Organic Process Research & Development **2010,** *14,* 1032–1036

Risk Assessment of Potentially Genotoxic Impurities within the Framework of Quality by Design

Adam R. Looker,*,† Michael P. Ryan,† Bobbianna J. Neubert-Langille,† and Redouan Naji‡

Chemical Development and Analytical Development, Vertex Pharmaceuticals Incorporated, 130 Waverly Street, Cambridge, Massachusetts 02139, U.S.A.

Abstract:

A strategy for the risk assessment of potentially genotoxic impurities is described that utilizes Quality by Design in an effort to furnish greater process and analytical understanding, ultimately leading to a determination of impurity criticality. By identifying the risks and parameters that most influence those risks, an enhancement of both product and process control is attained that mitigates the potential impact of these impurities. This approach calls for the use of toxicological testing where necessary, chemical fate arguments when possible, multivariate analyses to develop design space, and use of spiking data to support specifications. Strong analytical support, especially with the development of low-level detection methods, is critical. We believe that this strategy not only aids in the development of a robust API process but also delivers on the identification and subsequent mitigation of risks to a class of impurities that are of high interest in the field.

Introduction

Genotoxic impurities are compounds that have established *in vitro* or *in vivo* capability to damage DNA, potentially leading to tumor development. Examples of compound classes that contain genotoxic impurities are alkyl halides,[1] various alkyl sulfonate esters,[1] and hydrazine.[2] In these cases, regulatory guidelines stipulate that these impurities need to be controlled to levels consistent with the Threshold of Toxicological Concern (TTC),[3] and possibly to even lower levels depending on the age of the patient population.[4] Beyond known genotoxins, drug development often uncovers impurities best classified as *potentially* genotoxic impurities (PGIs). These are identified during development with software packages that screen compounds for potential genotoxicity using structure−activity relationships (SAR).[5] Those compounds showing a positive relationship are referred to as "structural alerts" for genotoxicity.

Regulatory agencies have addressed how PGIs should be treated during drug development, ranging from specifying and controlling levels in the drug substance below the TTC, to performing toxicological tests on these impurities to determine whether they have established *in vitro* genotoxicity.[3,4] The question of whether the Quality by Design (QbD)[6,7] paradigm could be used to establish even greater process understanding (and therefore control) over these PGIs became apparent during the recent development of a drug candidate. PGIs were identified within the framework of a nitroaromatic reduction process. Herein we outline our use of QbD to risk assess PGIs, allowing greater process understanding and potential for regulatory flexibility.

Results and Discussion

Overall Risk Assessment Strategy. The general strategy for risk assessment is based on the current guidelines surrounding genotoxic impurities and PGIs.[3b,4] Figure 1 outlines the flowchart used to sequentially risk assess each impurity identified during development. The impurity is first evaluated using an applicable software package, such as DEREK (Deductive Estimation of Risk from Existing Knowledge), to determine which are classified as potentially genotoxic. Those impurities that do not trigger an alert in DEREK are treated as routine impurities in the synthetic process according to ICH Q3A.[8] Compounds that generate an alert are classified as a PGI, and are further categorized as either *observed* or *potential*.[9] Observed

* Author for correspondence. E-mail: adam_looker@vrtx.com.
† Chemical Development.
‡ Analytical Development.

