Integrated PharmaConsulting

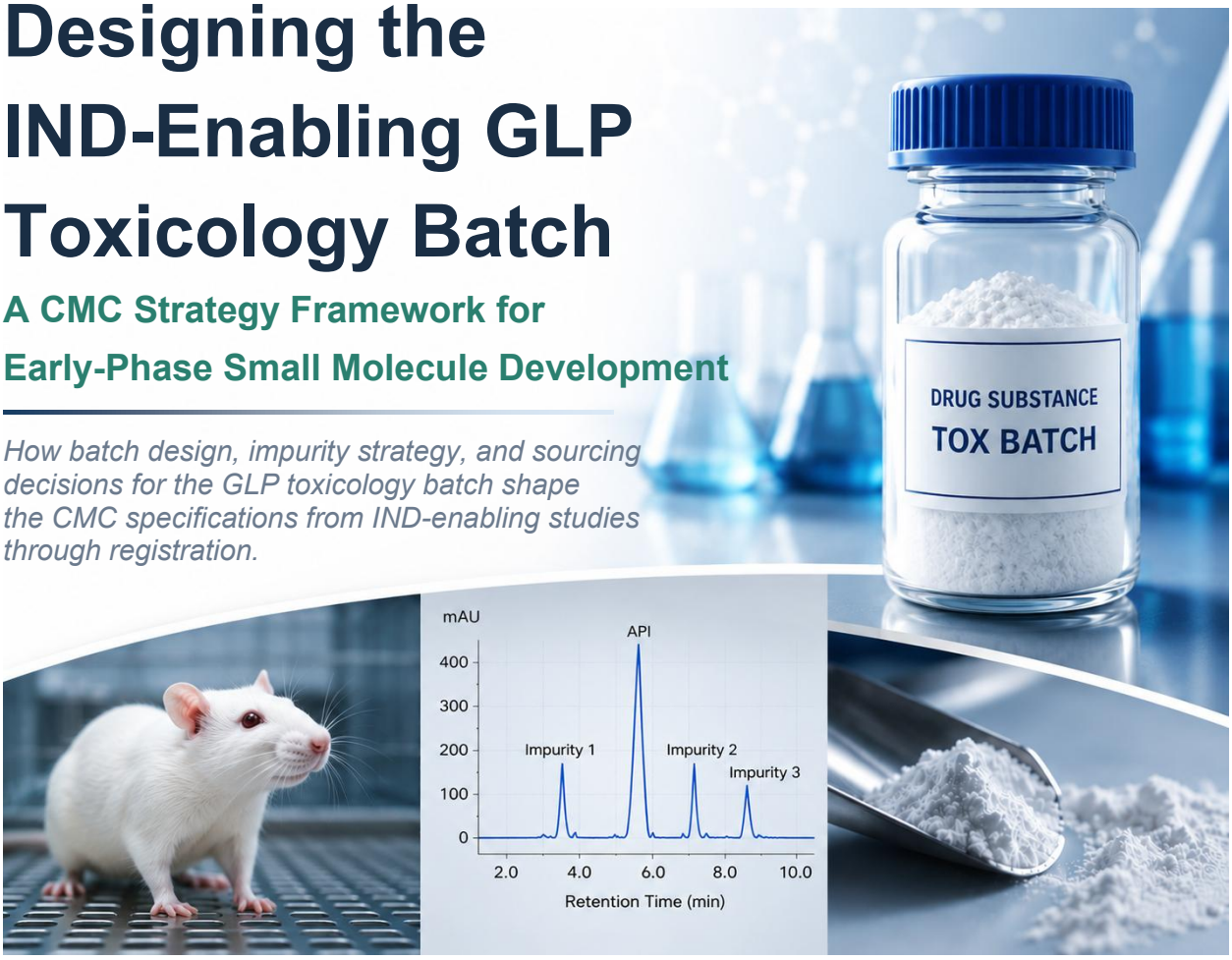
AUTHOR: Chad Hopkins, PhD

WHITE PAPER TOPIC: CMC & Toxicology | JULY 2026

Designing the IND-Enabling GLP Toxicology Batch

A CMC Strategy Framework for Early-Phase Small Molecule Development

How batch design, impurity strategy, and sourcing decisions for the GLP toxicology batch shape the CMC specifications from IND-enabling studies through registration.



