Ten cross-disciplinary experiments for egg replacement & off-flavor reduction

A shortlist of benchtop-scale experiments generated by mining the seams *between* sub-literatures — enzyme chemistry, brewing, pharma formulation, cheese ripening — that individually study alt-protein problems but rarely cite one another.

Corpus: 80 peer-reviewed papers (40 off-flavor · 40 egg)
Method: keyword-scored literature mine → LLM synthesis, cited to paper IDs
Selection: top 5 experiments + 1 data program per field **Status:** unbuilt — for specialist review

The premise is that the highest-leverage unexplored experiments in alternative-protein R&D are not new mechanisms — they are **known mechanisms that have never been introduced to each other.** A phospholipase used for decades on egg yolk has never been pointed at mung-bean protein. Brewing solved aroma-loading of a protein matrix a century ago; plant-protein flavor science doesn't cite it. So rather than ask a model to invent chemistry, we built a small, tightly-filtered corpus and forced every proposed experiment to name the specific papers — and specific fields — it stitches together.

COLLECT **~750 candidates** OpenAlex + Europe PMC, two themes → FILTER **80 papers** keyword-density score, top-40 / theme → SELECT **20 seed papers** chosen for *mechanism diversity*, not score → SYNTHESIZE **25 hypotheses** LLM, each cited to paper IDs

THE GENERATIVE MOVE

The seed papers were hand-picked for *coverage of distinct intervention classes* (breeding, processing, enzymatic, fermentation, non-thermal, hybrid, precision fermentation) — deliberately *not* the top-scored papers, which all cluster on one well-studied intervention. Cross-paper synthesis has more surface area when the inputs don't overlap. The model was then asked for combinations tested individually but never together, and for mechanisms importable from adjacent fields.

HOW TO READ THE RANKING

Each experiment carries a rough portfolio score — the product of five 1–5 judgments: *insight, probability of success, commercial value, breadth,* and *neglect.* Within each field the five benchtop experiments come first, ordered best-first — then one *data-infrastructure* program (shown on a dashed card). Both scored highly; they aren't lab experiments but shared datasets every experiment above them would draw on.

> **Where the ideas come from — the recurring pattern is a citation gap between two mature fields:**
> phospholipase / egg-yolk chemistry × plant-protein design · brewing hop-stand aroma-loading × pea-protein flavor · cheese-rind ripening cultures × textured plant protein · pharma cyclodextrin/MIP capture × protein isolation · precision-fermented ovalbumin × plant-protein backbones · wine/tea tannin–aldehyde chemistry × protein slurries

§ The ten experiments at a glance

#	EXPERIMENT	CROSS-FIELD GRAFT	EST. COST	DIFFICULTY	P(SUCCESS)
	EGG REPLACEMENT — HARDEST UNSOLVED FUNCTIONS FIRST				
E1	H14 · **PLA2-hydrolyzed mung-bean liquid egg**	enzyme chem → plant protein	$25–50k	Moderate	High
E2	H15 · **PLA2 + transglutaminase cake-binder cascade**	2-enzyme cascade	$20–40k	Moderate	High
E3	H16 · **2–5% ovalbumin spike on a plant backbone**	precision ferm. × plant	$40–80k	Moderate	High*
E4	H20 · **Hot-set + cold-set “boiled egg white” mimic**	colloid gel cascade	$25–50k	Moderate	Moderate
E5	H17 · **Aggregate-engineered mung foam for meringue**	colloid sci → foam	$20–40k	Moderate	Mod–High
E6	H21 · **Cross-protein egg-function benchmarking atlas** ◇ DATA PROGRAM	shared measurement layer	$100–200k	Moderate	Very high
	OFF-FLAVOR REDUCTION — HIGHEST-LEVERAGE LEVERS FIRST				
O1	H1 · **Glycosidase knockdown for residual 1-octen-3-ol**	glycoside pathway × breeding	$50–100k	Moderate	Moderate
O2	H2 · **Sequential RF + cold-plasma deactivation**	two non-thermal levers	$25–75k	Moderate	High
O3	H7 · **Cyclodextrin volatile-scavenging at isolation**	pharma excipient → protein	$10–20k	Simple	High
O4	H10 · **Glutaminase → alcalase → TGase enzyme cascade**	3-enzyme series	$30–60k	Moderate	Moderate
O5	H8 · **Three-stage flavor-*building* fermentation**	cheese/cured-meat cultures	$30–60k	Moderate	High
O6	H4 · **Variety × harvest × storage GC-MS atlas** ◇ DATA PROGRAM	lot-selection dataset	$150–300k	Moderate	High