(1) Sobol, E.; Engel, M. E.; Rubitski, E.; Ku, W. W.; Aubrecht, J.; Schiestl, R. H. *Mutat. Res.* **2007**, *633*, 80.
(2) Leakakos, T.; Shank, R. C. *Toxicol. Appl. Pharmacol.* **1994**, *126*, 295.
(3) (a) Muller, L.; Mauthe, R. J.; Riley, C. M.; Andino, M. M.; Antonis, D. D.; Beels, C.; DeGeorge, J.; De Knaep, A. G.; Fagerland, J. A.; Ellison, D.; Frank, R.; Fritschel, B.; Galloway, S.; Harpur, E.; Humfrey, C. D.; Jacks, A. S.; Jagota, N.; Mackinnon, J.; Mohan, G.; Ness, D. K.; O'Donovan, M. R.; Smith, M. D.; Vudathala, G.; Yotti, L. *Regul. Toxicol. Pharmacol.* **2006**, *44*, 198. (b) *Guideline on the Limits of Genotoxic Impurities*; EMEA Committee for Medicinal Products for Human Use, CPMP/SWP/5199: London, England, 2006. (c) Subpart B: Food Additive Safety, Section 170.39: Threshold of Regulation for Substances Used in Food-Contact Articles, Revised April 1, 2007. Code of Federal Regulations; Title 21, Food and Drugs; Office of the Federal Register, National Archives and Records Service, General Services Administration: Washington, DC, 2007; Chapter 1, Vol. 3. (d) Jacobsen-Kram, D.; McGovern, T. *Adv. Drug Delivery Rev.* **2007**, *59*, 38.
(4) *Guidance for Industry, Genotoxic and Carcinogenic Impurities in Drug Substances and Products: Recommended Approaches*, Draft; U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER): Rockville, MD, 2008.
(5) (a) Richard, A. M. *Mutat. Res.* **1998**, *400*, 493. (b) Wang, S.; Milne, G. W. *Chem. Res. Toxicol.* **1993**, *6*, 748.
(6) (a) ICH Q8. Pharmaceutical Development. In *International Conference on Harmonisation, Harmonised Tripartite Guideline*; ICH Expert Working Group: Geneva, Switzerland, 2006.). (b)(b) ICH Q9. Quality Risk Management. In *International Conference on Harmonisation, Harmonised Tripartite Guideline*; ICH Expert Working Group: Geneva, Switzerland, 2006.
(7) The approach described in this paper is likely similar to others industry wide. In designing this methodology, focus was directed at generating strong process understanding through the use of QbD whereby robust control points can be identified, thus ensuring patient safety.
(8) ICH Q3A(R2). Impurities in New Drug Substances. In *International Conference on Harmonisation, Harmonised Tripartite Guideline*; ICH Expert Working Group: Geneva, Switzerland, 2006.
(9) The following are definitions for observed and potential impurities. *Observed impurity*: impurity detected in isolated batches of starting material, intermediate, or API. *Potential impurity*: impurity seen during process, but not detected in starting material, intermediate, or API.

