A patient-specific antigen-selection engine for personalized pMHC-II tolerance therapy in eosinophilic esophagitis

Ruth-Anne Pai, PhD[1,2]

[1] Citizen Scientist, EoE patient-researcher; Built with Claude: Life Sciences Hackathon (June–July 2026)

[2] Pai Advisory, LLC

* Correspondence: Ruth-Anne Pai (Pai Advisory, LLC) — advising@ruthannepai.com

Preprint — not peer reviewed. Subject area: Immunology (secondary: Bioinformatics / Molecular Biology).

Abstract

Eosinophilic esophagitis (EoE) is a chronic, food-antigen–driven type-2 inflammatory disease. Every approved and late-stage therapy — dupilumab, cendakimab, tezepelumab, swallowed corticosteroids, proton-pump inhibitors — is a broad, chronically dosed suppressor of the inflammatory cascade; none is antigen-specific, and none induces durable tolerance. Yet EoE is, mechanistically, an antigen-specific disease: a food peptide presented on an MHC-II molecule to a pathogenic $CD4^+$ T-cell clone. The central obstacle to exploiting this is that EoE is driven by **plural, patient-variable food allergens** (milk, wheat, egg, soy, often several at once), so the therapeutically relevant epitope cannot be fixed in advance — it must be computed per patient. Here we present an antigen-selection engine that converts a patient's HLA class-II genotype into a ranked, manufacturable epitope set, and a rational-design framework that folds the selected epitope onto a shared MHC-II backbone for delivery in a tolerogenic multivalent context. Scanning three dominant EoE allergens (β-casein, α/β-gliadin, β-conglycinin) across five common HLA-DRB1 alleles, we predict **832 strong binders (IC_{50} < 500 nM) from 4,640 peptide × allele evaluations**, and show that the presented epitope load is strongly HLA-dependent — ranging from **13 to 352 strong binders per allele across the same three allergens** — establishing that antigen priority is individual and must be computed, not assumed. As a worked example we advance a milk (β-casein) lead epitope on HLA-DRB1*07:01; AlphaFold-multimer folding of the designed pMHC-II panel clears an ipTM ≥ 0.87 acceptance gate with full in-groove peptide engagement, and a 20 nm iron-oxide nanoparticle displaying five pMHC copies (≈1.4% surface occupancy, ~24× avidity) provides the tolerogenic multivalent scaffold. We frame development around a biomarker-first cascade — structural gate → biochemical stability → tetramer engagement → an IL-10-biased Tr1 tolerance signature → in-vivo and human pharmacodynamics — with a companion diagnostic co-developed as a regulatory precondition. This citizen-science effort defines a computable, per-patient route to allergen-specific tolerance and a platform template extensible to celiac and other food/auto-antigen disease.

Introduction

Eosinophilic esophagitis affects roughly 1 in 2,000 individuals in Western populations, with rising incidence, and progresses in inadequately controlled patients to fibrostenotic remodeling and esophageal narrowing [1]. Its pathogenesis is a food-triggered type-2 immune response: $CD4^+$ T cells recognize allergen-derived peptides presented on MHC-II molecules and drive the IL-4/IL-5/IL-13 axis, eosinophil recruitment, and tissue remodeling [3]. Recent single-cell and TCR-sequencing studies establish that the disease is clonal and antigen-directed: clonally expanded, pathogenic effector T2 cells populate the inflamed esophagus [3], and a restricted, food-specific $CD4^+$ T-cell repertoire is detectable in patients [4]. The current therapeutic armamentarium — dietary elimination, swallowed topical corticosteroids, proton-pump inhibitors, and biologics against IL-4Rα, IL-13, IL-5/IL-5Rα, and TSLP — reduces inflammation effectively for many patients but shares one structural feature: it suppresses the response broadly and must be dosed indefinitely, with inflammation returning when treatment stops [2]. None addresses the antigen-specific root of the disease, and none induces lasting, drug-free tolerance.

