Per-Residue Immunogenicity Risk Classification

of Manufacturing Impurities in Semaglutide

Substitution, Deletion, and Insertion Sensitivity Profiles
via Mint_IRI Pipeline Scoring

WHITE PAPER

Mint Precision Analytical

Zain Tariq, MD

S. Mustafa Zaidi, MD

April 2026

Abstract

Manufacturing impurities in therapeutic peptides can alter MHC binding profiles, potentially introducing neo-epitopes that trigger unwanted immune responses. Semaglutide, a 31-amino-acid GLP-1 receptor agonist, is among the most widely prescribed peptide therapeutics globally, with patient exposure continuing to expand across metabolic and cardiovascular indications. As compounded and generic versions enter production, the scale of patient exposure elevates the importance of characterizing whether manufacturing impurities introduce immunogenic neo-epitopes not present in the parent sequence. This white paper presents a computational methodology for mapping per-residue immunogenicity risk across three classes of sequence-altering SPPS impurities: amino acid substitutions (589 variants), single-residue deletions (31 variants), and self-insertions (31 variants).

All 651 impurity variants were scored through the Mint_IRI platform using NetMHCpan 4.2 and NetMHCIIpan 4.3 across a 14-allele HLA panel covering greater than 95% of major world populations. The three impurity types produce strikingly different risk landscapes, with deletion the most uniformly dangerous and substitution providing the richest risk differentiation. Cross-impurity analysis identifies positions 11 through 15 (the SSYLE segment) as the highest-priority region in the semaglutide sequence: all three impurity types produce elevated Class I risk, and substitution and deletion independently produce elevated Class II risk at these positions.

The sensitivity maps represent the hazard component of the immunogenicity risk equation. When specific impurities are subsequently identified by LC-MS in manufacturing batches, the same pipeline scores the observed variant individually, producing a batch-specific risk classification. Both analytical modes employ identical scoring logic through the same platform.

For manufacturing, the maps identify the coupling cycles and sequence positions where process control has the most direct relationship to patient immunological safety. The universal hotspot at positions 11 through 15 identifies where process control has the most direct relationship to patient immunological safety. Position-specific risk tiers can support scientifically justified acceptance criteria that reflect predicted immunogenic consequences at each position, rather than uniform thresholds applied across the sequence. The complementary variable is position-specific impurity frequency under real manufacturing conditions which is not available in the published literature and represents a field-level gap that CDMO manufacturing data would address.

The algorithmic architecture has been retrospectively evaluated against T-cell-confirmed immunogenicity outcomes from two independent, FDA-funded experimental datasets. Under the current prediction engine, 6 of 8 T-cell-confirmed immunogenic impurities were classified at MODERATE or above. The two misses are attributable to a defined scope limitation (TCR-mediated immunogenicity) and a geometric constraint (near-terminal fragment coverage), respectively. Neither reflects a failure of MHC binding prediction.

1. Introduction

1.1 The Gap: No Advance Intelligence on Position-Level Immunogenicity Risk

Peptide therapeutics manufactured via solid-phase peptide synthesis (SPPS) contain process-related impurities at every stage of production. Deletion sequences, substitution variants, insertion sequences, and chemical degradation products are well-documented in the SPPS impurity literature and are routinely detected by LC-MS during quality control.

Current approaches to assessing these impurities are reactive. Impurities are identified analytically, characterized, and evaluated on a case-by-case basis. This workflow answers the question ‘what impurities are present in this batch?’. It does not address a more fundamental question: where in the peptide sequence is a manufacturing error most likely to produce an immunogenically dangerous impurity?

This distinction matters because not all impurities carry equal immunogenicity risk. Peptide impurities that alter the amino acid sequence of a therapeutic can change how the molecule is processed by the immune system. Major Histocompatibility Complex (MHC) molecules on the surface of antigen-presenting cells bind short peptide fragments and display them to T cells. If an impurity introduces a new peptide fragment that binds an MHC molecule where the parent therapeutic did not, this creates a neo-epitope: a new immune-visible signal that can initiate an adaptive immune response.

This is the mechanism by which manufacturing impurities can trigger anti-drug antibody (ADA) formation. ADA formation is driven by CD4+ T helper cell activation via MHC Class II presentation, though MHC Class I presentation to CD8+ cytotoxic T cells can also contribute to the overall immune response. Conversely, an impurity that does not alter MHC binding is silent from the perspective of MHC-mediated adaptive immunity, though innate immune activation through mechanisms such as aggregation or pattern recognition remains possible and is outside the scope of this analysis.

A per-residue risk classification map addresses this gap by computationally scanning every position in the peptide sequence before any specific impurity is observed. The output identifies positions where quality control should be most vigilant. It enables risk-informed acceptance criteria for batch release, and provides CDMOs and compounding pharmacies with actionable data for process optimization.

The per-residue sensitivity maps and the assessment of individual observed impurities represent two complementary analytical modes addressing the same underlying question through different entry points. The sensitivity maps presented in this white paper operate in advance of manufacturing: they exhaustively scan the full space of hypothetical impurities to identify positions where the immunogenic potential of a manufacturing error is highest. This is a hazard identification exercise. When a specific impurity is subsequently detected by LC-MS in a manufacturing batch, the same architecture scores that specific impurity against the same HLA panel and binding thresholds, producing a risk classification for the actual observed variant at its

actual position. The first mode identifies where to look; the second mode evaluates what is found. Both modes are generated by the same production platform and employ identical scoring logic, ensuring consistency between the advanced intelligence and the batch-specific assessment.

1.2 Scope of This Work

This white paper presents the methodology and results for generating per-residue risk classification maps across three impurity types for semaglutide: substitution (589 variants), deletion (31 variants), and insertion (31 variants). Each variant is scored through the Mint_IRI platform, producing independent MHC Class I and Class II risk classifications. The three impurity types employ different scoring architectures: substitution uses binder-gated paired comparison against the parent sequence, while deletion and insertion use novel fragment assessment for frame-shifted fragments that have no parent counterpart.

2. Methodology

2.1 Peptide Under Analysis

Peptide Name	Semaglutide
Length	31 amino acids
Sequence	H-Aib-EGTFTSDVSSYLEGQAAKEFIAWLVRGRG
Non-Standard Residue	Aib at position 2
Proxy Handling	Aib represented as Alanine in both parent and impurity
Manufacturing Method	Solid-phase peptide synthesis (SPPS)

The semaglutide sequence contains Aib at position 2, a non-standard amino acid that cannot be directly processed by MHC binding prediction algorithms. This residue is represented as Alanine in both parent and impurity sequences. Because the same proxy substitution appears on both sides of the parent-versus-impurity comparison, the proxy approximation error cancels in the delta calculation (delta cancellation principle). This is documented via an informational flag but does not affect risk scoring or confidence.

Semaglutide has demonstrated low clinical immunogenicity, with anti-semaglutide antibody formation observed in approximately 1% of patients across clinical trials (Ozempic prescribing information; MedCentral drug monograph). This low baseline immunogenicity is relevant to the impurity risk assessment. The 1% ADA rate indicates that the semaglutide API sequence itself presents minimal immunogenic risk. Manufacturing impurities that introduce novel MHC binding events not present in the parent sequence would therefore represent a qualitatively distinct source of immunogenic risk that is not captured by the clinical ADA profile of the unmodified API. By contrast, for a highly immunogenic peptide such as salmon calcitonin, where over 60% of patients develop neutralizing anti-drug antibodies, impurity-derived epitopes are introduced into an already active immune environment. The sensitivity maps presented here identify positions where manufacturing impurities could introduce MHC binding events that are absent from the low-immunogenicity parent API.

2.2 Impurity Simulation Design

The three impurity types profiled in this analysis (substitution, deletion, and insertion) represent the major classes of sequence-altering manufacturing impurities in SPPS. These impurities change the amino acid identity, length, or reading frame of the peptide, directly altering the set of fragments available for MHC binding. A separate class of SPPS impurities involves chemical modifications that alter amino acid side-chain chemistry without changing the sequence length or reading frame. These include deamidation (Asn→Asp, Gln→Glu), oxidation (Met, Trp, Cys), racemization (all chiral residues), and incomplete side-chain deprotection. Deamidation can be modeled within the substitution scoring framework because the products (Asp and Glu) are standard amino acids with well-characterized MHC binding properties; this extension is under development as a separate module. Oxidation and racemization introduce non-standard amino acid forms that fall outside the 20-residue alphabet of current MHC binding prediction algorithms and require different modeling approaches. The present analysis is therefore comprehensive for sequence-altering impurities but does not address the full chemical modification landscape.

