Organic Process Research & Development **2008,** *12,* 632–636

Identification of Critical Process Impurities and Their Impact on Process Research and Development

Yande Huang,*,† Qingmei Ye,† Zhenrong Guo,‡ Venkatapuram A. Palaniswamy,† and John A. Grosso†

Analytical Research and Development and Process Research and Development, Bristol-Myers Squibb Company, New Brunswick, New Jersey 08903-0191, U.S.A.

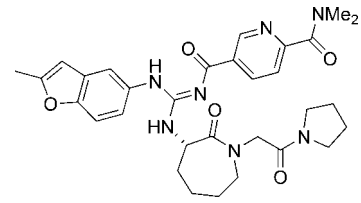
Abstract:

The identification of low-level critical process impurities and degradants encountered during pharmaceutical development is crucial to the process development, but can often be challenging and can negatively impact the timeline of the developmental program. This is demonstrated during the early stage of process research and development of a Factor Xa inhibitor, the caprolactam 1. Details focusing on rapid identification of impurities in the active pharmaceutical ingredient (API), recognition of their root causes of formation, and the impact on process development are described.

Introduction

The goal of pharmaceutical development is to develop process understanding and control which will yield procedures that consistently deliver products possessing the desired key quality attributes. To achieve this, the quality by design (QbD) paradigm has been employed in combination with process-risk assessment strategies to systematically gather knowledge through the application of sound scientific approaches.[1] Ganzer et al. recently published an article about critical process parameters and API synthesis.[2] The article presented an in-depth discussion of a stepwise, process risk assessment approach to facilitate the identification and understanding of critical quality attributes, process parameters, and in-process controls. The primary benefit of working within the QbD conceptual framework and employing process risk assessment strategies is the reproducible delivery of high-quality active pharmaceutical ingredient (API). However, a secondary benefit is the ability to obtain regulatory flexibility with respect to filing requirements.[3]

The control of impurities observed in an API is critical in delivering an API of high quality. Identification and understanding of the mechanism of formation of process-related impurities are critical pieces of information required for the development of control strategies. In addition, to ensure a continuing supply of API for drug product clinical manufacture, timely identification of key impurities is essential. These synthesis-related impurities and their precursors are considered as critical impurities because they directly affect the quality and impurity profile of the API. It is our practice that critical impurities be identified if practicable. Therefore, the timely identification of critical impurities becomes an integral part of process development.



1

Figure 1. **Structure of caprolactam 1.**

There are different approaches to the identification of impurities. We describe, herein, a general strategy that we have used in our laboratory, which leads to the rapid identification of impurities. To identify the structure of a low-level unknown impurity, we usually use liquid chromatography/mass spectrometry (LC/MS)/high-resolution MS (HRMS) and tandem MS (MS/MS) for molecular weight (MW) determination, elemental composition, and fragmentation patterns. On the basis of the mass spectrometric data and knowledge of the process chemistry, one or more possible structure(s) may be assigned for the impurity, with definitive structure information obtained by inspection of the HPLC retention time, UV spectrum, and MS profile of an authentic compound. If an authentic sample is not available, the isolation of a pure sample of the impurity is undertaken for structure elucidation using NMR spectroscopy. The isolation of low-level impurities is usually conducted using preparative HPLC chromatography. This approach was utilized for the identification of three unknown impurities present in an API batch of a Factor Xa inhibitor clinical candidate, the caprolactam **1** (Figure 1).

Caprolactam **1** was in the early stage of development as a Factor Xa inhibitor for the prevention or treatment of thrombosis, coronary artery disease, or cerebrovascular disease.[4] The synthetic route[5] depicted in Scheme 1 was used to prepare a

* Author to whom correspondence should be sent. E-mail: yande.huang@bms.com. Telephone: 732-227-7405. Fax: 732-227-3934.

† Analytical Research and Development.

‡ Process Research and Development.

(1) *International Conference on Harmonisation (ICH) Guidelines*; Q8, Pharmaceutical Development, 2005; Q9, Quality Risk Management, 2006.

(2) Ganzer, W. R.; Materna, J. A.; Mitchell, M. B.; Wall, L. K. *Pharm. Technol.* **2005**, (July 2), 1–12.