Executive Summary

The batch of drug substance manufactured to support IND-enabling GLP toxicology studies is one of the highest-leverage decisions in a small molecule development program. Its impurity profile, solid form, and analytical characterization become the reference point against which subsequent GMP batches are measured. This white paper, prepared by Integrated Pharma Consulting, outlines a practical framework for designing, sourcing, and qualifying the GLP tox batch, and compares strategic approaches to early-phase impurity control, including how expectations differ between ICH M3 and ICH S9 (oncology) programs, and what the published regulatory record shows about how these strategies actually fare with global health authorities.

KEY INSIGHTS

- The tox batch sets the reference point for GMP specifications; specifications should tighten, not loosen, as development progresses.
- Analytical release methods should be stability-indicating and consistent with the methods ultimately used for DS/DP GMP release.
- A purpose-built “dirty” tox batch allows impurity qualification flexibility and keeps the tox study off the GMP critical path; using part of the first GMP batch as the tox batch removes that flexibility.
- ICH Q3A(R2) thresholds, while written for commercial-stage material, remain a useful aspirational target even for the first GMP batch.
- Regulatory acceptance of elevated early-phase impurity limits varies by country, and even within the same agency across programs. Plan for negotiation, not certainty, for any impurity not already qualified in GLP tox studies.

1. Why the GLP Tox Batch Matters

If a single batch punches above its weight in a small molecule development program, it is the one used for IND-enabling GLP toxicology studies to support first-in-human (FIH) trials. Its impurity profile, solid form, and potentially even formulation do not just support one study, they become the benchmark that every later process change, specification, and regulatory justification in early development gets compared against. Getting this batch right sets up a smooth path to FIH dosing; getting it wrong creates consequences that can persist for years of development. A guiding principle throughout is that specifications should tighten, not loosen, as development progresses: loosening a limit later signals to a regulator that the process was not well understood or controlled, and invites added scrutiny.

This is why batch design deserves the same level of strategic planning normally reserved for process development or formulation, and should not be treated as a routine manufacturing exercise. This paper covers organic, non-mutagenic impurities (NMIs) in drug substance (DS) and drug product (DP). Mutagenic (genotoxic) impurities are governed separately under ICH M7 and are outside this paper's scope. If your program is dealing with mutagenic or potentially mutagenic impurities, reach out to Integrated Pharma Consulting to discuss strategies.

2. Core Design Principles for the Tox Batch

2.1 Representative Final Form and Formulation

Wherever feasible, the tox batch should use the intended final solid form (neutral, salt, or co-crystal) and its relevant polymorph. For emulsions, polymer supports (e.g. SDDs or HMEs) or an intravenously administered drug, the actual intended formulation should be used. This is especially critical when a novel formulation or an excipient that is not Generally Recognized as Safe (GRAS) is involved, since the toxicology study then needs to qualify the safety profile of that excipient, not just the active pharmaceutical ingredient (API). For more formulation guidance, refer to Integrated Pharma Consulting’s Formulation white paper (www.integratedpharmaconsulting.com/resources).

2.2 Analytical Method Development

The analytical release method used for the tox batch matters nearly as much as the batch itself. It should be stability-indicating (able to detect and resolve degradation products, not merely assay the API) and as close as possible, ideally identical, to the method ultimately used for DS and DP GMP release. Switching methods later makes it harder to bridge impurity data back to the material that was actually tested in animals (refer to Section 2.4).