* mechanistically high; economics hinge on recombinant-ovalbumin price. “Novelty” throughout means *not found in this 80-paper corpus* — several ideas may exist in patents (excluded here) or industry trade secret.

Egg replacement

5 experiments + 1 data program

Emulsification, structural gelling, and foam are three different egg jobs no single plant protein does well. These target the hardest ones — liquid egg, boiled-egg-white, meringue — with mechanisms borrowed from enzyme chemistry and colloid science.

1 · EGG · H14 · FLAGSHIP EXPERIMENT
PLA2-hydrolyzed mung-bean liquid egg with tailored emulsion architecture

Treating a mung-bean protein + phospholipid mixture with phospholipase A2 reproduces the lyso-phospholipid + protein architecture that gives PLA2-treated egg yolk its emulsifying power — yielding a plant liquid-egg whose emulsion stability matches or beats hydrolyzed yolk.

COST **$25–50k** · TIMELINE **~3 mo** · DIFFICULTY **Moderate** · P(SUCCESS) **High** · NOVELTY **High in corpus** · INPUTS **all food-grade, off-shelf**

CROSS-DISCIPLINARY ORIGIN
egg_017 (2024): PLA2 on egg yolk lifts protein solubility 27%→91% by disintegrating lipoprotein micelles. egg_026: polysaccharide-conjugated mung protein emulsifies well. egg_029: pH-shift + TGase is today's best plant liquid-egg. No one has pointed the yolk enzyme at mung protein.

MECHANISM
PLA2 cleaves the sn-2 fatty acid from phosphatidylcholine, yielding lyso-PC (a potent emulsifier) + free fatty acid. In yolk these reorganize the lipoprotein micelle; in mung protein + added soy lecithin the same chemistry should self-assemble emulsion-active structures.

THE EXPERIMENT
Mung isolate + soy lecithin at 5:1 / 10:1 / 20:1 protein:lecithin, dosed with commercial food-grade PLA2 (e.g. Lecitase Ultra). Incubate; benchmark against native mung, PLA2-treated yolk, and commercial mayo.

READS OUT ON
emulsion stability index · droplet size (SLS) · freeze–thaw cycles · heat-set gel G′

KILL IF PC→lyso-PC conversion <30% at 2 h, droplet size >2× the yolk control, or ESI no better than the pH-shift/TGase baseline.

2 · EGG · H15
PLA2 + transglutaminase cascade as a single cake-binder ingredient

A two-enzyme cascade — PLA2 to generate emulsifier in situ from added lecithin, then TGase to cross-link the protein network — reproduces *both* the emulsifying and structural-gelling roles of egg in one bakery ingredient.

COST **$20–40k** · DIFFICULTY **Moderate** · P(SUCCESS) **High** · NOVELTY **High in corpus** · BUILDS ON **E1 substrate stage**

CROSS-DISCIPLINARY ORIGIN
egg_017 (PLA2 emulsification) unified with egg_005, egg_006, egg_029 (TGase functionality gains in mung / pea protein). Single-enzyme studies handle emulsification and gelling separately; the cascade does both in sequence.

MECHANISM
PLA2 first makes lyso-PC + FFA emulsifier and better-solubilizes the protein; TGase then cross-links Lys–Gln residues on the now-accessible network. Order matters — PLA2 doesn't consume TGase substrates, so emulsification precedes structure-building.

THE EXPERIMENT
Apply PLA2 then TGase to mung isolate + lecithin slurry; bake into a fixed pound-cake formula against either enzyme alone, whole egg, and a commercial egg replacer.

READS OUT ON
cake volume · crumb image analysis · texture profile (TPA) · 7-day moisture retention

WIN BAR Matches whole-egg cake volume & crumb while beating the commercial replacer on staling.