2.2.1 Substitution Variants (589 total)

At each of the 31 amino acid positions, all 19 alternative standard amino acids are tested as hypothetical single-residue substitution impurities. Each simulated variant differs from the parent sequence at exactly one residue while preserving the original sequence length. The total substitution simulation space is 31 positions times 19 substitutions, yielding 589 impurity variants. For each position, the worst-case risk across all 19 variants determines the position-level risk classification.

2.2.2 Deletion Variants (31 total)

At each of the 31 positions, the native residue is removed, producing an impurity sequence of length 30. Each deletion variant is scored once through the pipeline. Unlike substitution, there is no multi-variant aggregation: each position receives a single direct risk classification from the pipeline. Deletion impurities create frame-shifted novel fragments downstream of the deletion site that have no positional counterpart in the parent sequence.

2.2.3 Insertion Variants (31 total)

At each of the 31 positions, the native residue is duplicated in place (self-insertion), producing an impurity sequence of length 32. For example, position 9 (Asp) generates the sequence with two adjacent Asp residues at that site. Each insertion variant is scored once through the pipeline. Self-insertion preserves the amino acid identity at the modification site while frame-shifting all downstream fragments. This design isolates the immunogenic impact of the frame shift from the impact of introducing a foreign amino acid.

2.3 HLA Allele Panel

MHC binding predictions are performed across a 14-allele HLA panel providing greater than 95% estimated cumulative coverage of major world populations. The 8 MHC Class I alleles were selected based on the supertype classification of Sette and Sidney (1999), which demonstrated that nine major HLA Class I supertypes account for the vast preponderance of HLA-A and HLA-B polymorphism. The 6 MHC Class II alleles were selected based on the functional classification of Greenbaum et al. (2011). Class I and Class II results are scored independently throughout the analysis pipeline.

Table 1. MHC Class I Allele Panel (8 alleles)

Allele	Supertype	Prediction Tool
HLA-A*01:01	A01	NetMHCpan 4.2
HLA-A*02:01	A02	NetMHCpan 4.2
HLA-A*03:01	A03	NetMHCpan 4.2
HLA-A*24:02	A24	NetMHCpan 4.2
HLA-B*07:02	B07	NetMHCpan 4.2
HLA-B*08:01	B08	NetMHCpan 4.2
HLA-B*15:01	B62	NetMHCpan 4.2
HLA-B*40:01	B44	NetMHCpan 4.2

Table 2. MHC Class II Allele Panel (6 alleles)

Allele	Prediction Tool
HLA-DRB1*01:01	NetMHCIIpan 4.3
HLA-DRB1*03:01	NetMHCIIpan 4.3
HLA-DRB1*04:01	NetMHCIIpan 4.3
HLA-DRB1*07:01	NetMHCIIpan 4.3
HLA-DRB1*15:01	NetMHCIIpan 4.3
HLA-DPA1*01:03/DPB1*04:01	NetMHCIIpan 4.3

The Class II panel emphasizes HLA-DR alleles, consistent with the established dominance of HLA-DR in the presentation of peptide therapeutics to CD4+ T cells and with the allele focus of the FDA-funded validation datasets used for retrospective concordance (Section 4.5). A single HLA-DP allele (DPA101:03/DPB104:01) is included to capture DP-restricted presentation. HLA-DQ alleles are not included in the current panel. While DQ-restricted epitopes have been documented for some antigens, their contribution to peptide drug immunogenicity has not been systematically characterized, and DQ representation in MHC binding prediction training data remains sparser than for DR. Expansion of the panel to include DQ alleles is feasible within the

existing pipeline architecture and would be warranted if DQ-restricted responses are identified as clinically relevant for GLP-1 analogs.

2.4 Fragment Generation

For each simulated impurity variant, overlapping peptide fragments are generated using a sliding-window approach at biologically relevant lengths: 8, 9, 10, and 11-mer for MHC Class I, and 15-mer for MHC Class II.

For substitution variants, fragments are generated from both parent and impurity sequences, and only fragments overlapping the modification position are included. For deletion and insertion variants, fragments are categorized positionally into tracks: PRE-modification fragments (identical to parent, automatic LOW risk) and SPANNING/POST-modification fragments (novel, assessed against absolute binding thresholds).

2.5 MHC Binding Prediction

Binding predictions are performed using locally installed DTU standalone software: NetMHCpan 4.2 for MHC Class I and NetMHCIIpan 4.3 for MHC Class II. Both tools return three values per prediction: Eluted Ligand percentile rank (EL %Rank), binding affinity in nanomolar (IC50), and Binding Affinity percentile rank (BA %Rank). The primary metric used for all scoring decisions is the EL %Rank.

NetMHCpan 4.2 and NetMHCIIpan 4.3 were selected as the prediction engines for this pipeline because they are the current releases of the most widely validated publicly available MHC binding prediction tools, trained on the largest available corpora of mass-spectrometry-identified eluted ligand data and binding affinity measurements. The use of publicly available prediction tools (NetMHCpan 4.2, NetMHCIIpan 4.3) ensures that the underlying binding predictions are generated by independently validated, peer-reviewed algorithms, and that the prediction layer can be independently audited. The risk classification architecture that includes the binder-gated scoring logic, novel fragment assessment, epitope density evaluation, and concordance matrix is proprietary to Mint Precision Analytical.

2.6 Risk Classification via Production Pipeline

Each impurity variant is scored through the Mint_IRI platform. This is the same pipeline used for real client impurity assessments.

2.6.1 Substitution: Binder-Gated Risk Scoring

For substitution variants, matched parent-impurity fragment pairs are evaluated for binder-gated threshold crossing. A strong new binder event is defined as: parent fragment EL %Rank greater than 2.0% (non-binder) AND impurity fragment EL %Rank less than or equal to 0.5% (strong binder). A weak new binder event uses the 2.0% threshold for the impurity. Class I and Class II are assessed independently. Any strong new binder on any allele triggers automatic HIGH risk (severity rule). 4 or more alleles with new binders triggers HIGH (breadth rule). 2 to 3 alleles trigger ELEVATED. 1 allele triggers MODERATE. 0 alleles result in LOW.

2.6.2 Deletion and Insertion: Novel Fragment Risk Scoring

Deletion and insertion variants use a two-track scoring architecture. Track A covers PRE-modification fragments that are identical to the parent sequence (automatic LOW). Track B covers SPANNING and POST-modification fragments that are novel (no parent counterpart exists). Track B fragments are assessed against absolute binding thresholds, with Class I and Class II scored independently. The combined risk is MAX(Track A, Track B). Confidence baseline for deletion and insertion is MODERATE (lower than substitution's HIGH baseline), reflecting the inherent uncertainty of frame-shifted novel fragment assessment.

The absolute threshold architecture for deletion and insertion scoring is inherently less specific than binder-gated scoring. Without parent gating, any novel fragment that crosses the binding threshold is flagged regardless of whether the parent sequence also produces comparable binders at overlapping positions through a different reading frame. This architectural difference contributes to the MODERATE confidence baseline assigned to deletion and insertion classifications, and it means that the false-positive rate for novel fragment assessment is expected to exceed that of binder-gated substitution scoring. Systematic characterization of this differential would require scoring deletion and insertion variants on control peptides with known immunogenicity profiles and comparing the predicted-to-observed concordance rates between the two scoring modes. The available experimental datasets (four TPT deletions, two SCT insertions) provide a preliminary basis but are insufficient for powered comparison. Until additional experimental datasets become available, the confidence baselines (HIGH for substitution, MODERATE for deletion/insertion) serve as the primary mechanism for communicating the differential analytical certainty between the two scoring architectures.

2.6.3 Epitope Density Assessment and Concordance

The risk classifications reported in this white paper are not produced by the primary scoring algorithms alone. Each impurity variant passes through an additional epitope density assessment that asks a complementary question: how does the overall burden of MHC binding fragments per residue of sequence length change between parent and impurity?

The primary scoring algorithms (Sections 2.6.1 and 2.6.2) identify specific threshold-crossing events: individual fragments that gain MHC binding capacity the parent did not have. Epitope density measures the aggregate strong binder count normalized by sequence length for both parent and impurity, computed independently for Class I and Class II. A positive delta indicates the impurity presents a denser epitope landscape than the parent.

