Project Tolera — Scientific Development Plan

From AI/ML design to first-in-patient: building and delivering a personalized pMHC-II tolerance vaccine for EoE

Ruth-Anne Pai, PhD — PhD Scientist & Advisor | Bridging Science and Advocacy · person living with EoE
advising@ruthannepai.com · ruthannepai.com · linkedin.com/in/ruth-anne-pai *Developed during the "Built with Claude: Life Sciences Hackathon", July 2026.*

An independent scientific roadmap — a scoped brief for any biotech partner or funder who wants to take it forward, not a clinical protocol. Preclinical-stage concept. Non-dairy epitope activity is a computational prior pending wet-lab validation; all in-patient statements are development goals, not clinical claims. Not medical advice.

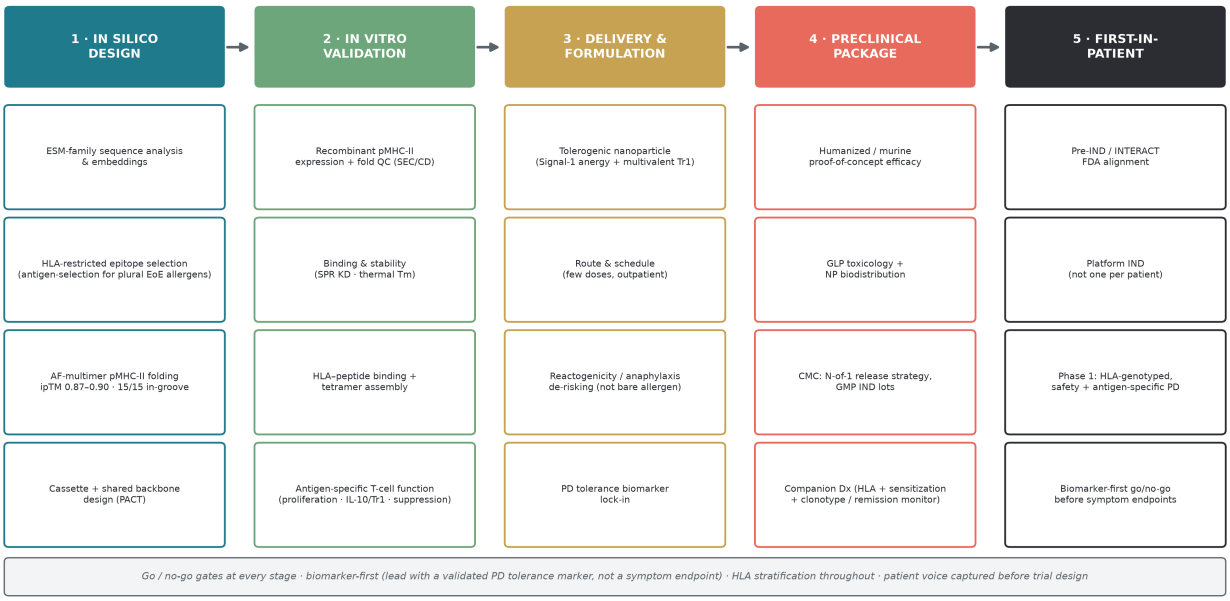
---

0. Design principles — the five lessons that shape every decision

The antigen-specific-immunotherapy landscape has produced exactly one approved agent and a field of instructive failures. Five transferable lessons are wired into this plan as first-order constraints, not afterthoughts:

1. **Biomarker-first.** Programs that led with a validated pharmacodynamic (PD) tolerance biomarker advanced (TAK-101, teplizumab); the one that bet a symptom endpoint before locking a PD marker failed (Nexvax2). We define and qualify an EoE tolerance PD marker **before** committing to a symptom endpoint.
2. **Deliver antigen in a tolerogenic context, never as bare allergen.** Nexvax2's peptide *was* the pathogen and reproduced the disease. For food allergens, anaphylaxis/reactogenicity risk is paramount — we favor a pMHC-II scaffold on a tolerogenic multivalent nanoparticle over free peptide.
3. **Stratify by HLA/sensitization from Phase 1.** A failed antigen (GAD-alum) was rescued by HLA-DR3-DQ2 responder selection and a changed delivery route. Our program is genotype-defined by construction.
4. **Target early, tolerance-tractable disease with a progression-relevant endpoint.** Teplizumab won approval by redefining the population (at-risk) and endpoint (delay of onset). We consider early / pre-fibrostenotic and pediatric EoE, where tolerance induction is most plausible.
5. **Co-build the screening/companion-diagnostic pathway.** Tzield's uptake is gated by screening, not demand. The companion diagnostic is co-developed on the clinical timeline, not behind it.