3 · EGG · H16
A 2–5% precision-fermented ovalbumin spike on a plant-protein backbone

Adding only 2–5% recombinant ovalbumin on top of a 95% plant-protein base buys egg-mimic cake structure at a fraction of a pure-recombinant cost — reframing ovalbumin as a functional cross-linker additive, not a wholesale egg replacer.

COST **$40–80k** · DIFFICULTY **Moderate** · P(SUCCESS) **High, price-sensitive** · NOVELTY **Moderate** · NEEDS **ovalbumin sample access**

CROSS-DISCIPLINARY ORIGIN
egg_013 (2025) names downstream-processing cost as the binding constraint on recombinant ovalbumin. egg_035 shows ovalbumin's SH groups are the *structure-forming* element in egg-white cake. Spend the expensive protein only where it's uniquely needed — cross-linking — and let plant protein do the rest.

MECHANISM
Ovalbumin's free thiols form disulfide cross-links during bake, supplying the structural backbone egg white contributes. A small spike delivers cross-linker density without paying for ovalbumin to also emulsify, where plant proteins are already adequate.

THE EXPERIMENT
Fixed chickpea or mung backbone + ovalbumin at 0/1/2/5/10% of total protein, baked against whole-egg control — run alongside a cost model parameterized on ovalbumin $/kg.

READS OUT ON
cake volume & crumb · sensory vs whole egg · cost / functional unit

PRIZE Egg-equivalent structure at ≤5% spike — a 10–50× cut in required recombinant loading.

4 · EGG · H20
Hot-set + cold-set gel cascade for a plant “boiled egg white”

Pairing a hot-set protein (potato or mung) with a cold-set gelling agent (κ-carrageenan / low-acyl gellan) plus TGase yields a plant boiled-egg-white that holds shape after both cooking and chilling — an unsolved, wide-open format.

COST **$25–50k** · DIFFICULTY **Moderate** · P(SUCCESS) **Moderate** · NOVELTY **Mod–High** · WHITE SPACE **standalone egg mimic**

CROSS-DISCIPLINARY ORIGIN
egg_020 (polysaccharide + plant-protein gels), egg_021 (TGase-crosslinked plant-dairy gels) and egg_031 (casein–soy TGase gels) each build networks — but none target the boiled-egg-white brief, which uniquely needs hot-set *and* cold-stable *and* the right bite at once.

MECHANISM
Plant protein supplies the thermal gel formed during cook; carrageenan + Ca^{2+} supplies a cold-set network that survives chilling; TGase adds covalent cross-links robust across both temperature extremes.

THE EXPERIMENT
Factorial: protein (potato / mung / soy isolate) × polysaccharide (κ-carrageenan / gellan / mix) × TGase on/off. Cook 90 °C × 20 min, chill to 4 °C, profile texture.

READS OUT ON
TPA hardness / springiness · cohesiveness · water release · bite vs hard-boiled egg

TARGET A slice that survives salad handling with hard-boiled-egg-white bite.

5 · EGG · H17
Aggregate-engineered mung foam for a full meringue replacer

The exceptional foaming of phase-separated mung-protein colloids can be amplified by deliberate aggregate-forming pre-treatments — mild Maillard or controlled fermentation — to fully replace egg white in meringue, the canonical hardest foam test.

COST **$20–40k** · DIFFICULTY **Moderate** · P(SUCCESS) **Mod–High** · NOVELTY **High in corpus**

CROSS-DISCIPLINARY ORIGIN
egg_054 (2024): liquid–liquid phase-separated mung colloids reach 324% foamability with a 400-min half-life, far beyond mildly-purified mung. The colloid-science mechanism (aggregates thicken the film, block lamellar drainage) predicts that *any* aggregate-promoting treatment should push it further — a link the food-science meringue literature hasn't made.

MECHANISM
Larger, denser aggregates in the thin lamella slow drainage. Mild Maillard cross-links raise protein–protein interaction; fermentation-driven aggregation does the same via pH-shift and bacterial-protease modification.