10.1021/op900338g
Published on Web 04/07/2010

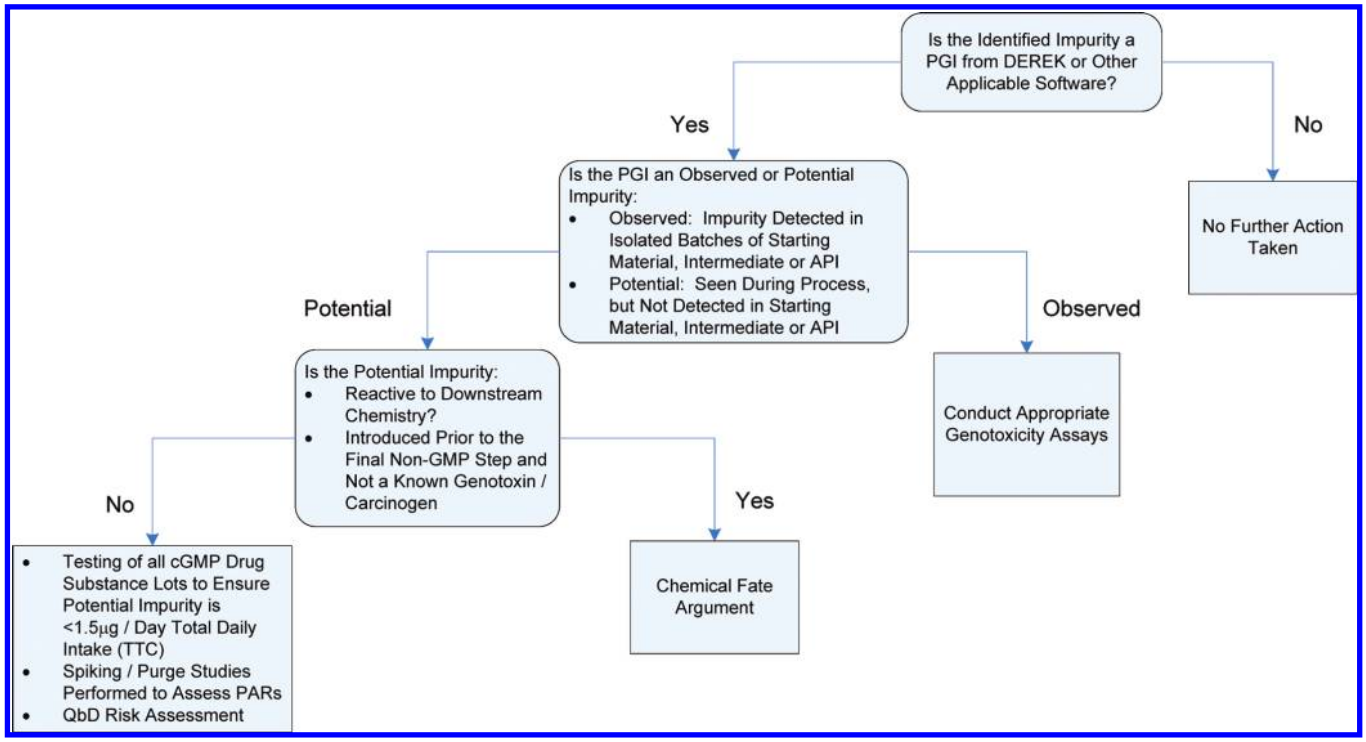


Figure 1. **Flowchart for the risk assessment of drug substance process impurities.**

impurities are evaluated using the proper toxicological testing due to their greater likelihood of being present in the drug substance. Impurities defined as potential are further risk assessed on the basis of their reactivity to downstream processing and according to how many purification operations (such as crystallization) are in the synthetic sequence that can realistically be used for purging prior to isolation of the final API. If the potential impurity is reactive to downstream processing, or introduced prior to the final non-GMP step, then the risk of being present in drug substance is determined to be low due to chemical fate. Otherwise, more intensive studies are warranted.[10]

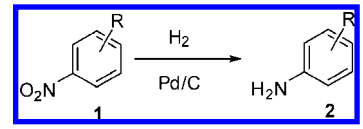
In supporting these studies, a suitable analytical method that can detect the PGIs at levels consistent with the TTC (dependent on predicted dosage, duration, and possibly the age of the patient population) is critical. Testing for PGIs at the TTC often calls for quantitation at ppm levels or lower. These methods allow for the collection of (a) spiking data, (b) design space experimental data, and (c) the determination of historical levels in all representative cGMP drug substance batches. Without these methods, it would be impossible to identify proven acceptable ranges (PARs) of the PGI.

Our approach uses these data to determine the criticality of the PGI, establish process controls (such as intermediate specifications or in-process testing criteria), and ultimately determine the API specifications. Having a PAR ≫ normal operating range (NOR) supports considering the impurity as noncritical.

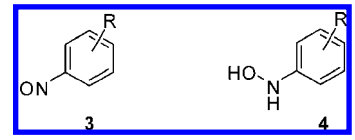
(10) Alternatively, one could perform the appropriate genotoxicity assays on these higher-risk PGIs prior to conducting further studies. If negative in these assays, no further work would be required. If positive, however, agencies would likely require their specification in the API. Under our approach, if the PGI is ultimately deemed "noncritical", specification in the API would not be necessary, thus avoiding longer-term low-level detection requirements during commercial production.

Example: Nitroaromatic Reduction. Scheme 1 outlines the general structure of nitroaromatic **1** and its reduction via Pd-catalyzed hydrogenation to aniline **2**. This process represents a synthetic step producing a key intermediate in an API synthesis. Two impurities that were identified during development were nitroso **3** and hydroxylamine **4** (Scheme 2). This was not entirely a surprise, as these are known intermediates in the hydrogenation process.[11]

Scheme 1. **Reduction of nitroaromatic 1 to aniline 2**



Scheme 2. **Potential process impurities**



All four compounds raised structural alerts using the DEREK for Windows program, version 10. Two of these (**1** and **2**) are deemed higher risk because they are intermediates in the synthetic sequence. As a result, and according to the flowchart in Figure 1, both of these impurities were classified as *observed* and thus underwent the proper toxicological assays, including the Ames test in the presence of S9 mix. Both were shown to be negative and could be treated as routine impurities. Nitroso **3** and hydroxylamine **4** were deemed lower risk of being present in drug substance due to their potential for control in the