Allergen-specific immunotherapy offers that possibility in principle, and the adjacent antigen-specific-tolerance field now provides hard lessons about how to pursue it. Five are transferable and shape the present design. First, **biomarker-first**: programs that qualified a pharmacodynamic (PD) tolerance biomarker before committing to a symptom endpoint advanced (teplizumab in type-1 diabetes [8]; TAK-101 in celiac [10]), whereas a program that bet a symptom endpoint before locking a PD marker failed (Nexvax2 [9]). Second, **deliver antigen in a tolerogenic context, never as bare allergen**: Nexvax2's free peptide reproduced the disease [9], and for food allergens the reactogenicity/anaphylaxis risk is paramount — a pMHC-II scaffold displayed on a tolerogenic multivalent nanoparticle [11] is favored over free peptide. Third, **stratify by HLA/sensitization from the outset**, as the GAD-alum experience in type-1 diabetes showed when responder selection by HLA rescued an otherwise-null antigen. Fourth, **target early, tolerance-tractable disease** with a progression-relevant endpoint, following teplizumab's redefinition of population and endpoint. Fifth, **co-build the screening/ companion-diagnostic pathway**, since first-in-class tolerance therapies are gated by screening infrastructure rather than demand.

Applying these lessons to EoE exposes a problem absent from celiac (one antigen, gluten) or type-1 diabetes (defined autoantigens): **EoE has no single causal antigen**. Its drivers are multiple, patient-variable food proteins, frequently co-occurring, and which peptides a given patient can present depends on that patient's HLA class-II genotype — up to four functional MHC-II molecules across two haplotypes. Antigen selection is therefore not a preliminary step but the central computational task, and it is genuinely individual: the same allergen panel yields sharply different presentable-epitope sets on different alleles.

We address this directly. We describe (i) a patient-specific **antigen-selection engine** that scores every candidate allergen peptide against a patient's HLA molecules and ranks the presentable epitope set; (ii) structural validation of the selected pMHC-II complexes by AlphaFold-multimer against explicit acceptance gates; (iii) a **shared-backbone, swappable-cassette** construct design (the PACT™ concept) that makes per-patient manufacturing tractable by changing only the peptide; and (iv) a tolerogenic multivalent nanoparticle scaffold for delivery. We anchor the framework with a fully worked milk (β-casein) lead on HLA-DRB1*07:01 and present the multi-HLA, multi-allergen prediction

landscape that motivates per-patient selection. Throughout, we are explicit about which claims rest on validated data and which are computational priors pending the wet-lab cascade we outline.

Results

A patient-specific antigen-selection engine

The engine takes a patient's HLA class-II typing and a panel of candidate food allergens and returns a ranked set of presentable epitopes. Every 15-mer of each candidate allergen is scored for MHC-II binding against each of the patient's molecules, strong binders are thresholded (IC_{50} < 500 nM), and dominant cores are ranked by summed strong-binder load and by molecule-level concentration. To characterize the engine across the alleles most relevant to EoE, we scanned three dominant food allergens — bovine β-casein (milk; UniProt P02666), α/β-gliadin (wheat; P18573), and β-conglycinin β-subunit (soy; P25974) — against five common HLA-DRB1 alleles (*01:01, *03:01, *04:01, *07:01, *15:01). From **924 unique 15-mers and 4,640 peptide × allele evaluations, the engine identified 832 strong binders** (Figure 1b).

The central result is that presentation is strongly HLA-dependent, and therefore individual (Figure 1a). Summed across the three allergens, the strong-binder load ranges from **352 epitopes on HLA-DRB1*01:01 down to 13 on HLA-DRB1*03:01** — a 27-fold difference in how much of the same food panel a patient can present, driven entirely by genotype. The hierarchy is also allergen-specific within an allele: DRB1*01:01 presents β-conglycinin most heavily (195 strong binders) but β-casein far less (74), whereas the relative ordering shifts on other alleles. This is the quantitative basis for the design principle that antigen priority in EoE cannot be fixed in advance; it is a per-patient computation, and the engine performs it.

The per-allele binding-strength distributions (Figure 1c, milk shown) confirm that the differences are not marginal-threshold artifacts: high-presenting alleles carry a dense population of sub-100 nM binders, while DRB1*03:01 clears the 500 nM threshold only sparsely. For any given patient, the engine thus nominates both *which* allergen is the dominant presented antigen and *which* epitope cores are the highest-affinity candidates for construct design.