THE EXPERIMENT
Four mung variants — baseline, colloid mix, Maillard-conditioned (mild glucose, 65 °C × 6 h), LAB-fermented + heat-aggregated — compared on foam and baked meringue.

READS OUT ON
foam overrun · half-life / drainage curves · interfacial viscosity · meringue volume & crispness

TARGET A baked meringue matching egg-white structure and crisp set.

◇ DATA-INFRASTRUCTURE PROGRAM · SHARED MEASUREMENT LAYER, NOT A BENCHTOP EXPERIMENT

6 · EGG · H21 · DATA PROGRAM
Cross-protein × cross-function egg-replacement benchmarking atlas

One normalized dataset measuring every candidate protein against every egg function at common conditions — so formulators *pick* the right protein for each role instead of guessing across papers that can't be compared.

COST **$100–200k** · DIFFICULTY **Moderate** · P(SUCCESS) **Very high (engineering)** · NOVELTY **The project, not the science** · VALUE **compounds over time**

CROSS-DISCIPLINARY ORIGIN
egg_001, egg_018 and egg_027 each benchmark proteins — but at different concentrations, methods and applications, so no two results line up. The data exists; it has never been put on one axis.

WHAT MAKES IT DIFFERENT
Not a hypothesis to falsify — a measurement substrate. It tests no mechanism; it removes the guesswork *under* every card above, telling you which protein to reach for before you run any of them.

THE BUILD
One standardized protocol across ~20 commercial plant + dairy ingredients × 8 egg functions (emulsification, foaming, gelling, water- and oil-holding, binding, browning, opacity), released as a public database.

DELIVERS
protein × function matrix · ESI · overrun · G′ · MGC · WHC · OHC · open database

PAYOFF Ingredient selection becomes an evidence lookup instead of a guess — and the reference grows more valuable as rows are added.

Off-flavor reduction

5 experiments + 1 data program

The beany / green / earthy signature of plant proteins is a mix of LOX-derived volatiles, non-LOX enzymatic sources, and formed compounds. These attack it at different points — precursor, process, isolation, and fermentation — with levers imported from breeding, brewing, pharma, and cheese-making.

1 · OFF-FLAVOR · H1 · HIGHEST UPSIDE
Glycosidase knockdown to eliminate residual 1-octen-3-ol

Knocking down the soybean β-glucosidase that releases 1-octen-3-ol from its glycoside precursor eliminates the dominant mushroom / earthy off-note that *survives* LOX-null breeding — the residual gap every LOX-knockout program hits.

COST **$50–100k (inhibitor route)** · DIFFICULTY **Moderate** · P(SUCCESS) **Moderate** · NOVELTY **High in corpus** · UPSIDE **field-wide step change**

CROSS-DISCIPLINARY ORIGIN
Three papers, never connected: off_047 — 1-octen-3-ol is released from a glucoside *independent of LOX*; off_007 — it still ranks top by odor-activity value *after* LOX knockout; off_030 — residual off-flavor in triple-LOX-null soy traces to a heat-labile "LOX-like enzyme." Together they name a druggable glycosidase as the next target.

MECHANISM
A β-glucosidase active during soak/maceration liberates volatile 1-octen-3-ol from a non-volatile glucoside. Anything that inactivates it before hydration — heat, an inhibitor, or a low-activity genotype — should suppress the note.

THE EXPERIMENT
(a) Screen varietal panels for naturally low β-glucosidase activity; (b) in parallel, add food-compatible β-glucosidase inhibitors at the soak step vs a heat-inactivated control; (c) quantify the full odor-active panel by HS-SPME-GC-MS.

READS OUT ON
1-octen-3-ol OAV · mushroom/earthy panel score · protein solubility / foaming

TARGET 1-octen-3-ol OAV <0.1× the LOX-null baseline without damaging protein function.

2 · OFF-FLAVOR · H2
Sequential radio-frequency + cold-plasma deactivation

RF seed pretreatment (kills the precursor-making enzyme) followed by DBD cold plasma on the wet slurry (oxidizes the volatiles that slip through) gives additive off-flavor reduction while staying below thermal protein-damage thresholds.