Although analytical method development is outside the scope of this white paper, there are a couple points worth mentioning. A robust method has to be capable of actually detecting what is there. For example, a single, symmetric peak on a UPLC trace is not proof of purity. A co-eluting impurity hidden beneath the main peak can be functionally invisible

unless the method is designed to expose it. Although the data in Figure 1 is hypothetical, it is based on an actual case the author worked on. Extracting the UV-Vis spectrum from the leading, middle, and trailing thirds of a peak and comparing their spectral shape and λ_{max} is a simple, effective check. Divergent spectra across the width of an apparently clean peak are a strong signal of impurity co-elution and indicates the need for further analytical development.

Wavelength selection deserves the same scrutiny. Monitoring at a wavelength on the slope of a UV-Vis spectrum, rather than at or near a λ_{max}, can under- or over-represent an impurity's relative response and distort the apparent impurity profile. Occasionally, multiple wavelengths, or even alternative detection methods, may need to be used for release testing if the λ_{max} of the impurities are significantly different or have poor, or nonexistent, absorbance. Figure 1 illustrates both failure modes using a simulated peak that looks perfectly clean in the chromatogram but resolves into three spectrally distinct species once the UV-Vis spectrum is examined across its width.

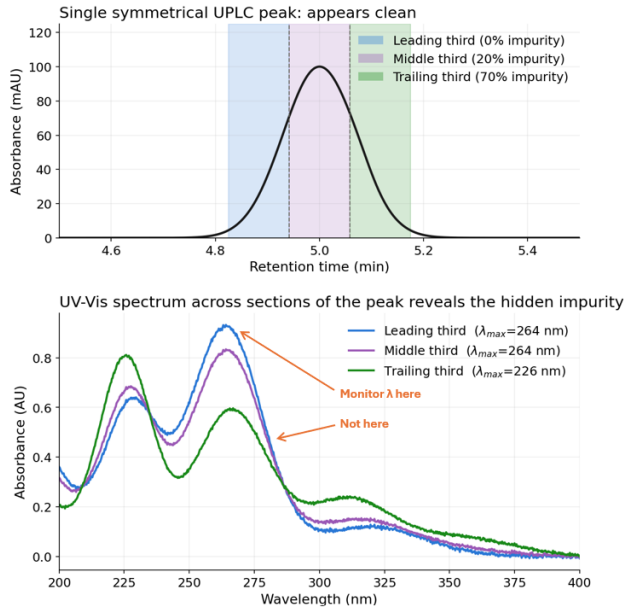


Figure 1. A single, symmetric UPLC peak can hide a co-eluting impurity. Extracting the UV-Vis spectrum from the leading, middle, and trailing thirds of the peak reveals three distinct spectra (λ_{max} of 264, 264, and 226 nm) despite the peak appearing pure at a single wavelength, underscoring the value of spectral purity checks and careful wavelength selection during method development.

2.3 Forced Degradation and Impurity Spiking

Running pilot forced degradation studies early is the best way to anticipate which degradation products are likely to appear later in development, ensuring the toxicology package and analytical methods are ready for them rather than playing catch-up. Spiking is an underused complementary tool: deliberately introducing known or predicted impurities into the tox batch at elevated levels is a practical way to qualify them with margin above expected GMP levels, without requiring a dedicated standalone qualification study. Abbreviated fate and purge studies also provide an additional tool for supporting inclusion or exclusion of potential impurities.

2.4 Sample Retention

A portion of the tox batch should always be retained and stored at low temperature (-20 °C or lower). Retained reference material is invaluable for analytical method development, impurity identification, or bridging a new process batch back to what was administered in the toxicology study.

2.5 Beyond Organic Impurities: The Broader Release Panel

Organic impurities are only one part of a tox batch release specification. A complete panel typically also includes appearance, assay, water content, residual solvents, elemental impurities, and residue on ignition (ROI), along with solid-form. Several of these are governed by frameworks separate from ICH Q3A such as ICH Q3C (residual solvents), ICH Q3D (elemental impurities), and compendial general chapters (e.g. ROI). Each carries its own limits and qualification logic and cannot simply inherit the organic-impurity strategy described in this paper.