(11) For discussions on the mechanistic pathway of nitroaromatic hydrogenation chemistry, please see: Augustine, R. L. *Heterogeneous Catalysis for the Synthetic Organic Chemist*; Marcel Dekker, Inc., 1996; p 473.

hydrogenation process. Although observed in process, both compounds were consistently not detected (<0.03%) by our assay methods in isolated batches of aniline **2**. Using the flowchart in Figure 1, impurities **3** and **4** were therefore defined as *potential* impurities. Since both were also introduced within the GMP portion of the synthesis, additional testing was recommended.

Quality by Design Development for Aniline 2 Process. In the case of the reduction of nitroaromatic **1**, early development work was able to identify a robust process that consistently produced aniline **2** meeting all required specifications. This process (generically) can be described as: (a) Hydrogenation of **1** in the presence of Pd/C and solvent; (b) removal of Pd via filtration; (c) concentration of the resultant filtrate to desired volume; (d) crystallization via antisolvent addition; and (e) filtration/drying of the isolated solid **2**. This early development work played an important role in helping to define the NOR within which the parameters can be controlled and still successfully produce **2**.

An intensive study was undertaken to explore the process design space and define the PAR's. Such information would enable the identification of parameters critical to product quality, and also to fine-tune process set-points. First, a risk assessment was conducted for all parameters in the master batch record using failure modes and effects analysis (FMEA). Risk was judged on the basis of each parameter's impact with respect to quality or for its potential to interact with another parameter (based on historical data and experience). Examples of process parameters that were assessed are shown in Table 1. Those parameters deemed to have low risk for impacting the process or for interacting with other process parameters were further investigated, as needed, via range finding (stressing) studies. Parameters with known or potential interactions as well as those with a higher risk of impacting the process were further studied using multivariate analysis (in this case design of experiments (DoE)).

Table 1. **Identified process parameters**

parameter	proposed action
hydrogen pressure[a]	range-finding studies
reaction volume	
distillation volume	
isolation temperature	
number, volume and temperature of washes	
temperature of drying	
vacuum pressure of drying	
amount of Pd (wt %)	multivariate analysis
reaction temperature	
reaction time	

[a] Pressure was evaluated linearly as a result of equipment limitations and the fact that this reaction requires pressures easily achieved and controlled at scale.

The assessment for the reduction of nitroaromatic **1** to aniline **2** indicated that catalyst loading, temperature, and time of hydrogenation were to be best studied using DoE. Our goal was to fully understand the limits of the hydrogenation process itself, especially with respect to nitroso **3** and hydroxylamine **4**, while establishing in-process controls that would lead to levels of these compounds below the TTC after downstream purge points. Ensuring reaction robustness at the IPC built in an extra level of control in advance of crystallization, further mitigating risk. In addition, using in-process purity profiles instead of evaluating isolated solids reduced design variables (isolation parameters and sources of variance) and greatly reduced the experimental burden of the design.[12] Spiking experiments were used to establish IPC levels for each compound that would result in none detected levels (ND, <0.03%) in the isolated solids under normal processing.

The DoE was generated using the custom design function of SAS's JMP software, version 7, with initial focus on temperature and catalyst. The design was later augmented to include reaction time. The resulting experimental plan allowed for the estimation of all main effects and interactions while also checking for curvature. All other parameters of the reactions were kept within process NORs. Figure 2 is a graphical representation of the experimental design, while Table 2 outlines the factors and response results. The responses chosen for the design were the IPC levels (% area) of aniline **2** nitroaromatic **1**, nitroso **3**, and hydroxylamine **4**.