COST **$25–75k** · DIFFICULTY **Moderate** · P(SUCCESS) **High** · NOVELTY **High in corpus** · RIGS **pilot-scale available**

CROSS-DISCIPLINARY ORIGIN
off_018 (2026): RF at ~75 °C inactivates 80% of LOX with no yield loss. off_021 (2026): DBD cold plasma at 60% O_2 degrades 1-octen-3-ol, 2-pentylfuran and others — compounds formed *after* extraction. Both are 2026; the obvious combination hasn't had time to publish.

MECHANISM
The two act on different timescales (seed vs slurry) and different compound classes (precursors vs formed volatiles). Independent failure modes → multiplicative, not merely additive, reduction.

THE EXPERIMENT
2×2 factorial (RF on/off × plasma on/off) on one yellow-pea source: dehull → RF → wet fractionation → cold plasma at 60% O_2, vs conventional control.

READS OUT ON
hexanal / nonanal · 1-octen-3-ol / 2-pentylfuran · sensory triangle test · solubility & foaming

TARGET Combined arm beats either single treatment on volatiles with functionality intact.

3 · OFF-FLAVOR · H7 · CHEAPEST, CLEANEST
Cyclodextrin in-situ scavenging of pyrazines & aldehydes at isolation

Adding β- or γ-cyclodextrin during pea-protein extraction traps methoxypyrazines and key aldehydes as inclusion complexes that partition off at ultrafiltration — stripping volatiles while leaving protein structure untouched.

COST **$10–20k** · DIFFICULTY **Simple** · P(SUCCESS) **High mechanistically** · NOVELTY **High in corpus** · REGULATORY **food-grade (E459/E468)**

CROSS-DISCIPLINARY ORIGIN
off_014 (2023) pins methoxypyrazines as major pea-aroma contributors — trace, polar, and stubbornly bound in protein hydrophobic pockets. Cyclodextrins are the standard pharma/cosmetic tool for exactly this capture problem, and are entirely absent from the plant-protein literature.

MECHANISM
Cyclodextrins form 1:1 inclusion complexes with small hydrophobic guests. Pyrazines, hexanal, nonanal and 1-octen-3-ol all fit β/γ-CD cavities; CD-bound volatiles leave with the permeate at ultrafiltration, unchanged protein stays in the retentate.

THE EXPERIMENT
β-CD and γ-CD at 1/2/5% w/v added post-extraction, pre-UF, vs untreated control. Profile volatiles in both retentate and permeate; confirm no structural change by SDS-PAGE.

READS OUT ON
retentate volatiles ↓ · permeate volatiles ↑ · CD residue (HPLC) · solubility / SDS-PAGE

TARGET Meaningful volatile drop at legal CD dose with an unchanged protein fingerprint.

4 · OFF-FLAVOR · H10
Glutaminase → alcalase → transglutaminase enzyme cascade

Running three enzymes in series on dry-fractionated pea protein hits a trio usually treated as incompatible: lose beany volatiles, gain solubility, *and* recover gelling — because each enzyme fixes what the previous one leaves.

COST **$30–60k** · DIFFICULTY **Moderate** · P(SUCCESS) **Moderate** · NOVELTY **High in corpus** · UNIFIES **3 separate process choices**

CROSS-DISCIPLINARY ORIGIN
Each enzyme has a champion paper and a trade-off: glutaminase (off_008) removes hexanal/2-pentylfuran but not function; alcalase (off_024) cuts off-flavor but peptides lose gelling; TGase (egg_005) rebuilds gelling. No one runs them as one ordered cascade.

MECHANISM
Glutaminase deamidates surface Gln/Asn (raises solubility, frees off-flavor binding sites); alcalase cleaves remaining hydrophobic peptides (more flavor release); TGase re-crosslinks the network. Order is load-bearing — TGase substrate availability depends on the prior hydrolysis.

THE EXPERIMENT
Single-variable scan on one dry-fractionated pea: each enzyme alone, all pairwise, and the full cascade — measuring flavor, sensory and function after every step.

READS OUT ON
hexanal / 2-pentylfuran ↓ · solubility & foaming ↑ · gel hardness / G′ held

PRIZE One ingredient that is simultaneously de-flavored, soluble, and gelling.