A few recommendations are worth noting at this stage. Unidentified impurities are commonly tracked by relative retention time (RRT) until a structure is confirmed, which allows a consistent thread from the tox batch through later GMP batches before identification is complete. Seemingly routine attributes, such as appearance or water content, can carry useful information about form, hydration state, or degradation that becomes important when bridging batches later. No specification is necessarily required for these assays and may be listed as “report” or noted “for informational purposes only.”

Every program is unique. Each may have its own challenges and should be evaluated on a case-by-case basis.

3. Sourcing the Tox Batch: Purpose-Built vs. GMP-Derived

A common approach is to use a portion of the first GMP batch as the tox batch. While it all but guarantees, in the absence of meaningful

degradation, that the first GMP batch is within specification, it removes the flexibility to qualify higher impurity levels in a standalone “dirty” batch. The first GMP batch is not the place to intentionally push impurity levels higher. As such, we do not recommend this as the first option.

> **CAUTION**
> Using the first GMP batch as the tox batch also tends to stretch the development timeline, since the toxicology study must wait on the GMP batch instead of running in parallel with process development.

A purpose-built tox batch, intentionally pushed higher on select impurities through spiking or process variants, frees up the critical path and gives the program room to qualify a wider impurity range without putting GMP material at risk.

4. Purity Targets and ICH Q3A(R2) Benchmarks

As a working target for the tox batch, we look for 95-97% purity with no single impurity above 1%. These specs are generally suitable to support qualification arguments with margin to spare. Qualification here means establishing that an impurity level is justified by GLP tox studies. This range is deliberate: a batch that is too clean leaves few impurities to qualify, which caps the levels you can later justify in GMP material, the same rationale behind the purpose-built “dirty” batch in Section 3. At the same time, the batch should not be so impure as to risk severe adverse events in the tox study.

> **CAUTION**
> Do NOT use a highly pure tox batch (i.e. ≥99%). It will be hard to justify qualification levels in GMP batches beyond those set by ICH Q3A(R2).

The 95-97% purity and not more than (NMT) 1% single-impurity targets above are pragmatic internal targets informed by the ICH Q3A and Haber’s Law framework discussed below, not explicit regulatory requirements. ICH Q3A(R2) is written for commercial-stage DS, but its thresholds remain a useful target even for the first GMP batch. Aiming early GMP batches toward these thresholds, combined with the wider qualification window the tox batch provides, gives a program flexibility while the manufacturing process is still evolving.

Table 1. ICH Q3A(R2) reporting, identification, and qualification thresholds for non-mutagenic organic impurities

Maximum Daily Dose	Reporting Threshold	Identification Threshold	Qualification Threshold
≤ 2 g/day	0.05%	0.10% or 1.0 mg/day (whichever is lower)	0.15% or 1.0 mg/day (whichever is lower)
> 2 g/day	0.03%	0.05%	0.05%

5. Two Strategic Paths for Early Impurity Qualification

Beyond the batch itself, sponsors generally choose between two broad strategies for qualifying impurities in early clinical development. Each carries its own benefits and risks, and the right choice depends on the program, the indication, and the sponsor's appetite for risk and regulatory pushback.

5.1 Blanket Specification Approach

This approach sets a single blanket specification. For example, proposal of 95% or 97% purity with no single impurity above 1% regardless of what is actually observed. There are successful precedents for this approach, particularly in oncology. But, it carries real risk given how inconsistently global health authorities respond, as the regulatory case studies discussed in Section 7 illustrate. Pairing a higher purity floor with a lower single-impurity cap can hedge somewhat, but the risk of rejection by one or more agencies remains.

5.2 Iterative Qualification via Subsequent Tox Studies

The alternative is to defer qualification of new or higher-level impurities to subsequent sub-chronic and chronic toxicology studies, rather than attempting to qualify everything up front in the initial GLP tox batch. There will inevitably be situations where new or elevated impurities are observed over the course of process research & development (PRD) and qualification during later tox studies buys early-stage time and flexibility, at the cost of carrying open qualification questions later in development. Additionally, early Phase 1 human data can be incorporated into the qualification strategy based on actual human dose (e.g. compare human maximum dose to what was actually dosed in the tox studies) and exposures (e.g. metabolites).