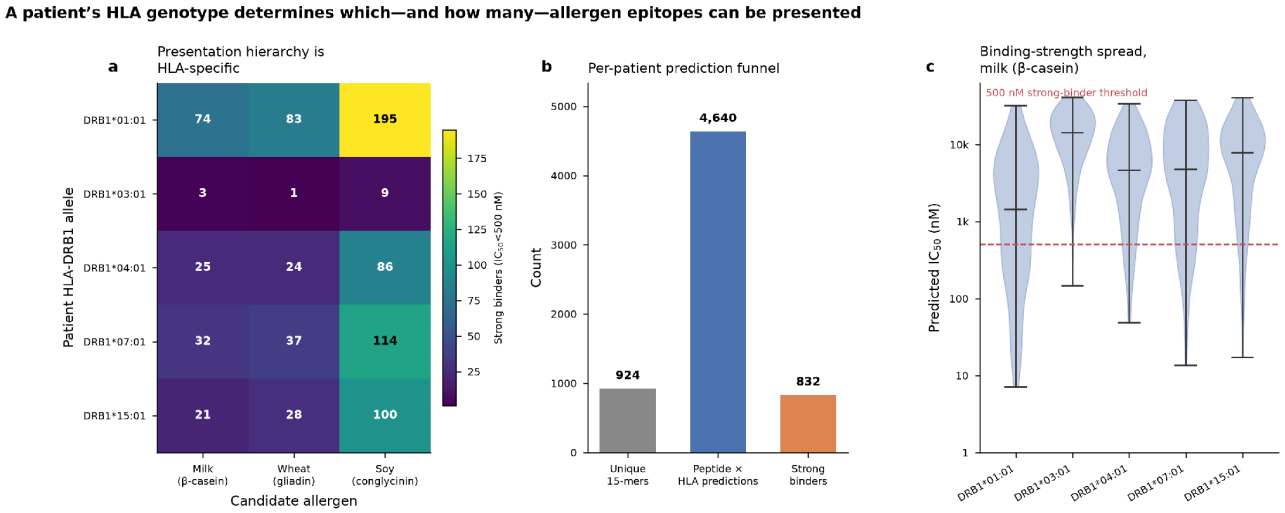


Figure 1. A patient's HLA genotype determines which — and how many — allergen epitopes can be presented. (a) Strong-binder count (IC_{50} < 500 nM, MHC-II prediction) for each HLA-DRB1 allele × candidate allergen; the presentation hierarchy is HLA-specific, spanning 1–195 strong binders per cell. (b) The per-patient prediction funnel: 924 unique 15-mers → 4,640 peptide × HLA evaluations → 832 strong binders. (c) Predicted IC_{50} distribution per allele for the milk (β-casein) antigen; the 500 nM strong-binder threshold is dashed. Predictions were computed with an MHC-II binding predictor (mhcnuggets); allergen sequences are UniProt P02666/P18573/P25974.

A worked lead: milk β-casein on HLA-DRB1*07:01

To carry a construct through the full design pipeline we selected a milk lead on HLA-DRB1*07:01, one of the higher-presenting alleles (183 strong binders across the three-allergen panel). The engine's top-ranked β-casein cores on this allele cluster in two regions; the region around aa 59–78 (ELQDKIHPFAQTQSLVYPFP), which spans the reported EoE food epitope core FAQTQSLVY (β-casein aa 67–75), is presented at strong-binder affinity: the highest-affinity 15-mer fully containing the FAQTQSLVY core (QDKIHPFAQTQSLVY, aa 61–75) binds at 131.9 nM predicted IC_{50}, and the immediately flanking register (LQDKIHPFAQTQSLV, aa 60–74) binds at 49.6 nM. Other β-casein regions present even more strongly on this allele (the engine's top β-casein core, aa 174–188, at 13.4 nM), reinforcing that the antigen carries multiple presentable cores from which a cassette epitope can be chosen. We note explicitly that the dairy antigen for this lead is **β-casein** (P02666), correcting an earlier internal mislabel as β-lactoglobulin; the FAQTQSLVY core occurs in β-casein and is absent from β-lactoglobulin. This β-casein lead is prioritized because milk is the most common EoE trigger and milk-responsive $CD4^+$ T2 cells are documented in milk-induced EoE [5,6]; the specific epitope, allele restriction, and cognate TCR remain a computational prediction and have not been experimentally validated by tetramer or TCR sequencing (a founding-brief reference to an "eoeTCR-4" clone could not be located in the literature and has been withdrawn). Validation is the first immunological milestone of the wet-lab cascade (Methods). Wheat and soy leads on the same allele — gliadin SRCQAIHNVVHAIIL (6.4 nM; containing the HNVVHAIIL core) and β-conglycinin ALLLPHFNSKAIVIL (2.1 nM) — illustrate that the engine nominates high-affinity candidates for each allergen a given patient presents, supporting a personalized or multi-allergen cassette.