5 · OFF-FLAVOR · H8
Three-stage fermentation that *builds* flavor, not just removes it

Adding a third, cured-food culture stage to an existing two-stage deflavoring ferment simultaneously strips beany volatiles *and* adds meaty / umami aroma — flipping fermentation from cost-of-defect-removal into a flavor-creation asset.

COST **$30–60k** · DIFFICULTY **Moderate** · P(SUCCESS) **High chem · Mod sensory** · NOVELTY **Mod–High**

CROSS-DISCIPLINARY ORIGIN
off_003 (2025) runs two LAB/yogurt stages that reduce volatiles but add nothing; off_034 shows fermentation + protease yields amino-acid-derived aromas; off_037 names rapid-acidifying strains. The natural third stage — surface-ripened-cheese and dry-cured-ham cultures — is unrun on plant protein.

MECHANISM
Stages 1–2 (L. plantarum → yogurt culture) drop pH and degrade volatiles. Stage 3 (Brevibacterium linens or Staphylococcus carnosus + neutral protease) generates Strecker aldehydes and short-chain acids — the same chemistry behind cheese-rind and cured-ham aroma.

THE EXPERIMENT
Common 9% pea substrate; identical stages 1–2; four stage-3 candidates (B. linens, S. carnosus, D. hansenii, control) with added protease. Track volatile evolution by SIFT-MS at each stage.

READS OUT ON
stagewise volatile fingerprint · beany ↓ / umami · meaty ↑ · free amino acids (HPLC)

TARGET Net-positive aroma profile with beany notes suppressed and savory notes built.

◇ DATA-INFRASTRUCTURE PROGRAM · LOT-SELECTION DATASET, NOT A BENCHTOP EXPERIMENT

6 · OFF-FLAVOR · H4 · DATA PROGRAM
Variety × harvest-year × storage → volatile-fingerprint atlas

A public dataset mapping how variety, harvest year, location and storage drive pea-protein volatile load — so formulators *select* low-off-flavor lots rather than spend yield and function process-correcting them.

COST **$150–300k yr-1** · DIFFICULTY **Moderate** · P(SUCCESS) **High (engineering)** · NOVELTY **High in corpus** · VALUE **compounds over time**

CROSS-DISCIPLINARY ORIGIN
off_040 (harvest year & location dominate LOX activity), off_039 (storage reshapes volatiles over a year) and off_001 (varietal fatty-acid profile → off-note load) together imply lot selection can outperform process intervention — yet no selection infrastructure exists.

WHAT MAKES IT DIFFERENT
Not a process to test — a selection layer. Where the five experiments above *correct* off-flavor, this lets you *avoid* it upstream at zero yield cost, and quantifies how much of the problem is lot choice versus processing.

THE BUILD
3–5 processor partners each contribute ~20 lots/year with metadata; standardized PPI extraction + HS-SPME-GC-MS fingerprint; open-data release; a regression predicting volatile load from metadata, validated on held-out lots.

DELIVERS
volatile fingerprint per lot · metadata → volatile model · open dataset

PAYOFF Predict hexanal / nonanal / 1-octen-3-ol on unseen lots (R^2>0.7 is useful) — buy the good lots, skip the correction.

READING & PROVENANCE

Paper IDs (e.g. **egg_017**, **off_047**) are stable references into the project corpus; each maps to a title, abstract and DOI in the underlying dataset, so any claim above traces back to primary literature. **“Novelty (in corpus)”** means the exact experiment was not found among these 80 papers — it does not rule out prior art in patents (deliberately excluded) or unpublished industry work; treat every card as a hypothesis to be checked, not a claim of first-invention. **Costs, timelines and success odds** are order-of-magnitude buckets for triage, not quotes. **The sixth item in each field** sits on a dashed card because it is a different kind of thing — a public dataset (a variety-×-harvest-×-storage GC-MS *atlas*; a cross-protein functional-benchmarking *database*) rather than a benchtop test. Both score highly and would de-risk every experiment above them, but they should be judged as engineering programs, not hypotheses to falsify.

Alt-Protein Paper Scout · hypothesis brief v1 · generated for specialist review · not a funding submission.