Because both metrics are derived from the same binding prediction data but measure different properties of that data, they can agree or disagree. A concordance step reconciles the two signals into the final baseline risk and baseline confidence reported in the results tables. The key rules governing this reconciliation are:

The primary algorithm risk is dominant and cannot be overridden downward by epitope density. If the primary algorithm classifies an impurity as HIGH, the baseline risk is HIGH regardless of the density signal.

Epitope density can escalate risk in specific cases where the primary algorithm detects minimal threshold-crossing activity but the density assessment identifies a substantial increase in overall

epitope burden. This captures the binder-strengthening pattern: impurities where existing weak MHC binders become strong binders without creating new binding events that the threshold-crossing logic would detect.

When both signals agree, confidence increases. When they disagree, confidence decreases. For deletion and insertion variants, the concordance step starts from a MODERATE confidence baseline (reflecting the inherent uncertainty of novel fragment assessment), which is why the majority of deletion and insertion positions in Tables 5 and 6 report LOW or MODERATE confidence even when the risk tier is HIGH.

The risk tier and confidence level are complementary dimensions of the assessment and should be read together. A HIGH-risk, LOW-confidence classification indicates that the binding prediction layer identified strong novel MHC binders, but the analytical certainty is lower than for a substitution variant at the same risk tier.

3. Results

3.1 Summary Statistics

All 651 impurity variants were scored successfully through the production Mint_IRI pipeline with zero failures. 1,211 automated tests pass with zero regressions.

Table 3. Per-Position Risk Distribution by Impurity Type

	LOW	MOD	ELEV	HIGH
Subst. Class I (worst of 19)	1	2	3	25
Subst. Class II (worst of 19)	10	10	6	5
Deletion Class I	6	2	4	19
Deletion Class II	26	0	3	2
Insertion Class I	5	6	6	14
Insertion Class II	29	2	0	0

The three impurity types produce fundamentally different risk landscapes. Substitution shows the richest Class II differentiation (four-tier spread). Deletion is the most uniformly dangerous (20 of 31 positions HIGH on combined Mint_IRI seen in table 5). Insertion shows a near-even four-tier spread on Class I (5/6/6/14) while Class II is nearly silent (29 of 31 LOW).

An important interpretive distinction applies across impurity types. Substitution classifications carry a HIGH confidence baseline, reflecting the paired parent-versus-impurity comparison that inherently filters binding noise. Deletion and insertion classifications carry a MODERATE confidence baseline, reflecting the inherent uncertainty of assessing novel frame-shifted fragments against absolute binding thresholds without parent gating. After reconciliation with the epitope density assessment, the majority of deletion and insertion positions report LOW confidence (Tables 5 and 6).

3.2 Substitution Sensitivity Map

All 589 substitution variants were scored successfully. Class I shows 25 of 31 positions at HIGH worst-case risk, reflecting the high sensitivity of the binder-gated severity rule. Class II shows substantially more differentiation: 10 LOW, 10 MODERATE, 6 ELEVATED, and 5 HIGH positions. A contiguous dual-class hotspot spans positions 7 through 16, with position 19 also reaching dual-class HIGH. Position 9 (Asp) is the dominant Class II hotspot, with 9 of 19 substitutions producing HIGH Class II risk.

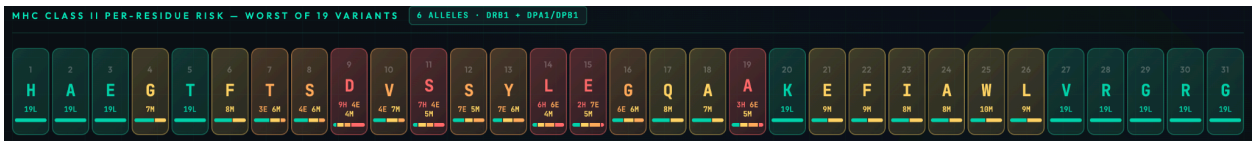

Figure: MHC Class II Per-Residue Substitution Risk Heatmap

Table 4. Complete Per-Position Substitution Risk Classification

Pos	AA	CI	CII	IRI	CI Dist	CII Dist
1	H	LOW	LOW	LOW	19L	19L
2	A	MOD	LOW	MOD	17L 2M	19L
3	E	MOD	LOW	MOD	15L 4M	19L
4	G	HIGH	MOD	HIGH	7L 8M 2E 2H	15L 4M
5	T	HIGH	LOW	HIGH	4L 6M 5E 4H	19L
6	F	HIGH	MOD	HIGH	3L 3M 4E 9H	18L 1M
7	T	HIGH	ELEV	HIGH	1L 4M 8E 6H	12L 5M 2E
8	S	HIGH	ELEV	HIGH	4L 9M 5E 1H	13L 5M 1E
9	D	HIGH	HIGH	HIGH	4L 7M 4E 4H	5L 3M 2E 9H
10	V	HIGH	ELEV	HIGH	9L 6M 3E 1H	9L 9M 1E
11	S	HIGH	HIGH	HIGH	5L 4M 5E 5H	4L 3M 5E 7H
12	S	HIGH	ELEV	HIGH	2L 2M 7E 8H	10L 5M 4E
13	Y	HIGH	ELEV	HIGH	2L 4M 2E 11H	12L 4M 3E
14	L	HIGH	HIGH	HIGH	1L 6M 7E 5H	2L 3M 8E 6H
15	E	HIGH	HIGH	HIGH	2L 3M 5E 9H	13L 4M 2H
16	G	HIGH	ELEV	HIGH	4L 1M 8E 6H	3L 14M 2E
17	Q	HIGH	MOD	HIGH	2L 1M 14E 2H	11L 8M
18	A	HIGH	MOD	HIGH	1L 3M 11E 4H	17L 2M
19	A	HIGH	HIGH	HIGH	1L 5M 4E 9H	10L 4M 1E 4H
20	K	HIGH	LOW	HIGH	1L 1M 11E 6H	19L
21	E	HIGH	MOD	HIGH	2L 4M 5E 8H	13L 6M
22	F	HIGH	MOD	HIGH	3L 5M 9E 2H	18L 1M
23	I	HIGH	MOD	HIGH	7L 9M 2E 1H	17L 2M
24	A	HIGH	MOD	HIGH	4L 7M 5E 3H	18L 1M
25	W	HIGH	MOD	HIGH	1M 7E 11H	18L 1M
26	L	HIGH	MOD	HIGH	10L 6M 1E 2H	16L 3M
27	V	ELEV	LOW	ELEV	17L 1M 1E	19L
28	R	HIGH	LOW	HIGH	8L 7M 1E 3H	19L
29	G	HIGH	LOW	HIGH	13L 2M 3E 1H	19L
30	R	ELEV	LOW	ELEV	11L 5M 3E	19L
31	G	ELEV	LOW	ELEV	17L 1M 1E	19L

Note: CI Dist and CII Dist show the distribution of risk classifications across all 19 substitution variants at each position. For example, "9L 4M 4E 2H" indicates 9 variants classified LOW, 4 MODERATE, 4 ELEVATED, and 2 HIGH. The position-level classification (CI, CII, IRI columns) reflects the worst-case variant at each position.

3.3 Deletion Sensitivity Map

All 31 single-residue deletion variants were scored successfully with zero failures. Deletions are far more uniformly alarming than substitutions: 20 of 31 positions score HIGH on combined Mint_IRI. Class I shows 19 HIGH (versus 25 for substitution worst-case), while Class II shows only 2 HIGH and 3 ELEVATED, with 26 of 31 positions at LOW. The dual-class hotspot narrows to positions 11 through 15, where both classes independently show ELEVATED or HIGH.