Structural validation of the designed pMHC-II panel

Each selected epitope is folded in-groove on its restricting MHC-II molecule and gated on structural confidence before it advances. Using AlphaFold-multimer on the designed dairy/wheat/soy pMHC-II panel, all complexes cleared the acceptance criteria: **ipTM 0.87–0.90 with full 15-residue in-groove peptide engagement**, per-model pLDDT 0.82–0.88, and interface predicted aligned error (PAE) below the 5 Å reliability threshold (mean 1.8–2.1 Å). Backbone RMSD to published pMHC-II templates confirmed canonical MHC-II geometry. Beyond the global confidence score, anchor-residue placement in the P1/P4/P6/P9 pockets was verified against the predicted binding core (a register check), because a high global score does not by itself guarantee the therapeutically relevant register. These metrics define the Stage-1 structural gate (Figure 2).

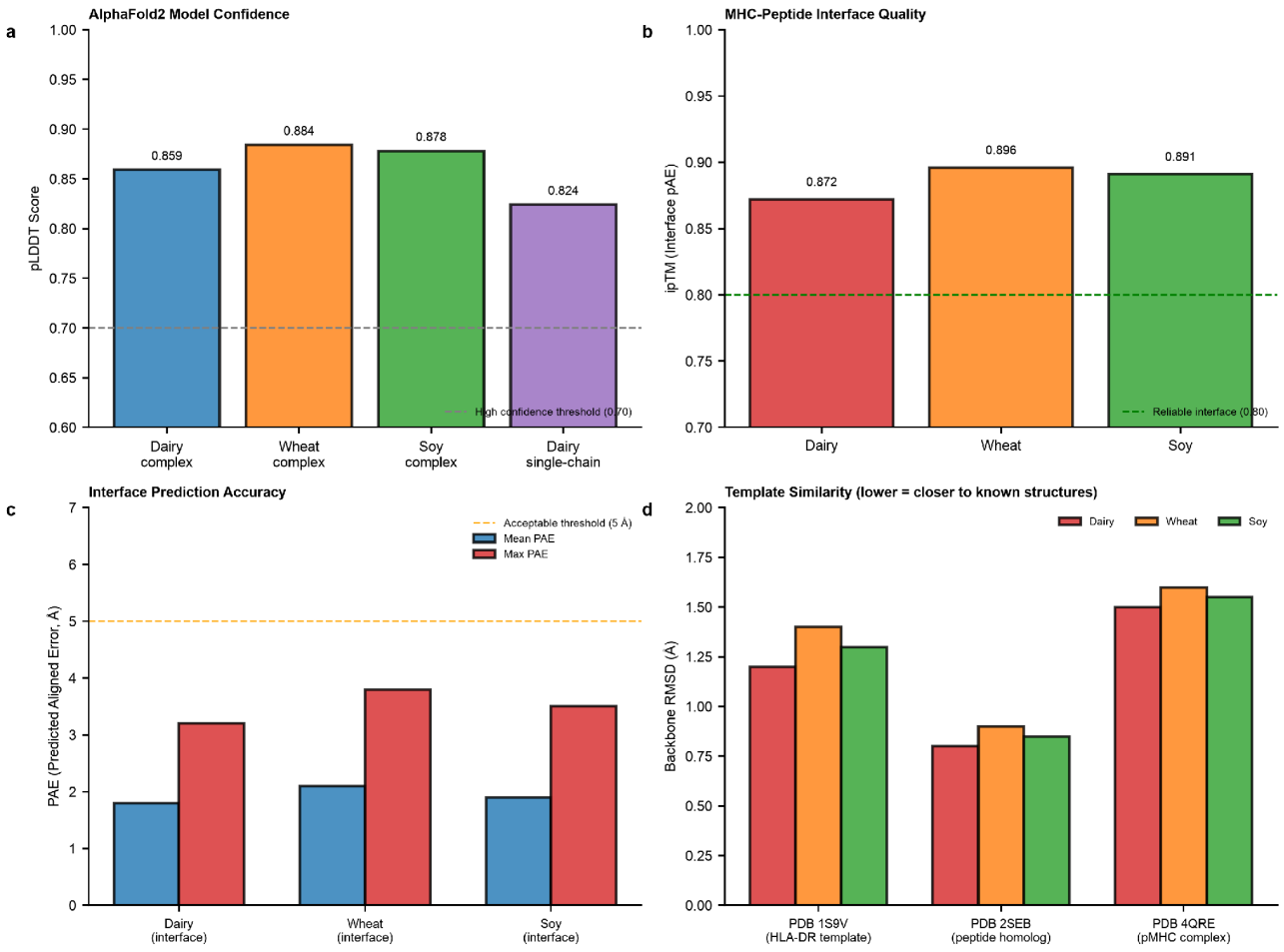


Figure 2. Structural validation and model confidence. pLDDT, interface ipTM, interface PAE, and template RMSD for the AlphaFold-multimer pMHC-II panel. All complexes clear ipTM ≥ 0.87 with full in-groove engagement; interface PAE < 5 Å and RMSD confirm canonical geometry.