From AI/ML design to first-in-patient: the Tolera development cascade



Stage 1 — In silico design engine (AI/ML)

The design engine converts a patient's molecular profile into a manufacturable pMHC-II construct. It is the part of the platform that is already built and computing.

1.1 Antigen & epitope selection — the central EoE problem

Unlike celiac (one antigen, gluten) or T1D (defined autoantigens), EoE is driven by **plural, patient-variable food allergens** (milk, wheat, egg, soy most common), often several at once. Antigen selection is therefore the first and hardest computational task:

- **Input:** the patient's HLA class-II typing (2 haplotypes → up to 4 functional MHC-II molecules) plus an allergy/eosinophil genetic screen.
- **Epitope prediction:** every 15-mer of the candidate allergen panel is scored for MHC-II binding against *all four* of the patient's molecules (class-II binding affinity, e.g. mhcnuggets/NetMHCIIpan-class predictors). In the worked index case, 2,424 unique 15-mers × 4 molecules = 9,696 predictions yielded 1,087 strong binders (IC50 < 500 nM), with a clear per-allergen and per-molecule presentation hierarchy.
- **Prioritization:** dominant cores are ranked by summed strong-binder load and molecule-level concentration. The single most patient-specific signal in the index case — heavy gliadin presentation on the celiac-associated DQ2.2 heterodimer — illustrates that antigen priority is genuinely individual and must be computed, not assumed.

1.2 Sequence-based analysis with ESM-family models

ESM-family protein language models support (a) embedding-based similarity and self-vs-food discrimination to flag cross-reactive or self-mimicking epitopes, (b) variant-effect and stability scoring of candidate constructs, and (c) prioritization of designed cassette sequences for expressibility and developability before any structure is computed.

1.3 Structural validation — AlphaFold-multimer pMHC-II folding

Each candidate peptide is folded in-groove on its restricting MHC-II molecule with AlphaFold-multimer:

- **Acceptance gate:** ipTM ≥ 0.87 and full 15-residue groove engagement. The current 9-construct dairy/wheat/soy panel folds at **ipTM 0.87–0.90 with 15/15 peptides in-groove**.
- **Register check:** anchor-residue placement (P1/P4/P6/P9 pockets) is verified against the predicted binding core, not merely the global confidence score.

1.4 PACT™ cassette + shared backbone design

Constructs are built as a **shared, validated MHC-II backbone with a swappable peptide cassette** — the manufacturing keystone: only the peptide changes per patient, so backbone expression, folding QC, and formulation are standardized once. GPU design status: the dairy backbone is validated; the wheat/soy cassettes are computational priors pending the wet-lab work below.

Stage 1 go/no-go: a patient-matched construct passes only if (a) a dominant epitope clears the binding threshold on a typed molecule, (b) the folded pMHC-II clears ipTM/register gates, and (c) ESM developability flags are clear.

Stage 2 — In vitro validation cascade

This is the assay cascade that converts a computational prior into a validated cassette — the largest single line in the Seed budget and the primary Seed value driver. The cascade is designed as a **gated funnel**: a construct must pass each tier to advance, and failures are informative (they reprioritize the epitope list).

2.1 Tier A — biochemical: does the pMHC-II fold and hold the peptide?

Assay	Readout	Pass criterion (target)
Recombinant expression (mammalian/insect)	Yield, monodispersity	Expressible, single species
Size-exclusion chromatography (SEC)	Aggregation state	Monomeric / defined oligomer, no aggregation
Circular dichroism (CD)	Secondary-structure fold	Correct MHC-II fold signature
Peptide–MHC binding (SPR)	**K_D**	High-affinity, stable complex (target K_D consistent with a stable tetramer)
Thermal stability (DSF/nano-DSC)	**T_m**	Peptide-loaded complex thermostabilized vs. empty

Tier A answers the most basic manufacturability question — *does this construct exist as a stable, correctly-folded pMHC-II?* — and is the direct experimental test of the AlphaFold prediction.

2.2 Tier B — immunological assembly: can it engage cognate T cells?

Assay	Readout	Purpose
HLA–peptide binding confirmation	Competitive binding / stabilization	Confirms the predicted restriction element
Tetramer assembly & staining	Cognate $CD4^+$ T-cell detection (patient PBMC)	Detects and quantifies the allergen-specific clone — the platform's core reagent
Clonotype identification (TCR-seq)	Paired αβ TCR sequences	Defines the pathogenic clone; feeds the companion Dx clonotype monitor

Tetramer staining is the pivotal Tier-B assay: it is simultaneously a **validation reagent** (does the construct bind the intended T cells?) and a **diagnostic reagent** (how many such cells does this patient carry?).