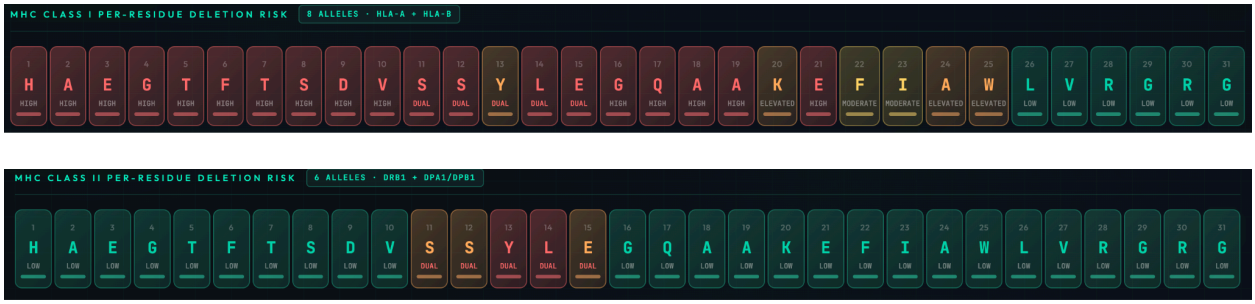

Figure: MHC Class I and Class II Per-Residue Deletion Risk Heatmap

Table 5. Complete Per-Position Deletion Risk Classification

Pos	AA	CI	CII	IRI (Max CI, CII)	Confidence	Key Flags
1	**H**	HIGH	LOW	HIGH	LOW	NSB, Terminal
2	**A**	HIGH	LOW	HIGH	LOW	NSB, Terminal
3	**E**	HIGH	LOW	HIGH	LOW	NSB, Terminal
4	**G**	HIGH	LOW	HIGH	LOW	NSB, Terminal
5	**T**	HIGH	LOW	HIGH	LOW	NSB, Terminal
6	**F**	HIGH	LOW	HIGH	LOW	NSB, Terminal
7	**T**	HIGH	LOW	HIGH	LOW	NSB, Terminal
8	**S**	HIGH	LOW	HIGH	LOW	NSB, Terminal
9	**D**	HIGH	LOW	HIGH	LOW	NSB, Terminal
10	**V**	HIGH	LOW	HIGH	LOW	NSB, Terminal
11	**S**	HIGH	ELEV	HIGH	LOW	NSB, Dual_Class, Terminal
12	**S**	HIGH	ELEV	HIGH	LOW	NSB, Dual_Class, Terminal
13	**Y**	ELEV	HIGH	HIGH	LOW	NSB, Dual_Class, Terminal
14	**L**	HIGH	HIGH	HIGH	LOW	NSB, Dual_Class, Terminal
15	**E**	HIGH	ELEV	HIGH	LOW	NSB, Dual_Class, Terminal

16	**G**	**HIGH**	**LOW**	**HIGH**	LOW	NSB
17	**Q**	**HIGH**	**LOW**	**HIGH**	LOW	NSB, Terminal
18	**A**	**HIGH**	**LOW**	**HIGH**	LOW	NSB, Terminal
19	**A**	**HIGH**	**LOW**	**HIGH**	LOW	NSB, Terminal
20	**K**	**ELEV**	**LOW**	**ELEV**	LOW	NSB, Terminal
21	**E**	**HIGH**	**LOW**	**HIGH**	LOW	NSB, Terminal
22	**F**	**MOD**	**LOW**	**MOD**	MOD	Terminal
23	**I**	**MOD**	**LOW**	**MOD**	MOD	Terminal
24	**A**	**ELEV**	**LOW**	**ELEV**	MOD	NSB, Terminal
25	**W**	**ELEV**	**LOW**	**ELEV**	MOD	NSB, Terminal
26	**L**	**LOW**	**LOW**	**LOW**	HIGH	Terminal
27	**V**	**LOW**	**LOW**	**LOW**	HIGH	Terminal
28	**R**	**LOW**	**LOW**	**LOW**	HIGH	Terminal
29	**G**	**LOW**	**LOW**	**LOW**	HIGH	Terminal
30	**R**	**LOW**	**LOW**	**LOW**	LOW	Terminal
31	**G**	**LOW**	**LOW**	**LOW**	LOW	Terminal

The C-terminal positions 26 through 31 are LOW on both classes, reflecting the reduced novel fragment generation near the terminus. Confidence is predominantly LOW for deletion variants, reflecting the MODERATE baseline (frame-shift uncertainty) further reduced by epitope density concordance factors. The near terminal flag fires at 30 of 31 positions (all except position 16), indicating that most deletion sites produce incomplete fragment coverage due to proximity to a terminus on a 31-residue peptide.

3.4 Insertion Sensitivity Map

All 31 self-insertion variants were scored successfully with zero failures. Insertions show a near-even four-tier spread on Class I: 5 LOW, 6 MODERATE, 6 ELEVATED, 14 HIGH. Class II is nearly silent: 29 of 31 positions are LOW, with only positions 9 (Asp) and 10 (Val) reaching MODERATE. No insertion variant reaches ELEVATED or HIGH on Class II.

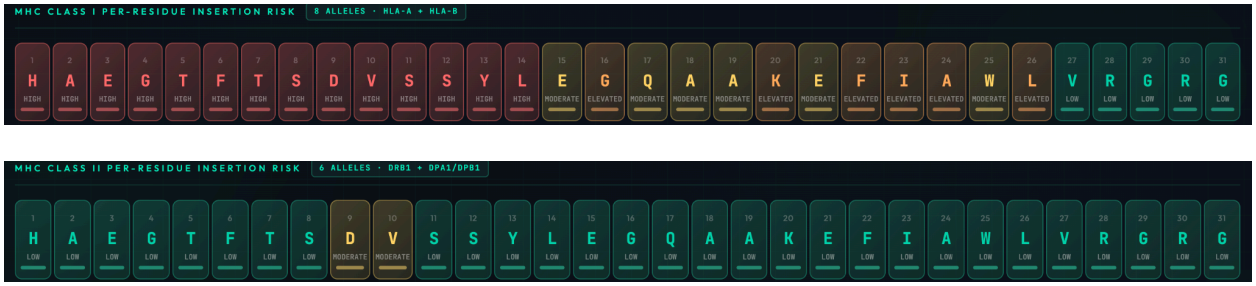

Figure: MHC Class I and Class II Per-Residue Insertion Risk Heatmap

Table 6. Complete Per-Position Insertion Risk Classification

Pos	AA	CI	CII	IRI	Conf	Key Flags
1	**H**	HIGH	LOW	HIGH	LOW	NSB, Terminal
2	**A**	HIGH	LOW	HIGH	LOW	NSB, Terminal
3	**E**	HIGH	LOW	HIGH	LOW	NSB, Terminal
4	**G**	HIGH	LOW	HIGH	LOW	NSB, Terminal
5	**T**	HIGH	LOW	HIGH	LOW	NSB, Terminal
6	**F**	HIGH	LOW	HIGH	LOW	NSB, Terminal
7	**T**	HIGH	LOW	HIGH	LOW	NSB, Terminal
8	**S**	HIGH	LOW	HIGH	LOW	NSB, Terminal
9	**D**	HIGH	MOD	HIGH	LOW	NSB, Terminal
10	**V**	HIGH	MOD	HIGH	LOW	NSB, Terminal
11	**S**	HIGH	LOW	HIGH	LOW	NSB, Terminal
12	**S**	HIGH	LOW	HIGH	LOW	NSB, Terminal
13	**Y**	HIGH	LOW	HIGH	LOW	NSB, Terminal
14	**L**	HIGH	LOW	HIGH	LOW	NSB, Terminal
15	**E**	MOD	LOW	MOD	MOD	Terminal
16	**G**	ELEV	LOW	ELEV	LOW	NSB
17	**Q**	MOD	LOW	MOD	MOD	Terminal
18	**A**	MOD	LOW	MOD	MOD	Terminal
19	**A**	MOD	LOW	MOD	MOD	Terminal

20	**K**	**ELEV**	**LOW**	**ELEV**	MOD	Terminal
21	**E**	**MOD**	**LOW**	**MOD**	MOD	Terminal
22	**F**	**ELEV**	**LOW**	**ELEV**	MOD	NSB, Terminal
23	**I**	**ELEV**	**LOW**	**ELEV**	MOD	NSB, Terminal
24	**A**	**ELEV**	**LOW**	**ELEV**	MOD	NSB, Terminal
25	**W**	**MOD**	**LOW**	**MOD**	HIGH	Terminal
26	**L**	**ELEV**	**LOW**	**ELEV**	MOD	NSB, Terminal
27	**V**	**LOW**	**LOW**	**LOW**	HIGH	Terminal
28	**R**	**LOW**	**LOW**	**LOW**	HIGH	Terminal
29	**G**	**LOW**	**LOW**	**LOW**	HIGH	Terminal
30	**R**	**LOW**	**LOW**	**LOW**	LOW	Terminal
31	**G**	**LOW**	**LOW**	**LOW**	LOW	Terminal

A clear N-to-C risk gradient is visible on Class I: positions 1 through 13 are all HIGH, transitioning through ELEVATED and MODERATE in the middle of the sequence, with positions 27 through 31 at LOW. Position 25 (Trp) reaches only MODERATE despite being mid-sequence, while adjacent positions 22 through 24 are ELEVATED. For semaglutide, the novel fragments generated by self-insertion variants do not produce strong Class II binding events, making insertion the lowest-risk impurity type from a Class II perspective for this particular peptide.