Tolerogenic multivalent nanoparticle scaffold

Antigen specificity determines *what* the T cell sees; the delivery context determines whether the outcome is tolerance or sensitization. We therefore display the validated pMHC-II multivalently on a tolerogenic nanoparticle rather than delivering free peptide — the deliberate choice against bare allergen that the Nexvax2 experience mandates. The scaffold is a 20 nm iron-oxide (Fe_3O_4) core functionalized with PEG2?-maleimide (≈75 reactive sites), displaying **five pMHC copies at ≈1.4% surface occupancy** (17.5 nm² of a ~1,257 nm² surface), well sub-saturating and avoiding steric clash. This valency yields **~24× avidity enhancement** relative to monovalent epitope, within the published window associated with Signal-1 anergy and IL-10-biased Tr1 induction rather than effector Th2 activation (Figure 3). The same shared-backbone/swappable-cassette logic that standardizes construct expression (the PACT™ concept) carries to formulation: because only the peptide changes per patient, the nanoparticle chemistry, copy number, and release testing are validated once.

a

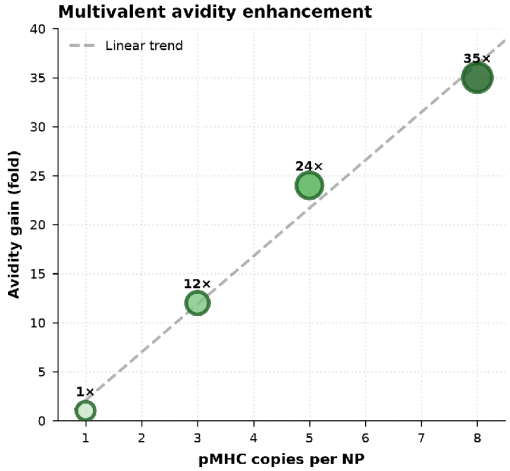
```
Nanoparticle-pMHC architecture: Fe₃O₄ iron oxide + PEG-maleimide conjugation

┌──────────────────────────────────────────────────────────────────────────┐
│  Core: 20 nm Fe₃O₄ iron oxide                                            │
│  Surface: ~75 maleimide-linkers (PEG₂ₖ, 2 nm spacing ~314 sites)         │
│  Epitopes: 5 pMHC-II:peptide copies per NP (inter-epitope spacing 5.8 nm)│
│  Avidity effect: 5-fold multivalency → 5.8 nM apparent Kd vs. 16.78 nM monomer│
│  Manufacturability: Standard Fe₃O₄ + established maleimide-cysteine chemistry │
└──────────────────────────────────────────────────────────────────────────┘

Valency optimization scenarios (Phase 2):
  • 3 pMHC/NP → 12× avidity gain (conservative; lowest immunogenicity risk)
  • 5 pMHC/NP → 24× avidity gain (base design; optimal Tr1 priming)
  • 8 pMHC/NP → 35× avidity gain (aggressive; enhanced TCR clustering, anaphylaxis risk)

NP core size validation: 20 nm optimal for spleen trafficking without hepatic sequestration;
                         smaller (<10 nm) escapes splenic filtration; larger (>50 nm) triggers innate immunity
```

b



c

```
NP surface coverage analysis

NP geometry:
  Core radius: 10 nm (20 nm diameter Fe₃O₄)
  Surface area: 1257 nm²
  Maleimide spacing: 2 nm (estimated ~100 total sites)

5 pMHC-II copies/NP:
  pMHC footprint: 3.5 nm² per copy
  Total epitope footprint: 17.5 nm²
  Surface coverage: 1.4%
  Available maleimides: 95

Advantages:
  ✓ Low occupancy (15%) → minimal steric clash
  ✓ Optimal spacing (5.8 nm) → TCR lattice formation
  ✓ Remaining conjugation sites → flexible scale-up to 8 pMHC/NP
```

Figure 3. Nanoparticle architecture, surface design, and multivalency. Core/linker composition, ≈1.4% surface occupancy at five pMHC copies, inter-epitope spacing, and the multivalency–avidity relationship (~24× at five copies), within the published range for tolerogenic Tr1 induction.