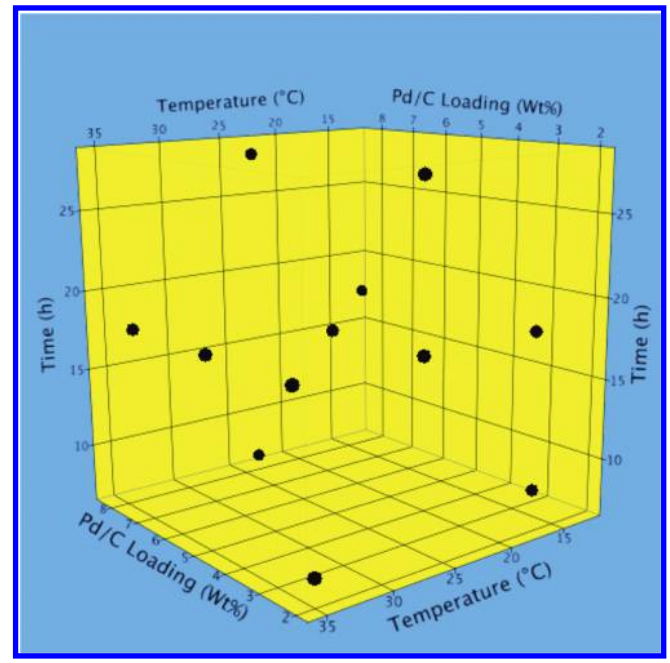


Figure 2. **Design space for the process step.**

The data for each of the responses were modeled using standard regression techniques to eliminate nonsignificant terms and interactions.[13] Factors with p values <0.05 were considered significant. The resulting prediction equations allowed the levels of each compound to be estimated for any given combination of time, catalyst loading, and temperature. These models demonstrated the existence of a time point at which an IPC specification for all compounds could be satisfied, regardless of the amount of palladium charged or the reaction temperature.

(12) The design spaces for the crystallization and isolation procedures were evaluated separately.

(13) For information on multiple regression analysis, please see: (a) Chatterjee, S.; Hadi, A. S. *Regression Analysis by Example*; John Wiley & Sons, Inc.: Hoboken, NJ, 2006. (b) Keith, T. Z. *Multiple Regression and Beyond*; Allyn & Bacon: Boston, 2005. (c) Sen, A.; Srivastava, M. *Regression Analysis: Theory, Methods, and Applications*; Springer-Verlag New York Inc.: New York, 1990.

Table 2. **Factors and response data**

exp	amt of Pd (wt % wet)	temp. (°C)	time (h)	IPC (% area) **2**	**1**	**4**	**3**
1	2.5	35	18	99.76	ND[a]	ND	ND
2	2.5	15	18	72.01	17.55	9.58	0.73
3	2.5	25	18	99.12	0.31	0.14	0.22
4	5	25	18	99.74	ND	ND	ND
5	5	35	18	99.70	ND	ND	ND
6	5	25	18	99.71	ND	ND	ND
7	7.5	15	18	99.66	ND	ND	ND
8	7.5	35	18	99.71	ND	ND	ND
9	7.5	25	8	99.71	ND	ND	ND
10	2.5	35	8	93.01	3.96	2.28	0.45
11	7.5	25	28	99.68	ND	ND	ND
12	2.5	15	8	50.99	42.12	5.65	1.16
13	2.5	25	28	99.77	ND	ND	ND

[a] ND: not detected (<0.03%).

Time was therefore judged to be the primary parameter controlling process robustness and quality, and the process was considered to have full design space in two dimensions (Pd amount and temperature, Figure 3). A stress test was performed at the edge of the design space using low Pd and low temperature to prove that these conditions will eventually meet each IPC. In fact, as Table 3 shows, that is the case. Aniline **2** of excellent quality was also isolated in this stress test to further validate the use of IPC data in the design.

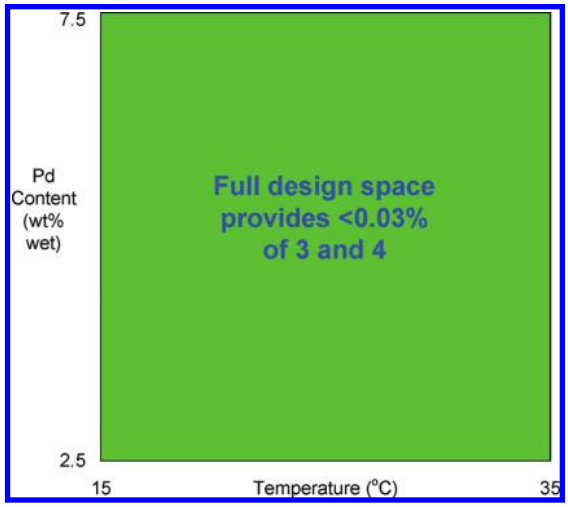


Figure 3. **Design space results.**

Table 3. **Stressing low Pd and temperature[a]**

time (h)	IPC (% area)	isolated purity of **2** (% area)
40	cmpd **2** - 99.76 cmpd **1** - ND cmpd **3** - ND cmpd **4** - ND	100.0

[a] Reaction used a Pd loading of 2.5 wt % (wet) at a temperature of 15 °C.