3.5 Cross-Impurity Comparison

Overlaying the three sensitivity maps reveals convergent and divergent risk patterns across the semaglutide sequence.

3.5.1 Cross-Impurity High-Sensitivity Region: Positions 11–15

Positions 11 through 15 represent the only region where all three impurity types converge on elevated Class I risk, and where substitution and deletion independently show elevated Class II signal. For substitution, all five positions are ELEVATED or HIGH on both Class I and Class II. For deletion, all five positions show Dual_Class_Risk flags, with position 14 (Leu) the only position reaching HIGH on both classes. For insertion, positions 11 through 14 are HIGH on Class I and position 15 is MODERATE, though insertion does not produce Class II signal at any position in the semaglutide sequence (29 of 31 positions LOW; Section 3.4). This hotspot corresponds to the SSYLE segment, and the dual-class convergence from substitution and deletion combined with uniform Class I signal from all three impurity types makes it the highest-priority region for manufacturing control across the full impurity landscape.

3.5.2 Class II Discriminating Power by Impurity Type

The Class II map provides the most useful risk differentiation for substitution variants (four-tier spread: 10/10/6/5). For deletion, Class II narrows to a tight hotspot at positions 11 through 15 with the rest at LOW (26/0/3/2). For insertion, Class II is essentially uninformative (29/2/0/0). This pattern reflects the underlying binding landscape: substitutions directly alter amino acid identity within the predicted binding fragments and produce the richest Class II differentiation; deletions generate novel frame-shifted fragments that produce concentrated Class II signal at positions 11 through 15; self-insertion variants, despite also generating frame-shifted novel fragments through the same scoring architecture, do not produce strong Class II binders for the semaglutide sequence.

The richer differentiation of substitution impurity type is partly architectural. The paired delta scoring used for substitution variants compares each impurity fragment directly against its parent counterpart. This inherently filters background binding noise: if both parent and impurity bind weakly, the delta is zero regardless. The result is higher-confidence classifications and finer-grained risk differentiation than the absolute threshold scoring required for frame-shifted deletion and insertion fragments. The near-complete Class II silence for insertion variants should therefore be interpreted with the caveat that moderate Class II binding signals that would be detected as threshold-crossing events in a paired comparison may fall below the absolute threshold cutoff for novel fragments. However, the magnitude of the Class II silence (29 of 31 positions at LOW) argues against a purely architectural explanation: if the absolute threshold were systematically masking moderate signals, a more gradual gradient from LOW to MODERATE would be expected rather than the near-uniform LOW observed.

A third contributing factor cannot be excluded: the current Class II panel omits HLA-DQ alleles (Section 5). If DQ-restricted presentation is relevant for frame-shifted insertion fragments in the semaglutide sequence, the current panel would be blind to these events. While DQ omission affects all three impurity types equally and therefore cannot fully explain why insertion Class II

silence is more pronounced than deletion Class II signal, DQ-restricted binding could contribute to an underestimate of Class II risk for insertion variants at specific positions. This possibility will be addressable when DQ alleles are incorporated into the panel.

3.5.3 C-Terminal Class I/Class II Divergence

Positions 26 through 31 are consistently LOW across all three impurity types on both MHC classes. Class I retains activity through position 25 (Trp, HIGH on substitution with 0 of 19 variants at LOW), while Class II is silent from position 20 onward for substitution and from position 16 onward for deletion and insertion.

This divergence is not a fragment coverage artifact. Within the receptor-binding core (positions 7 through 25), where Class II fragment coverage is uniform at or near maximum, fragment count shows no correlation with Class II hit rate (Spearman $\rho = -0.01$, $p = 0.92$). Fragment coverage explains 5.4% of the variance in Class II hit rate; sequence region explains 9.6%, and adding coverage to region adds effectively nothing ($\Delta R^2 = 0.02\%$). Position 20 (Lys) provides a direct proof point: 12 of 15 possible Class II fragments cover this position, all 19 substitutions were tested across 6 Class II alleles, and zero new Class II binders were detected. Class I classifies this position as HIGH.

The terminal 3 positions (29 through 31) fall within the Class II fragment-depleted zone, where the 15-mer sliding window generates progressively fewer fragments. At these positions, Class II silence cannot be cleanly attributed to biology versus reduced statistical power. The Terminal_Modification_LimitedFragments flag documents this reduced coverage. For manufacturing decisions, terminal impurities should not be cleared on the basis of a LOW Mint_IRI classification alone.

4. Discussion

4.1 Class II as the Primary Discriminating Map for Substitution

The Class I worst-case substitution map classifies 25 of 31 positions as HIGH, which at first glance may appear to offer limited differentiation. This outcome is an expected consequence of a deliberately conservative scoring architecture: the severity rule triggers HIGH on a single strong new binder event across any allele in the panel, and the worst-case aggregation reports the maximum risk observed across all 19 possible substitutions at each position. In practice, most positions have only a few substitutions that trigger HIGH out of 19 possible, meaning the specific substitution that occurs during manufacturing may not be the one that produces the worst-case classification. The exhaustive 19-variant scan is designed to identify the full immunogenic potential of each position rather than the expected risk under real-world process conditions. This conservative design is intentional: for a safety-oriented screening tool, the cost of missing a dangerous variant far exceeds the cost of flagging a position that, in practice, may only encounter benign substitutions. The Class II map, with its four-tier spread (10/10/6/5), provides the more operationally useful differentiation for prioritizing manufacturing controls, and is also the more clinically relevant pathway given that anti-drug antibody formation is driven by CD4+ T helper cell activation.

When a specific substitution impurity is identified analytically during batch release, it is scored individually through the same pipeline rather than reported as the worst-case classification for its position. A position classified as HIGH on the worst-case substitution map may contain a specific observed impurity that scores LOW or MODERATE, depending on which amino acid substitution actually occurred. The sensitivity map identifies positions requiring heightened analytical vigilance; the individual impurity assessment determines whether a specific detected impurity at that position warrants a further investigation.

4.2 Deletion: Uniformly Dangerous with a Narrow Dual-Class Hotspot

Deletions produce the highest proportion of HIGH-risk positions (20 of 31), reflecting the fundamental frame-shift mechanism: removing a residue creates entirely novel downstream peptide sequences with no parent counterpart. The dual-class hotspot narrows to positions 11–15, where Class II risk concentrates exclusively. Position 14 (Leu) is the only position reaching HIGH on both classes simultaneously, making it the single most critical deletion monitoring site in the semaglutide sequence.

4.3 Insertion: N-to-C Gradient and Class II Silence

Self-insertion variants produce a clear N-to-C risk gradient on Class I, reflecting the structural effect of frame-shift propagation length: N-terminal insertions disrupt more downstream fragments. The near-complete Class II silence (29 of 31 positions LOW) is the most distinctive finding across all three impurity types. Despite being scored through the same novel fragment assessment architecture as deletions, self-insertion variants in semaglutide do not produce novel fragments with strong Class II binding potential. This makes insertion impurities

substantially safer than deletions from a Class II perspective, though this finding is specific to the semaglutide sequence and should not be generalized to other peptides without independent analysis. This safety inference is further qualified by the HLA-DQ omission discussed in Section 3.5.2. The cross-peptide companion analysis confirms that C-terminal Class II silence is predominantly biological rather than methodological, with the caveat that terminal positions (within 7 residues of either end) have reduced Class II fragment coverage.

4.4 Biochemical Coherence of Substitution Patterns

Across all 589 substitution variants, the Class I results show clear biochemical stratification by substituting amino acid. Aromatic residues (Tyr 60% HIGH, Phe 52%) and bulky hydrophobics (Leu, Met, Ile in the 33–45% range) dominate the highest-risk outcomes, while small and polar side chains (Cys, Gly, Asn) never produce HIGH Class I risk. Class II shows a sparser signal overall, with no substituting amino acid exceeding approximately 11% HIGH rate. This biochemical stratification was not programmed into the pipeline; it emerges from the underlying NetMHCpan binding prediction models trained on experimental data. The reproduction of known MHC anchor residue biochemistry (Falk et al., 1991; Rammensee et al., 1995) provides methodological coherence.

4.4.1 Absence of Self-Proteome Filtering

The risk classifications presented in this white paper assess MHC binding potential without evaluating whether the predicted neo-epitopes resemble endogenous human peptides. Semaglutide is derived from human GLP-1, and a substantial fraction of its sequence is conserved with the endogenous hormone. Novel fragments generated by manufacturing impurities at some positions may be cross-reactive with peptides already presented to the immune system during thymic education, in which case the predicted MHC binding event would be met by existing immune tolerance rather than a naive T-cell response.