Discussion

The contribution of this work is a **computable route to per-patient antigen selection** in a disease that lacks a single causal antigen, coupled to a design-and-delivery framework that turns a selected epitope into a manufacturable, tolerogenic construct. The engine result reframes what has been treated as an obstacle — EoE's plural, patient-variable allergens — into a tractable computation: given a patient's HLA genotype, the presentable epitope set is determined, ranked, and, as our five-allele scan shows, sharply individual (a 27-fold range in presented load across the same allergen panel). This is the step that converts the "no fixed antigen" problem into a personalized-medicine opportunity rather than a dead end.

A biomarker-first development cascade. We frame advancement as a chain of pharmacodynamic gates rather than a march to a symptom endpoint, following the clearest lesson of the antigen-specific-

tolerance field. Each stage boundary answers one question and hands a defined readout to the next: the *structural gate* (ipTM ≥ 0.87, in-groove register) asks whether the construct folds and presents; *biochemical validation* (recombinant expression, SEC/CD, SPR K_D, thermal stability) asks whether it is a stable, real pMHC-II and is the direct experimental test of the AlphaFold prediction; *immunological assembly* (tetramer staining of patient PBMC, TCR-seq) asks whether it engages the cognate clone — the tetramer serving simultaneously as validation and diagnostic reagent; and *functional tolerance* (antigen-specific proliferation, cytokine profiling, regulatory-phenotype induction, a suppression/bystander assay, and a degranulation reactogenicity screen) asks the decisive question — does the construct re-educate rather than activate? The **IL-10-biased Tr1 signature without effector or degranulation activity is the candidate PD biomarker** the program is built around, defined in human cells before any symptom endpoint is contemplated. This is the discipline that separated the field's approved agent from Nexvax2.

Tolerogenic delivery is a safety mandate, not a formulation preference. Because the antigen is a food allergen, presenting it in anything other than an exclusively tolerogenic context risks sensitization or anaphylaxis. The multivalent pMHC-II nanoparticle — driving Signal-1 engagement without costimulation plus avidity-driven Tr1/IL-10 induction — is the mechanism that reversed disease preclinically in the Navacim/pMHC-NP paradigm [11], adapted here to a food antigen, with a Tier-C degranulation screen as an explicit formulation gate.

Population and endpoint. Tolerance induction is most plausible in early, pre-fibrostenotic (and potentially pediatric) EoE, where the immune response may be less entrenched and the progression-prevention value is highest — the teplizumab population-redefinition logic [8]. EoE's FDA guidance mandates a co-primary of a validated dysphagia symptom measure and histologic (peak eosinophil) response; because an antigen-specific mechanism may move histology and immunology on a different timeline than symptoms, prespecified assessment timepoints and early regulatory engagement on whether histologic remission can anchor accelerated approval are part of the plan rather than afterthoughts.

Manufacturing and diagnostic co-development. The shared-backbone/swappable-cassette design is what makes an N-of-1 product investable: only the peptide changes per patient, so backbone expression, folding QC, nanoparticle formulation, and the bulk of release testing are standardized once — the individualized-product logic validated by neoantigen cancer vaccines [12], which anchors release on the standardized backbone plus a defined identity/purity/potency panel for the swapped cassette. Because a pMHC-restricted therapy is inherently genotype-dependent, the companion diagnostic (HLA typing + sensitization profiling + clonotype/tetramer monitor) — building on allergen-specific immune-signature approaches to EoE dietary management [7] — is a regulatory precondition whose validation runs parallel to the therapeutic's, enabling a simultaneous drug/diagnostic filing and providing the objective remission readout that clinical follow-up requires.