2.3 Tier C — functional tolerance: does it re-educate, not activate?

This tier is the mechanistic proof-of-concept and the source of the **PD tolerance biomarker** the whole program is sequenced around.

Assay	Readout	Tolerance signature (goal)
Antigen-specific proliferation (CFSE/Ki67)	Cognate T-cell division	*Reduced* proliferation on re-challenge (anergy)
Cytokine profiling (multiplex/ intracellular)	IL-10, IFN-γ, IL-4/5/13, TGF-β	**IL-10-biased Tr1 signature**; suppression of effector Th2 cytokines
Regulatory-phenotype induction	Tr1 (IL-10$^+$, LAG-3$^+$CD49b$^+$), FoxP3	Induction of regulatory over effector phenotype
Suppression / bystander assay	Suppression of effector response	Functional tolerance, not just clonal silencing
Reactogenicity screen	Degranulation / basophil activation (allergen context)	**No** effector/anaphylactic activation — the Nexvax2 safety lesson

Stage 2 go/no-go (the Seed's central de-risking event): a cassette is "validated" (Tier-A/B/C clear) when it folds and holds peptide, assembles a tetramer that stains cognate cells, and drives an IL-10-biased Tr1/ suppressive signature **without** effector or degranulation activity. This functional-tolerance signature is the candidate **PD biomarker** carried into the clinic.

Stage 3 — Delivery & formulation

Antigen specificity is necessary but not sufficient; *how* the pMHC-II is delivered determines whether the outcome is tolerance or sensitization.

- **Tolerogenic multivalent nanoparticle.** The pMHC-II is displayed multivalently on a nanoparticle to drive **Signal-1 (TCR) engagement without costimulation → anergy**, plus avidity-driven **Tr1/IL-10 induction** — the Navacim/pMHC-NP mechanism that reversed disease preclinically, adapted to a food antigen. This is the deliberate choice *against* bare allergen peptide that the Nexvax2 failure mandates.
- **Route & schedule.** Designed for **few doses, outpatient administration** (subcutaneous/intradermal candidate routes), explicitly to avoid the adoption penalty of Tzield's 14 daily infusions — a direct input from the patient-burden analysis.
- **Reactogenicity/anaphylaxis de-risking.** Dose-escalation design, exclusion of high-anaphylaxis-risk patients early, and the Tier-C degranulation screen as a formulation gate. Allergen delivery carries real anaphylaxis risk; the formulation is engineered to present antigen in an exclusively tolerogenic context.
- **PD biomarker lock-in.** The tolerance signature from Stage 2 is formalized here into a clinical-grade assay (allergen-specific T-cell frequency + cytokine polarization + esophageal histology correlate) so it is deployable in Phase 1.

Stage 4 — Preclinical package (IND-enabling)

Workstream	Deliverable	Notes
Efficacy PoC	In vivo tolerance/efficacy signal	Humanized or murine model; demonstrate antigen-specific tolerance, ideally in an early-disease paradigm

Workstream	Deliverable	Notes
GLP toxicology	Safety package for FIH	Repeat-dose tox on the individualized product class
NP biodistribution	Where the nanoparticle goes	Tolerogenic-niche targeting; clearance
CMC	Individualized-manufacturing control strategy	Shared-backbone release + peptide-cassette identity/QC; GMP IND lots; N-of-1 release strategy
Companion diagnostic	CLIA-grade assay	HLA typing + allergen-sensitization + clonotype/remission monitor

The **CMC / N-of-1 release strategy is the make-or-break regulatory-science problem** for a personalized product: because only the peptide changes, release testing is anchored on the standardized backbone plus a defined identity/purity/potency panel for the swapped cassette — the individualized-product framework validated by neoantigen cancer vaccines.

Stage 4 go/no-go: a clean GLP tox package + a manufacturable, release-testable individualized product + an agreed FDA CMC/endpoint position = IND-ready.

Stage 5 — First-in-patient

- **Regulatory entry.** Pre-IND / INTERACT alignment on endpoints and the individualized-manufacturing control strategy; file a **platform IND** (not one IND per patient). Seek Orphan Drug → Fast Track, with Breakthrough on early signal and RMAT tested explicitly.
- **Phase 1 design.** HLA-genotyped participants; primary readouts are **safety/tolerability and antigen-specific PD** (the Stage-2 tolerance biomarker: allergen-specific T-cell frequency/polarization), *not* a symptom endpoint. This is the biomarker-first discipline that separated the field's winners from Nexvax2.
- **Population.** Prioritize early / pre-fibrostenotic (and potentially pediatric) EoE, where tolerance induction is most tractable and the progression-prevention value is highest — the teplizumab population-redefinition lesson.
- **Endpoint architecture toward Phase 2/3.** EoE's FDA guidance mandates a co-primary of a validated dysphagia symptom COA *and* histologic (peak eosinophil count) response. Because an antigen-specific mechanism may change histology/immunology on a delayed timeline relative to symptoms, we will prespecify assessment timepoints and engage FDA early on whether **histologic remission can anchor Accelerated Approval** with dysphagia confirmation as a post-marketing requirement.

