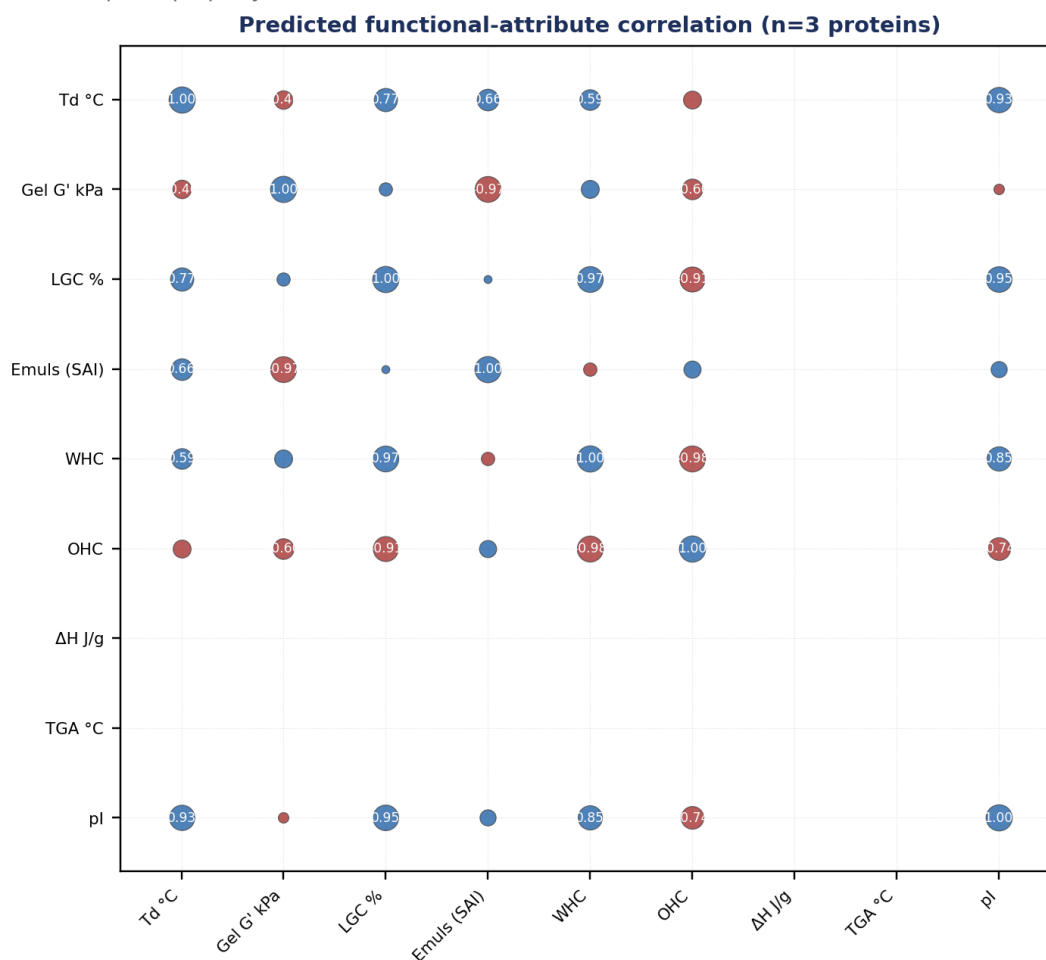
Functional clustering — Ward dendrogram

Cophenetic-correlation clustering: which novel proteins are functionally interchangeable with the soy/pea benchmarks.



## Functional-attribute correlation map (Pearson)

How the predicted properties co-vary across the panel (circle area  $\propto |r|$ , blue positive / red negative) — the trade-off structure a formulator reads to swap one property for another.



**Functional-attribute trade-offs** across the 3 proteins in this panel: WHC  $\leftrightarrow$  OHC (trade off); Gel G' kPa  $\leftrightarrow$  Emuls (SAI) (trade off); LGC %  $\leftrightarrow$  WHC (move together). With only 3 proteins these are read as DIRECTIONAL trade-offs, not a statistical correlation — a Pearson  $r$  on 3–4 points is near  $\pm 1$  by construction, so no coefficient is quoted here; the ranking, not the  $r$ , is the signal.

## Method & scope

Every value is computed live from the submitted structure by one fixed, deterministic physical model (zero training data — the same structure always returns the same result), with the structural mechanism named on each call. Where you supply measured labels the report scores itself against them above (with a label-shuffle control). This is a pre-bench shortlist / decision-support layer — it shows where to look first; it does not replace your panel or assays.

## Appendix — input / output JSON (machine-readable)

Input (as evaluated):  
{ "options": { "tier": 3 }, "proteins": [ { "id": "P02754", "name": "beta-lactoglobulin (whey, Bos d 5)", "conditions": {}, "sequence\_resolved\_a  
}  
Output (the read returned):  
[ { "name": "beta-lactoglobulin (whey, Bos d 5)", "denaturation\_C": 73.3, "pI": 4.56, "fold\_class": "LIPOCALIN" }, { "name": "soy glycinin 11S (  
]

## Provenance & integrity

This report is cryptographically bound to the submitted inputs (as standardized structures) and the engine version. Re-running the same inputs on the same engine reproduces the identical hash; any change in transit — an input or a computed result — changes it. Reproducible and tamper-evident from the hash alone.

| Field                     | Value                                                            |
|---------------------------|------------------------------------------------------------------|
| Prepared for              | bench_agent                                                      |
| Source file               | bench_agent_run_3338a79ded.json                                  |
| Records bound             | 3                                                                |
| Engine version (SHA-256)  | 44ebc09e38ed0800d6149578dc341e6e0b3fc7c7796d9c11bd993cf1d72f04eb |
| Merkle root — all results | e3836ab28d74b28fd4fe047710cc808c862b1c7043092c8c190686597d144f73 |
| Manifest hash (the seal)  | 83a90cfcabc2ee92046338ac6a7d340de89d085358a652c902608129b101a892 |
| Spec                      | RaRaMa-merkle-v1 (sha256, dup-last)                              |

Verify: recompute the Merkle root from the per-compound leaves in the sidecar (RaRaMa\_report\_<name>\_merkle.json) and confirm it matches the manifest hash above — no access to the engine is needed to check integrity.