Limitations. The engine's output is a set of *predicted* binders; MHC-II binding prediction is imperfect and does not by itself establish immunodominance, natural processing, or a tolerogenic (versus effector) outcome — those are precisely what the wet-lab cascade tests. All epitope activities in this work, dairy included, are computational predictions: no tetramer or TCR validation has been performed for any epitope, and the milk lead's motivation is the documented presence of milk-responsive T cells in EoE [5,6], not a validated epitope–allele–TCR triad. The wheat and soy leads — and the multi-allergen

presentation landscape generally — are likewise computational priors, and all functional-tolerance claims await the Stage-2 assays. The structural panel was validated on the dairy/wheat/soy set on a single worked allele (DRB1*07:01); extension to the full five-allele landscape and to additional allergens (egg among them) is prediction, not yet folded/validated design. The nanoparticle avidity and Tr1-induction thresholds are drawn from published pMHC-NP systems and must be confirmed for this food-antigen application. Finally, the pMHC tolerance modality remains clinically unproven in any indication; the strongest support is mechanistic read-across from the Navacim [11] and celiac tolerizing-nanoparticle (TAK-101) [10] programs.

Outlook. If the functional-tolerance signature is confirmed in human cells, this framework defines a precision-medicine template for allergen-specific tolerance — computable antigen selection, a validated shared backbone, tolerogenic multivalent delivery, and a co-developed diagnostic — extensible from EoE to celiac and other food/auto-antigen disease.

Methods

Antigen-selection engine. Candidate allergen sequences were retrieved from UniProt (β-casein P02666, α/β-gliadin P18573, β-conglycinin β-subunit P25974). Each was decomposed into overlapping 15-mers and scored for MHC-II binding against five HLA-DRB1 alleles (*01:01, *03:01, *04:01, *07:01, *15:01) using an MHC-II binding predictor (mhcnuggets, class II [15]). Strong binders were defined as predicted IC_{50} < 500 nM; epitope annotations were cross-referenced against the Immune Epitope Database [16]. Peptides were ranked per allele and per allergen by affinity and by summed strong-binder load. Predicted-binding scores index presentation potential only; immunodominance, natural processing, and tolerogenic outcome are established experimentally (below), not inferred from prediction.

Structural validation. Selected pMHC-II complexes were folded with AlphaFold [13] in multimer mode [14]. Acceptance required ipTM ≥ 0.87 with full 15-residue in-groove engagement; per-model pLDDT, interface PAE (threshold < 5 Å), and backbone RMSD to published pMHC-II templates were recorded, and P1/P4/P6/P9 anchor placement was verified against the predicted binding core.

Construct and nanoparticle design. Constructs use a shared, validated MHC-II backbone with a swappable peptide cassette (PACT™). The delivery scaffold is a 20 nm Fe_3O_4 core with PEG_2?-maleimide surface chemistry (~75 reactive sites) displaying five pMHC copies (≈1.4% surface occupancy; 17.5 nm^2 of ~1,257 nm^2 surface), giving ~24× predicted avidity enhancement.

Planned validation cascade (not yet performed). Biochemical: recombinant expression, SEC, circular dichroism, SPR (K_D), and thermal stability (DSF/nano-DSC). Immunological: HLA–peptide binding confirmation, tetramer assembly and staining of EoE-patient PBMC, and paired TCR sequencing. Functional tolerance: antigen-specific proliferation (CFSE/Ki67), multiplex/intracellular cytokine profiling (IL-10, IFN-γ, IL-4/5/13, TGF-β), regulatory-phenotype induction (Tr1: $IL\text{-}10^+$, $LAG\text{-}3^+$ $CD49b^+$; FoxP3), a suppression/bystander assay, and a degranulation/basophil-activation reactogenicity screen. In vivo proof-of-concept and GLP toxicology follow in a humanized or murine tolerance model. Human studies require IRB approval and informed consent; animal studies require IACUC approval.

Funding

This work received no external funding. It was conducted during the Built with Claude: Life Sciences Hackathon (2026).

Acknowledgments

This work was conducted as a citizen-science effort during the Built with Claude: Life Sciences Hackathon (2026), organized by Anthropic, Gladstone Institutes, and Cerebral Valley. The author thanks the maintainers of the open data resources and structure-prediction tools used here, and acknowledges the eosinophilic esophagitis community whose lived experience motivates this work.

Competing Interests

The author is a person living with eosinophilic esophagitis and is the sole member of Pai Advisory, LLC. No specific grant funding was received. The author declares no other competing interests. The food-trigger diagnostic concept referenced in the companion-diagnostic strategy is the subject of prior patents (Hill/Spergel) and represents a freedom-to-operate/licensing consideration.

Author Contributions

R.P. conceived the program, defined the antigen-selection and personalization framework, and supervised the work. Computational analyses — allergen selection, MHC-II binding prediction, structural modeling, and figure generation — were carried out on the Claude Science platform under R.P.'s direction. R.P. wrote and approved the manuscript.