This omission means the sensitivity maps may overestimate immunogenic risk at positions where novel strong binders are tolerated through self-recognition. The Mattei et al. (2025) teriparatide data illustrates the magnitude of this effect: JanusMatrix tolerance scores separated the nine tested impurities into distinct risk tiers that correlated with in vitro T-cell response rates. Des-Leu28, which retained high self-proteome homology (JanusMatrix 4.88, above the API baseline of 4.74), produced the lowest donor response rate (24%), while impurities with disrupted tolerance (JanusMatrix below 3.0) produced response rates of 43%.

Self-proteome homology assessment via the PEPMatch algorithm is being integrated into the production pipeline. This layer will evaluate whether predicted neo-epitopes match endogenous human peptide sequences, enabling filtering of positions where novel strong binders may be tolerated. Until this layer is operational, the sensitivity maps should be interpreted as upper-bound hazard estimates that do not account for immune tolerance. Positions classified as HIGH that generate novel fragments with high human proteome homology may carry lower effective immunogenic risk than the classification alone suggests.

4.5 Retrospective Concordance with T-Cell-Confirmed Immunogenicity Data

The risk classifications presented in this white paper are generated by an algorithmic architecture that has been retrospectively evaluated against T-cell-confirmed immunogenicity outcomes from two independent, FDA-funded experimental datasets: Roberts et al. (2024) on salmon calcitonin (SCT) impurities and Mattei et al. (2025) on teriparatide (TPT) impurities. Both studies employed EpiVax's IVIP naïve donor T-cell assay to establish functional immunogenicity, providing the highest evidentiary standard against which computational screening can be benchmarked. This section summarizes the concordance between Mint_IRI classifications and T-cell outcomes across eight impurities with functional data.

4.5.1 Salmon Calcitonin: 3 of 4 T-Cell-Positive Impurities Detected

Four SCT impurities were evaluated in the IVIP T-cell assay by Roberts et al. (n = 16 donors). Donor response rates ranged from 56% to 69%, all numerically above the 44% API baseline, though the differences in response intensity (IFN-γ SFC counts) between API and individual impurities did not reach statistical significance by Mann-Whitney or Wilcoxon tests in the original study (n = 16). Under the original prediction engine (NetMHCpan 4.1/NetMHCIIpan 4.0 via IEDB API), Mint_IRI classified three of four at MODERATE or above: Q20E (ELEVATED), LYS-AC18 (HIGH), and ENDO-GLY28 (MODERATE).
Under the updated engine (NetMHCpan 4.2/NetMHCIIpan 4.3, locally installed), Q20E reclassified to LOW and LYS-AC18 reclassified from HIGH to ELEVATED. ENDO-GLY28 remained MODERATE.

The Q20E reclassification requires careful interpretation from two distinct perspectives.
From an MHC-binding assessment perspective, the LOW classification under the updated engine is algorithmically correct. Q20E is a deamidation (Q→E) at a TCR contact position that alters T-cell recognition without changing MHC binding. The binder-gated method assesses whether an impurity creates a new MHC binding event that the parent did not have. For Q20E, the parent fragment already binds MHC at the relevant alleles; the deamidation does not create a new non-binder-to-binder threshold crossing. The binder-gate correctly identifies that no new MHC binding event occurred. In this sense, the updated engine provides a more accurate MHC-level assessment than the original engine, which classified Q20E as ELEVATED.
From a patient safety perspective, however, Q20E is a T-cell-confirmed immunogenic impurity (56% donor response rate versus 44% API baseline in the Roberts et al. IVIP assay) that the updated pipeline classifies as LOW. The immunogenicity is driven by the altered TCR-facing surface of a peptide that was already MHC-bound, which is a mechanism that operates downstream of MHC binding and that no MHC-binding-based algorithm can detect. This represents a true scope limitation rather than a prediction error: the pipeline correctly assesses that Q20E does not alter MHC binding, but the pipeline's scope does not extend to TCR-mediated immunogenicity.

For concordance accounting, Q20E is therefore classified as a scope-limited miss (SLM) under the updated engine: an impurity whose immunogenicity operates through a mechanism outside the pipeline's defined detection scope. This is distinct from the ENDO-THR31 false negative, which may involve MHC-mediated events that the pipeline fails to detect due to geometric constraints. Both outcomes underscore that a LOW Mint_IRI classification does not constitute a determination of non-immunogenicity; it indicates that no new MHC binding event was detected within the scope of the analysis.

The single false negative, ENDO-THR31, scored LOW on both engine versions despite producing the highest donor response rate (69%). This impurity is a near-terminal threonine duplication (position 31 of 32) that generates insufficient novel fragments to trigger any risk threshold. Two non-exclusive explanations exist for this discordance: the immunogenicity may operate through non-MHC pathways (altered processing, aggregation, or conformational effects) that MHC-binding-based algorithms cannot detect, or the reduced fragment coverage at position 31 of 32 may limit sensitivity to MHC-mediated events at near-terminal positions. Both explanations are consistent with the Terminal_Modification_LimitedFragments flag, which alerts reviewing scientists to reduced coverage at near-terminal positions. The relative contribution of each mechanism cannot be determined from available data.

The reduced fragment generation at near-terminal positions constitutes a genuine analytical limitation for MHC-mediated events. Near-terminal positions are flagged in the production pipeline to alert reviewing scientists to this reduced coverage, ensuring that terminal impurities are not cleared on the basis of a LOW classification alone.

Notably, Mint_IRI correctly differentiated the two acetylation impurities by position under both engine versions: LYS-AC18 was classified above LYS-AC11, consistent with the biological observation that position 18 lies within the immunogenic EpiBar frame, while position 11 occupies the low-epitope N-terminal region.

4.5.2 Teriparatide: 4 of 4 Deletion Impurities Detected with Gradient Concordance

Four TPT single-residue deletion impurities at positions 11–14 were evaluated in the IVIP T-cell assay by Mattei et al. (n = 21 donors). All four produced elevated donor response rates (38–48%) relative to the Forteo® baseline (19%). Mint_IRI classified three as HIGH and one (Des-Lys13) as ELEVATED, achieving 4 of 4 detection at ELEVATED or above.

Beyond binary detection, Mint_IRI's severity classifications produced a gradient that is directionally concordant with the experimentally measured tolerance-disruption gradient reported by Mattei et al. The two impurities with complete loss of human proteome homology at the Treg epitope (Des-Leu11, JanusMatrix score 1.42; Des-Gly12, JanusMatrix score 1.19) both scored HIGH with strong novel Class I binder signals. Des-His14, with moderate tolerance reduction (JanusMatrix 3.75), also scored HIGH. Des-Lys13, with the mildest tolerance reduction (JanusMatrix 3.61, still above the 3.0 tolerance threshold), was the only impurity classified below HIGH. This gradient suggests that the MHC-binding perturbation detected by Mint_IRI and the tolerance disruption measured by JanusMatrix are structurally correlated outcomes of the same underlying deletion event. However, the concordance is observed across only four impurities at adjacent positions on a single peptide. The directional alignment is notable but does not constitute statistical evidence of correlation; confirmation across additional peptides and impurity types would be required to establish a generalizable relationship between Mint_IRI severity gradients and tolerance disruption.

This TPT concordance is directly relevant to the semaglutide deletion sensitivity map (Section 3.3). The deletion algorithm architecture producing the semaglutide risk classifications is identical to the one that achieved 4 of 4 detection on TPT deletions. The TPT deletions at positions 11–14 overlap the same mid-sequence positional range where the semaglutide deletion map identifies its dual-class hotspot (positions 11–15). While the underlying biology differs between the two peptides (tolerance disruption for TPT versus novel epitope generation for semaglutide), the structural principle is shared: mid-sequence deletions produce the highest

proportion of novel frame-shifted fragments and the greatest potential for MHC-binding landscape disruption.

4.5.3 Combined T-Cell Concordance Summary

Table 7. Mint_IRI classifications versus T-cell-confirmed immunogenicity outcomes across two FDA-funded validation datasets.