Table 2. Blanket-specification vs iterative-qualification strategies

Strategy	Benefit	Risk
Blanket spec (e.g., 95 or 97% purity, ≤1% single impurity)	Simple and fast to implement; precedented in oncology programs	Inconsistent global acceptance; higher rejection risk absent impurity-specific justification
Deferred qualification via later tox studies	Preserves early-timeline flexibility; avoids impurity over-inclusion up front	Leaves open qualification questions later in development; risk of late-stage surprises

6. Population-Specific Considerations: Healthy Volunteers vs. Oncology

Expectations for the tox batch and its associated impurity strategy differ meaningfully depending on the patient population. For example, the risk-benefit for a healthy volunteer (HV) often used in early Phase 1 trials is not the same as for a patient with a life-threatening disease.

6.1 Healthy Volunteer FIH Programs

Non-oncology FIH programs in HVs fall under ICH M3(R2), and sponsors and agencies generally treat ICH Q3A/B thresholds as the reference point even though those guidelines are formally written for commercial-stage material. Because HVs receive no therapeutic benefit to offset impurity risk, impurity qualification tends to stay close to 1× ICH Q3A limits, sometimes with modest, justified headroom. Short single- and multiple-ascending-dose (SAD/MAD) study durations also keep the absolute daily intake of any impurity low, which helps keep programs within ICH Q3A limits without much need to stretch them.

6.2 Oncology and Other Life-Threatening Indications (ICH S9)

ICH S9 explicitly permits exceeding ICH Q3A/B limits when justified by factors such as disease severity, the parent compound's pharmacologic, genotoxic, and carcinogenic profile, treatment duration, and the impact of impurity control on manufacturing. In practice, sponsors have proposed limits of 3× to 7× ICH Q3A/B for early Phase 1 oncology material, on the basis that patients have a life-threatening disease with limited alternatives and that accelerated access is a priority. Abbreviated or accelerated nonclinical packages are also more common under S9, reinforcing the case for wider early impurity allowances.

Table 3. ICH M3 vs. ICH S9 (oncology) programs

Consideration	ICH M3	ICH S9
Typical impurity qualification	Close to 1× ICH Q3A/B, modest headroom	3×–7× ICH Q3A/B proposed early
Justification basis	Limited to no disease benefit to offset risk, especially for HVs	Disease severity, pharmacology, genotoxicity, treatment duration, manufacturing impact
Dose / duration	Short ascending-dose studies; low cumulative intake	Can involve longer or higher-dose regimens
Regulatory consistency	Generally consistent expectations	Highly variable by country and even by program

7. Regulatory Reality Check: Evidence from the Field

Two publications are especially useful for grounding early-phase impurity strategy in both toxicological rationale and real-world regulatory experience.

7.1 The Toxicological Case: Harvey et al. (2017)

Harvey and colleagues built the case for early-phase impurity limits, mining databases including Munro, RepDose, and the IMI eTOX database to support a 1 mg/day lifetime exposure threshold for non-mutagenic impurities as a negligible lifetime risk. They then applied a modified Haber's Law, using a 75-year lifetime exposure assumption and a 6-month early-phase exposure duration, to extend that threshold to 5 mg/day, or 0.7% (whichever is lower), for DS used in clinical trials under six months. For DP, they proposed 5 mg/day or 2%, whichever is lower, reflecting the additional quality-based margin already built into low-dose products.

7.2 The Regulatory Reality: Roberts et al. (2020)

Roberts and colleagues at Amgen published a case study of how proposed unspecified-impurity and

degradation-product limits actually fared across global health authorities for two Phase 1 oncology programs developed under ICH S9. The results show that a documented regulatory pathway for exceeding ICH Q3A/B limits does not translate into consistent acceptance.