After identifying time as a controlling process parameter and after establishing prediction formulas for the IPC levels of nitroaromatic **1**, nitroso **3**, and hydroxylamine **4**, the next objective was to determine what IPC values ensured isolation of aniline **2** meeting specifications. As previously mentioned, spiking experiments were performed by charging both **3** and **4** into the hydrogenation process to determine what levels could be rejected (Table 4). On the basis of these spiking data, the proper IPC criteria were selected to ensure PGI levels would always be <0.03% in isolated **2**. The process is therefore deemed to be of very low risk in producing an amount exceeding 0.03% of either potential impurity in the isolated intermediate. The risk of whether these impurities could be present in the drug substance at levels above the TTC, though, still needed to be addressed.

Table 4. **Purge capability of aniline process**

impurity amount (% area)	isolated purity of aniline **2** (% area)
cmpd **3** - 0.34	100.0
cmpd **4** - 0.83	100.0

Spiking Experiments and cGMP Batch Testing. In order to further risk assess the levels of nitroso **3** and hydroxylamine **4** in the drug substance, a method to detect them down to low ppm levels was required. Several methods were explored, and LC−MS was found to be ideal for obtaining a limit of detection of <1 ppm for **3**. Unfortunately, this method was not viable for detecting **4**, as the hydroxylamine was unstable to the method conditions. After investigations involving derivitization and other techniques, normal-phase LC−UV was settled upon as the best method for detecting **4** to <1 ppm. Interestingly, this method actually converts hydroxylamine **4** to nitroso **3** during sample preparation, and what is detected is the nitroso **3** peak. Spike recovery of 90% was obtained for this method at 1 ppm.

With the capability to detect well below the TTC of our drug substance, all representative cGMP batches of drug substance (current route of synthesis) were tested to determine the historical levels of both PGIs. In fact, both were <1 ppm (not detected) in all batches.

In order to close out the QbD risk assessment, a final evaluation of the PAR for levels entering the downstream chemistry was needed. Both compounds were spiked into the next process step where aniline **2** is the starting material, carrying forward through drying of the drug substance (two chemical steps and a crystallization of the API). These drug substance batches were then tested for the presence of nitroso **3** and hydroxylamine **4** (Table 5). Relative to typical levels of

Table 5. **Spike/purge capability of downstream processing**

impurity amount spiked	subsequent levels in the drug substance(ppm)
cmpd **3** - 0.54% area	<1 (not detected)
cmpd **4** - 0.34% area	<1 (not detected)

3 and **4** seen in cGMP batches of **2** (<0.03%), substantially higher levels are cleanly purged to <1 ppm in the drug substance.

As a result of the data generated, significant process understanding and control have been established with respect to PGIs **3** and **4**. Using time as a controlling parameter in the hydrogenation process, IPCs were defined that ensure both **3** and **4** are isolated in low levels in intermediate **2**, and these levels have been shown via spiking experiments to lead to amounts well within the TTC limits for the API. On the basis of this information, nitroso **3** and hydroxylamine **4** were assessed as noncritical attributes.

Conclusion

In conclusion, a QbD risk-based assessment of PGIs is presented that establishes process control and analytical control leading to a determination of criticality. In this particular example, four PGIs representing all facets of a nitroaromatic hydrogenation process (starting material, product, and pathway intermediates/byproduct) were risk assessed, and ultimately deemed to be noncritical.

Acknowledgment

Special thanks to Michael Shannon, Amy Zhang, Marjorie Egan, Michael Carver, Brian Phenix, and John French for their contributions and discussions related to the work described.

Received for review December 25, 2009.

OP900338G