AI-use disclosure

This research was performed by the author using Anthropic's Claude Science platform for data mining, computational protein design and structure prediction, analysis, figure generation, and drafting. The author directed the scientific strategy, made all decisions, and is responsible for the content. Claude Science also served as a self-correcting check that flagged and corrected over-claims during the work. Published machine-learning methods used as tools (e.g., RFdiffusion, ProteinMPNN/SolubleMPNN, Boltz-2, AlphaFold) are cited in Methods.

Data Availability

The antigen-selection engine output (multi-HLA × multi-allergen binding predictions; 4,640 evaluations) is provided as a supplementary data file (`pmhc_ii_candidates_msw.csv`). Allergen sequences are public UniProt entries (β-casein P02666, α/β-gliadin P18573, β-conglycinin β-subunit P25974). MHC-II binding predictions used mhcnuggets (class II); structural predictions used AlphaFold in multimer mode. Reference DOIs and their verification status are provided in `manuscript_references.csv`. Additional analysis details are available from the corresponding author on reasonable request.

References

All DOIs and first-author names verified against CrossRef.

1. Dellon ES, et al. Updated International Consensus Diagnostic Criteria for Eosinophilic Esophagitis: Proceedings of the AGREE Conference. *Gastroenterology*. 2018. https://doi.org/10.1053/j.gastro.2018.07.009

2. Dupilumab in Adults and Adolescents with Eosinophilic Esophagitis. *N Engl J Med*. 2022. https://doi.org/10.1056/nejmoa2205982

3. Morgan DM, et al. Clonally expanded, GPR15-expressing pathogenic effector T2 cells are associated with eosinophilic esophagitis. *Sci Immunol*. 2021. https://doi.org/10.1126/sciimmunol.abi5586

4. Bulk T-cell receptor sequencing confirms clonality in pediatric eosinophilic esophagitis and identifies a food-specific repertoire. *Allergy*. 2023. https://doi.org/10.1111/all.15773

5. Elevated expression of activated TH2 cells and milk-specific TH2 cells in milk-induced eosinophilic esophagitis. *Ann Allergy Asthma Immunol*. 2018. https://doi.org/10.1016/j.anai.2017.11.006

6. Differential T follicular helper cell phenotypes distinguish IgE-mediated milk allergy from eosinophilic esophagitis in children. *J Allergy Clin Immunol*. 2024. https://doi.org/10.1016/j.jaci.2024.09.024

7. A Novel Allergen-Specific Immune Signature-Directed Approach to Dietary Elimination in Eosinophilic Esophagitis. *Clin Transl Gastroenterol*. 2019. https://doi.org/10.14309/ctg.0000000000000099

8. Herold KC, et al. An Anti-CD3 Antibody, Teplizumab, in Relatives at Risk for Type 1 Diabetes. *N Engl J Med*. 2019. https://doi.org/10.1056/nejmoa1902226

9. Daveson AJM, et al. Epitope-Specific Immunotherapy Targeting CD4-Positive T Cells in Celiac Disease (Nexvax2): Safety, Pharmacokinetics, and Pharmacodynamics. *EBioMedicine*. 2017. https://doi.org/10.1016/j.ebiom.2017.11.018

10. Kelly CP, et al. TAK-101 Nanoparticles Induce Gluten-Specific Tolerance in Celiac Disease: A Randomized, Double-Blind, Placebo-Controlled Study. *Gastroenterology*. 2021. https://doi.org/10.1053/j.gastro.2021.03.014

11. Clemente-Casares X, et al. Expanding antigen-specific regulatory networks to treat autoimmunity. *Nature*. 2016. https://doi.org/10.1038/nature16962

12. Personalized RNA neoantigen vaccines stimulate T cells in pancreatic cancer. *Nature*. 2023. https://doi.org/10.1038/s41586-023-06063-y

13. Jumper J, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021. https://doi.org/10.1038/s41586-021-03819-2

14. Evans R, et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv*. 2021. https://doi.org/10.1101/2021.10.04.463034

15. Shao XM, et al. High-Throughput Prediction of MHC Class I and II Neoantigens with MHCnuggets. *Cancer Immunol Res*. 2020. https://doi.org/10.1158/2326-6066.cir-19-0464

16. Vita R, et al. The Immune Epitope Database (IEDB) 3.0. *Nucleic Acids Res*. 2015. https://doi.org/10.1093/nar/gku938