Data set	Impurity	T-cell Response	Mint_IRI (4.1/4.0)	Mint_IRI (4.2/4.3)	Outcome	Mechanistic Note
SCT	**Q20E**	56%	**ELEVATED**	**LOW**	**SLM (Scope Limited Miss)**	TCR-mediated; MHC binding unchanged. Binder-gated correctly identifies no new MHC binding event; immunogenicity driven by altered TCR recognition on existing MHC-bound peptide.
SCT	**LYS-AC 18**	56%	**HIGH**	**ELEVATED**	**TP (detected)**	3-allele Class I breadth (4.2); K→A proxy caveat applies
SCT	**ENDO-G LY28**	63%	**MODERATE**	**MODERATE**	**TP (detected)**	Unchanged across versions
SCT	**ENDO-T HR31**	69%	**LOW**	**LOW**	**FN (missed)**	Near-terminal; insufficient novel fragments. Unchanged across versions
TPT	**Des-Leu 11**	43%	**HIGH**	**HIGH**	**TP (detected)**	95% novel fragments; B*08:01 concentrated. Unchanged across versions
TPT	**Des-Gly 12**	43%	**HIGH**	**HIGH**	**TP (detected)**	5-allele breadth; highest concordance. Unchanged across versions
TPT	**Des-Lys 13**	38–48%	**ELEVATED**	**ELEVATED**	**TP (detected)**	Mildest tolerance disruption (JMX 3.61). Unchanged across versions
TPT	**Des-His 14**	38–48%	**HIGH**	**HIGH**	**TP (detected)**	Affinity concentration pattern. Unchanged across versions

Interpretive Note: These concordance results establish that the algorithmic architecture producing the semaglutide risk classifications has demonstrated directional concordance with T-cell-confirmed immunogenicity across two independent peptide systems. They do not constitute direct experimental validation of the semaglutide-specific predictions presented in this white paper.
All 8 concordance impurities were re-scored through the locally installed NetMHCpan 4.2 and NetMHCIIpan 4.3 to confirm classification stability across the prediction engine upgrade. All four TPT deletion impurities — the most directly relevant to the semaglutide sensitivity maps — retained identical risk across both engine versions. Among the SCT impurities, LYS-AC18 (HIGH → ELEVATED) reflects a single allele dropping below the binder-gate threshold; this impurity remains detected above MODERATE and the acetylation proxy caveat applies. ENDO-GLY28 remained MODERATE. Q20E (ELEVATED → LOW) is classified as a scope-limited miss under the updated engine: the binder-gated method correctly identifies no new MHC binding event, but the impurity's T-cell-confirmed immunogenicity operates through TCR-mediated mechanisms outside the pipeline's detection scope.
Under the original prediction engine (4.1/4.0), concordance was 7 of 8 (87.5%) at MODERATE or above. Under the updated engine (4.2/4.3), concordance is 6 of 8 (75%). Of the two misses, ENDO-THR31 is a geometric false negative at a near-terminal position, and Q20E is a scope-limited miss. Neither reflects a failure of the MHC binding prediction itself; both reflect defined boundaries of the analytical approach. From a patient safety standpoint, both impurities are immunogenic and both are classified LOW, which reinforces that a LOW classification should not be interpreted as a safety clearance without consideration of mechanisms outside the pipeline's scope. The sample size of 8 T-cell-tested impurities provides a basis for confidence in the structural logic of the risk stratification rather than a statistically powered sensitivity estimate.

4.6 Manufacturing Implications

The per-residue sensitivity maps establish a framework for risk-differentiated quality control in semaglutide manufacturing. Their primary function is to provide advanced intelligence: identifying the positions and impurity types where manufacturing excursions carry the greatest immunogenic potential, before any specific impurity is observed.

4.6.1 Position-Specific Risk Tiers

The convergence of all three impurity types at positions 11 through 15 identifies a universal high-sensitivity region where any detected impurity warrants investigation regardless of type. For process development, this region identifies the coupling cycles where efficiency has the most direct relationship to immunogenicity risk, and where interventions such as double-coupling protocols or extended reaction times are most justified from a patient safety perspective.

The broader deletion monitoring zone (positions 1 through 21, where all deletions produce HIGH or ELEVATED risk) and the substitution dual-class hotspot (positions 7 through 16 and 19) define regions where routine analytical methods should be most sensitive and where trending data should be reviewed with the greatest scrutiny. These risk tiers can also inform analytical method development by identifying the positions where detection sensitivity has the most direct relationship to patient immunological safety, enabling risk-proportionate allocation of analytical resources across the peptide sequence.

Positions 26 through 31 are consistently LOW across all three impurity types. Statistical analysis across the six-peptide GLP-1 family confirms that this C-terminal Class II silence is predominantly biological: within the full-coverage core, fragment count has zero correlation with Class II hit rate (Spearman $\rho = -0.01$, $p = 0.92$), and 72.5% of maximally covered positions across the family produce zero Class II binders under any substitution. However, the terminal 3 to 7 positions per peptide fall within the Class II fragment-depleted zone, and a LOW classification at these positions should not be interpreted as confirmed low immunogenicity risk. As discussed in Section 4.5.1, near-terminal impurities can be immunogenic through mechanisms the pipeline cannot detect.

Near-terminal positions are flagged in the production pipeline for both deletion and insertion variants to alert reviewing scientists to this reduced coverage. For manufacturing decisions, terminal impurities should not be cleared on the basis of a LOW Mint_IRI classification alone; orthogonal assessment methods remain necessary for impurities at these positions.

4.6.2 Toward Risk-Proportionate Acceptance Criteria

Current quality standards require that peptide impurities be controlled, but existing frameworks provide limited guidance for differentiating acceptance limits by immunogenicity risk. The per-residue sensitivity maps offer a quantitative basis for position-specific specification setting. An impurity at position 14 (dual-class HIGH across all three impurity types) presents a fundamentally different risk profile than the same impurity level at position 29 (LOW across all types). This data can support scientifically justified acceptance criteria that are proportional to the predicted immunogenic potential at each position, rather than uniform percentage thresholds applied identically across the sequence.

4.6.3 Concentration-Dependent Risk and the Threshold Question

The per-residue sensitivity maps classify immunogenic potential at the position level but do not, by themselves, define concentration thresholds at which a given impurity becomes clinically concerning. The most detailed regulatory framework for peptide impurity immunogenicity currently available is the FDA guidance for abbreviated new drug applications for synthetic peptide drug products, which establishes a 0.1% to 0.5% concentration threshold as the trigger for immunogenicity evaluation of impurities that differ between a generic product and its reference listed drug. While this threshold was developed for the ANDA pathway specifically, it represents the most concrete regulatory benchmark for the concentration at which peptide impurities warrant immunogenicity assessment, and is relevant as a reference point for any manufacturer of synthetic peptide therapeutics. However, this threshold functions as an administrative trigger for assessment rather than a biologically validated boundary between safe and unsafe concentrations.

The available experimental evidence suggests that impurity concentration modulates immunogenic response magnitude. Roberts et al. (2024) observed that salmon calcitonin impurities spiked into the reference drug product at their observed abundances (0.24% to 3.30% of the API) produced measurable increases in donor PBMC response rates relative to the drug product alone. Mattei et al. (2025) similarly reported that teriparatide impurity responses were attenuated at lower concentrations. These observations establish that concentration matters, but systematic concentration-response relationships mapped to specific sequence positions have not been published for any therapeutic peptide.

The sensitivity maps provide the prioritization framework that makes concentration studies tractable. Rather than requiring dose-response data for every possible impurity at every position, the maps identify the positions where concentration-dependent characterization has the greatest patient safety relevance. An impurity at position 14 (dual-class HIGH across all three impurity types) warrants more stringent concentration investigation than the same impurity type at position 29 (LOW across all types). This position-specific prioritization enables focused experimental designs: concentration-response studies for high-consequence positions, combined with the standard regulatory thresholds for low-consequence positions, would provide a risk-proportionate framework for specification setting that does not currently exist.

4.6.4 From Hazard Classification to Weighted Risk Assessment

The sensitivity maps presented in this white paper characterize immunogenic potential: the severity of the immunological outcome if a given impurity type occurs at a given position. This is the hazard component of a two-component risk equation. The complementary variable (position-specific impurity frequency under real manufacturing conditions) determines the exposure component. A complete risk assessment requires both.

Position-specific impurity frequency data is not available in the published literature for semaglutide or for most therapeutic peptides. The existing SPPS impurity literature characterizes impurity types and formation mechanisms but does not report position-resolved occurrence rates across manufacturing campaigns. Coupling efficiency is known to vary by amino acid identity, sequence context, resin loading, and solvent system, but these variables have not been systematically mapped to position-specific impurity frequencies for any GLP-1 analog at the level of resolution required for risk weighting.