---

6. The translational biomarker strategy (the spine of the plan)

Every stage boundary is a biomarker checkpoint, and they chain into one continuous PD story:

Stage	Biomarker / gate	Question answered
1	ipTM ≥ 0.87, in-groove register, binding IC50	Will it fold and present?
2A	K_D, T_m, SEC/CD	Is it a stable, real pMHC-II?
2B	Tetramer^{+} cognate frequency	Does it engage the right T cells?
2C	IL-10/Tr1 signature + suppression, no degranulation	Does it tolerize rather than activate?
3	Clinical-grade tolerance assay	Is the PD marker deployable?
4	In vivo tolerance + histology correlate	Does tolerance translate in vivo?
5	Antigen-specific PD in humans → histologic remission	Does it work in patients, biomarker-first?

The single sentence: *lead with a validated pharmacodynamic tolerance biomarker, deliver allergen in a tolerogenic (non-reactogenic) context, stratify by HLA/sensitization, target early disease with a progression-relevant endpoint, and co-build the screening pathway — doing what the field's winners did and avoiding what sank Nexvax2.*

---

7. Companion diagnostic — a scientific deliverable, not just a commercial one

The companion diagnostic is co-developed across Stages 2–4 and serves three scientific functions:

1. **Patient selection** — HLA typing + allergen-sensitization profiling defines who has a computable, addressable epitope (and on which molecule).
2. **Clone identification** — TCR-seq/tetramer identifies the pathogenic clonotype that the N-of-1 cassette must target.
3. **Remission monitoring** — the same tetramer/clonotype assay provides the objective, low-cost, repeatable readout that operationalizes outcomes-based reimbursement and, clinically, tracks durable tolerance off-therapy.

Because a pMHC-restricted vaccine is inherently genotype-dependent, the diagnostic is a **regulatory precondition** whose validation timeline runs parallel to the therapeutic's — enabling a simultaneous drug/diagnostic filing.

---

8. Key scientific risks & mitigations

Risk	Severity	Mitigation
No single dominant EoE allergen (antigen heterogeneity)	High	Omics-driven per-patient epitope prioritization; multi-allergen/personalized cassette; start milk/wheat-dominant subset
Anaphylaxis / reactogenicity from allergen delivery	High	Tolerogenic pMHC-NP (not bare peptide); Tier-C degranulation gate; dose-escalation; risk-based exclusions
No validated EoE tolerance PD biomarker exists yet	High	Build and qualify it in Stage 2C *before* any symptom endpoint (the program's defining discipline)
pMHC modality clinically unproven anywhere	Med–High	Navacim / celiac-NP (TAK-101, KAN-101) mechanistic read-across; strong preclinical package
N-of-1 CMC / release testing	Med–High	Shared-backbone control strategy; neoantigen-vaccine individualized-product precedent; early FDA CMC alignment
Co-primary symptom COA is demanding	Medium	Patient-co-designed PRO; pursue histologic surrogate under Accelerated Approval; prespecify kinetics timepoints
Wheat/soy epitopes are unvalidated priors	Medium	Explicit funded Seed wet-lab validation; dairy anchor (literature-reported) leads

---

9. The nearest, highest-value next step

The single most valuable experiment this roadmap points to converts **computational designs into human-cell-validated, IND-enabling data** and stands up the companion diagnostic — carrying the concept from a validated in-silico panel toward a **cleared platform IND and first-in-human dosing**. The gating scientific deliverable is the **Stage 2C functional-tolerance signature on Tier-A/B validated cassettes**: it would de-risk the mechanism in

human cells, define the clinical PD biomarker, and form the evidence base for Fast Track / Breakthrough engagement. This is where a researcher, clinician, or biotech partner could make the largest difference — contact: **advising@ruthannepai.com** · ruthannepai.com · linkedin.com/in/ruth-anne-pai.

All in-vitro pass criteria are target specifications and would be finalized with assay development. Non-dairy epitopes are computational priors pending the Stage-2 validation described here. The β-casein anchor's literature attribution is carried from the founding brief and flagged for independent citation verification. This is an independent research roadmap, not a clinical protocol. Not medical advice.