Table 4. Global health authority responses to proposed early-phase impurity limits (adapted from Roberts et al., 2020)

Oncology Program	Proposed Limit	Outcome
Program 1: DS impurity	7× ICH Q3A	Rejected in both countries filed; 3× ICH ultimately accepted
Program 2: DS impurity	7× ICH Q3A	Accepted in 3 of 5 countries; the other 2 countered with 3× and 1× ICH
Program 1: DP degradant	3× ICH Q3B	Accepted in all 3 countries filed
Program 2: DP degradant	5× ICH Q3B	Accepted in 3 of 5 countries; the other 2 countered with 3× and 1× ICH

Notably, the same health authority accepted a 7× limit for one oncology program while rejecting the identical proposal for another, evidence that specification and control-strategy expectations are not always internally consistent, even within a single agency.

Together, these two papers are a useful reminder: the toxicological science can support a materially wider early-phase impurity window, but regulatory acceptance of that window is negotiated country by country, and program by program. Sponsors should work closely with their internal regulatory affairs teams and build the case-specific justification using disease context, dose, duration, and manufacturing rationale into every early-phase filing rather than relying on the multiplier alone.

8. Closing Summary

The GLP tox batch is not a routine manufacturing checkpoint. It is a strategic CMC decision with consequences that echo through the rest of development. Three principles are worth carrying forward:

1) Design the batch deliberately. Use the intended final form or formulation (where appropriate), align analytical methods with eventual GMP release methods, run forced degradation early, and consider spiking to qualify impurities with margin.
2) Source it independently of the first GMP batch. A purpose-built tox batch preserves qualification flexibility and keeps the toxicology study off the critical path.
3) Anchor the impurity strategy in both toxicology and regulatory precedent. ICH Q3A(R2) and Haber's-Law-based extrapolations provide the scientific rationale; published case studies show that global acceptance still requires program-specific justification and active regulatory dialogue.

Aligning batch design, impurity strategy, and regulatory engagement early gives a program the best chance of a smooth path from IND-enabling toxicology through FIH and beyond.

References

1. Harvey, J.; Fleetwood, A.; Ogilvie, R.; Teasdale, A.; Wilcox, P.; Spanhaak, S. Management of Organic Impurities in Small Molecule Medicinal Products: Deriving Safe Limits for Use in Early Development. *Regul. Toxicol. Pharmacol.* **2017**, *84*, 116–123.
2. Roberts, S. W.; Cauchon, N. S.; Ma, M.; Bezemer, J.; Zhang, X.; Weilage, E.; Anson, O. Control Strategy Expectations in Early Clinical Phase Synthetic Oncology Programs: Two Global Regulatory Case Studies. *Org. Process Res. Dev*. **2020**, *24*, 96–100.
3. ICH Harmonised Tripartite Guideline. Impurities in New Drug Substances Q3A(R2). International Conference on Harmonisation, 2006.
4. ICH Harmonised Tripartite Guideline. Nonclinical Evaluation for Anticancer Pharmaceuticals S9. International Conference on Harmonisation, 2009.
5. ICH Harmonised Tripartite Guideline. Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals M3(R2). International Conference on Harmonisation.
6. Gaylor, D. W. The Use of Haber's Law in Standard Setting and Risk Assessment. *Toxicology* **2000**, *149*, 17–19.

If your program is approaching development candidate nomination, entering GLP toxicology-enabling studies, or planning impurity qualification and control strategy for FIH, reach out to Integrated Pharma Consulting at info@integratedpharmaconsulting.com. We have experience designing GLP tox batch strategy and impurity control packages across small molecule programs, from healthy-volunteer FIH through accelerated oncology development.

Integrated PharmaConsulting
integratedpharmaconsulting.com

About Integrated Pharma Consulting, LLC

Founded to give early-stage biotech companies and academic spinouts access to senior preclinical & CMC expertise typically found only inside larger pharmaceutical organizations. With over 15 years leading preclinical R&D at Zentalis and Kalyra Pharmaceuticals, IPC provides strategic depth across the full translational continuum from discovery science to regulatory submission.