This represents a field-level gap. Without position-resolved frequency data, it is not possible to distinguish positions where HIGH-risk impurities are generated at high frequency (true high-risk positions requiring the tightest control) from positions where HIGH-risk impurities are theoretically possible but rarely observed in practice (positions where the hazard classification overstates the operational risk). The sensitivity maps presented here provide the hazard framework that makes frequency-weighted risk assessment tractable: rather than requiring comprehensive frequency data at all 31 positions for all impurity types, the maps identify the specific positions where frequency data has the greatest impact on the risk calculation.

The finding that aromatic amino acids (Tyr, Phe, Trp) dominate HIGH-risk substitution outcomes (Section 4.4), combined with the identification of a dual-class hotspot where substitution risk is concentrated, illustrates this principle. If coupling efficiency data were available for each amino acid at each position in the semaglutide sequence, the positions where low coupling efficiency coincides with HIGH immunogenic consequence would emerge as the highest-priority targets for process optimization. The current maps identify the consequence side of this equation; the frequency side requires position-resolved impurity profiling from manufacturing campaigns.

As real-world impurity profile data is collected, position-specific impurity frequencies can be overlaid on the hazard maps to produce weighted risk scores reflecting actual process performance, enabling optimization efforts to be directed where they have the greatest impact on patient safety.

5. Known Limitations

MHC binding prediction does not capture TCR-facing modifications. Substitutions at T-cell receptor contact positions (positions 4, 5, and 6 in the 9-mer binding core) can alter T-cell recognition without changing MHC binding affinity. Such modifications are invisible to MHC-based prediction and represent a fundamental limitation of the approach.

The non-standard amino acid Aib at position 2 of semaglutide is modeled via a proxy Alanine substitution. While the delta cancellation principle ensures proxy error cancels in the parent-versus-impurity comparison, absolute binding predictions for fragments containing position 2 carry inherent approximation.

Class II fragment coverage is reduced at terminal positions due to the 15-mer sliding window. For semaglutide (31 residues), positions within 14 residues of either terminus are covered by fewer Class II fragments.

Substitution frequencies are assumed uniform across all 19 alternative amino acids. In reality, SPPS errors follow process-specific distributions. The exhaustive scan identifies the full immunogenic potential; real-world risk may be lower at positions where only rare substitutions produce high risk.

Insertion variants test only self-insertion (residue duplication). Insertions of non-native amino acids, which occur less frequently in SPPS but may arise from adjacent coupling contamination, would require additional simulation.

Deletion and insertion confidence baselines are MODERATE (lower than substitution's HIGH), reflecting the inherent uncertainty of novel fragment assessment without parent gating. The confidence level reported should be interpreted with this architectural difference in mind.

The risk classification reflects MHC binding potential only and does not account for upstream antigen processing steps that determine whether a predicted MHC binder is actually generated and presented in vivo. For Class I, this includes proteasomal cleavage and TAP transport. For Class II, this includes endosomal cathepsin processing. Prediction tools for some of these steps exist (e.g., NetChop for proteasomal cleavage) but were not incorporated because their predictive accuracy for short synthetic peptide fragments, as opposed to full-length protein antigens, has not been established. The gap between predicted MHC binders and naturally presented epitopes means the pipeline may flag fragments that are never generated by cellular processing, contributing to a conservative (over-calling) bias. T-cell repertoire diversity and immune tolerance mechanisms are also outside the current scope.

The current Class II panel omits HLA-DQ alleles. While NetMHCIIpan 4.3 supports DQ prediction and recent methodological advances have improved DQ prediction accuracy (Nilsson et al., 2023), the training data for DQ alleles remains sparser than for DR, and the contribution of DQ-restricted CD4+ T cell responses to peptide drug immunogenicity has not been systematically characterized. Including DQ alleles in the panel would increase Class II breadth counts and could influence risk classifications at positions near the MODERATE-ELEVATED boundary. Until DQ prediction accuracy and clinical relevance for peptide therapeutics are better

established, the panel design prioritizes the DR and DP loci where prediction confidence is highest. The pipeline architecture supports DQ expansion without modification, and DQ alleles will be incorporated when the evidence base supports confident interpretation of DQ-driven risk signals.

6. Future Directions

This methodology extends directly to other therapeutic peptides in the Mint_IRI peptide registry (47 registered peptides). As each peptide is profiled, cross-peptide patterns may emerge at conserved functional motifs.

Deamidation sensitivity maps (N→D and Q→E) can be generated for the subset of susceptible positions using the same pipeline infrastructure, completing the coverage of all major SPPS impurity classes.

As real-world impurity profile data is collected from CDMO manufacturing campaigns, position-specific impurity frequencies can be overlaid on the risk maps to produce weighted risk scores reflecting actual manufacturing conditions.

Self-proteome homology assessment via the PEPMatch algorithm is under development. This layer will evaluate whether predicted neo-epitopes resemble endogenous human peptides, enabling filtering of positions where novel strong binders may be tolerated through immune self-recognition.

The analysis covers sequence-altering impurities (substitution, deletion, insertion) but does not address chemical modification impurities such as oxidation, racemization, or incomplete side-chain deprotection. These modification types are among the most frequently observed SPPS impurities and may alter immunogenicity through mechanisms including changed MHC binding (for modifications at anchor positions), altered antigen processing (for modifications affecting protease cleavage sites), or changed aggregation propensity. The deamidation subset (Asn→Asp, Gln→Glu) is amenable to the existing substitution scoring framework and is being developed separately. For manufacturing quality decisions, the sequence-altering impurity risk maps should be interpreted alongside comprehensive impurity profiling that includes chemical modifications, rather than as a standalone assessment of the full impurity landscape.

References

1. Roberts BJ, Mattei AE, Howard KE, et al. Assessing the immunogenicity risk of salmon calcitonin peptide impurities using in silico and in vitro methods. Front. Pharmacol. 2024;15:1363139. PMID: 39185315.

2. Nilsson JB, Greenbaum J, Peters B, Nielsen M. NetMHCpan-4.2: improved prediction of CD8+ epitopes by use of transfer learning and structural features. Frontiers in Immunology. 2025;16:1616113.

3. Nilsson JB, Kaabinejadian S, Yari H, et al. Accurate prediction of HLA class II antigen presentation across all loci using tailored data acquisition and refined machine learning. Science Advances. 2023;9(47):eadj6367.

4. De Spiegeleer B, et al. Related impurities in peptide medicines. Journal of Pharmaceutical and Biomedical Analysis. 2014;101:2–30.

5. Gregg B, Swietlow A, Szajek AY, Rode H, Verlander M, Eggen I. Control strategies for synthetic therapeutic peptide APIs — Part III: Manufacturing process considerations. Pharmaceutical Technology. 2014;38(5).

6. Sette A, Sidney J. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and HLA-B polymorphism. Immunogenetics. 1999;50(3–4):201–212.

7. Greenbaum J, et al. Functional classification of class II human leukocyte antigen (HLA) molecules. Immunogenetics. 2011;63(6):325–335.

8. Lau JL, et al. Discovery of the once-weekly glucagon-like peptide-1 (GLP-1) analogue semaglutide. Journal of Medicinal Chemistry. 2015;58(18):7370–7380.

9. European Medicines Agency. Draft guideline on development and manufacture of synthetic peptide drug substances. EMA/CHMP/QWP/848945/2023.

10. Mattei AE, Roberts BJ, Lelias S, et al. (2025). Immunogenicity risk assessment of peptide-related impurities identified in generic teriparatide products. Front. Immunol. 16:1730346. doi: 10.3389/fimmu.2025.1730346.

11. De Groot AS, Roberts BJ, Mattei A, et al. (2023). Immunogenicity risk assessment of synthetic peptide drugs and their impurities. Drug Discov. Today 28:103714.

12. Falk K, Rötzschke O, Stevanović S, Jung G, Rammensee HG. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature. 1991;351(6324):290–296.

13. Rammensee HG, Friede T, Stevanović S. MHC ligands and peptide motifs: first listing. Immunogenetics. 1995;41(4):178–228.

14. Calis JJA, Maybeno M, Greenbaum JA, et al. Properties of MHC Class I presented peptides that enhance immunogenicity. PLoS Comput Biol. 2013;9(10):e1003266.

15. Semaglutide subcutaneous drug monograph. MedCentral. Available from: https://www.medcentral.com/drugs/monograph/174490-317067/semaglutide-subcutaneous. Accessed March 2026.