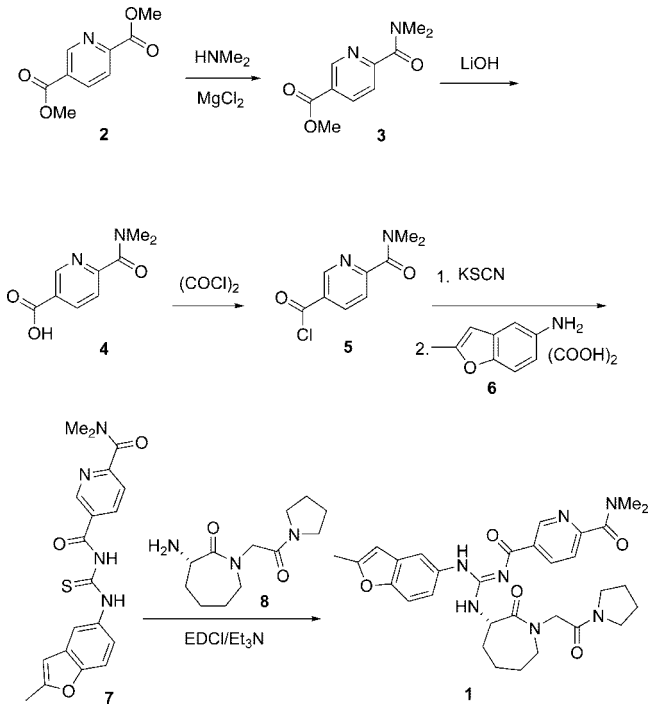
(3) Nasr, M. Drug Information Association Annual Meeting, Philadelphia, PA, June 19, 2006; *Pharmaceutical Quality Assessment System (PQAS) in the 21st Century*, 2006.

(4) Stein, P. D.; Bisacchi, G. S.; Shi, Y.; O'Connor, S. P.; Li, C. PCT Int. Appl. WO 2000047207, 2000; Application: A1 20000817; WO 2000-US2883, 2002.

(5) Guo, Z.; Li, W.-S.; Dowdy, E.; Polniaszek, R.; Kant, J.; Miller, M. 224th ACS National Meeting, Boston, MA, United States, August 18–22, 2002; Improved preparation of 5-{*N*′-(2-methyl-benzofuran-5-yl)-*N*″-[2-oxo-1-(2-oxo-2-pyrrolidin-1-yl-ethyl)-azepan-3-yl]-guanidinocarbonyl}-pyridine-2-carboxylic acid dimethylamide: A potent and selective Factor Xa Inhibitor *Abstracts of Papers*, 2002; ORGN-735.

10.1021/op800067v CCC: $40.75
Published on Web 06/21/2008

Scheme 1. **Synthetic route of caprolactam 1**



batch of the API for toxicology studies. In the HPLC profile of this batch, three unknown impurities were detected at significant levels (0.12−0.56 area percent (AP)). An expanded analytical HPLC profile is shown in Figure 2. The levels of these impurities exceed the 0.1% threshold of identification specified in ICH guidelines and hence were isolated and identified.[6] Ultimately, impurity structure identification and the determination of the mechanism of formation enhanced our understanding of the synthetic processes and aided our optimization of these processes to attain a controllable impurity profile and an API with the requisite key quality attributes.

Results and Discussion

The impurity identification strategy described above was implemented for API **1**. The LC/MS analysis of the API batch revealed a MW of 1134 Da for impurities A and B and 366 Da for impurity C. The high MW of impurities A and B suggests that they may be pseudodimers of **1** (2 M = 1174 Da). However, the 40 Da difference in mass between impurity A (or B) and the dimer of API did not readily suggest any possible structures. Impurity C is likely a substituted urea, an analogue of the substituted thiourea (penultimate **7**, Scheme 1) based on the MS data and process chemistry knowledge. As an authentic sample of the substituted urea was not available for structure confirmation, the isolation of impurities A, B, and C by preparative HPLC was undertaken.

Isolation. From the inspection of the analytical HPLC profile of **1** shown in Figure 2, it was apparent that this method would not be suitable to scale up for the isolation of impurities A and B because of their close elution with **1**. A new HPLC method with a shorter run time was developed, as described in the Experimental Section. Under the new HPLC conditions, impurities A and C and a few impurities at lower levels were coeluting at 9.3−10.1 min (collected as fraction-1). Impurity B eluted at 10.1−10.8 min (collected as fraction-2), while **1** eluted at 11.0−12.4 min. The elution order of **1**, impurity B, and impurity C in the preparative HPLC method were reversed in comparison with the analytical HPLC method, which is preferred for isolation. Impurity B was isolated as an off-white powder with a purity of 97 AP. Fraction-1 was concentrated under reduced pressure and placed in a refrigerator. Impurity C precipitated as a white solid from concentrated Fraction-1 and was isolated by filtration (98 AP). Impurity A was obtained at 90 AP by further purification of the filtrate of concentrated fraction-1 using a shallower linear gradient HPLC method. The isolated impurities A, B, and C were further analyzed using the analytical HPLC method to ensure they had not undergone any structural transformation during isolation, as shown in Figure 2.

Structure Elucidation of Impurity A. The molecular formula of impurity A is established from the isolated sample as $C_{59}H_{70}N_{14}O_{10}$ using exact mass of $[M + H]^+$ (m/z = 1135.5455) determined by positive ESI HRMS (calcd for $C_{59}H_{71}N_{14}O_{10}$: 1135.5477). Comparison of the NMR data of impurity A with those of **1** indicates that impurity A contains two sets of ^1H and ^{13}C resonances corresponding to two arginine core structures with the same caprolactam and pyridinedicarboxylate moieties as those in **1**. The four aromatic protons (at δ 6.33, 7.04, 7.36, and 7.28) corresponding to the benzofuran moiety that are the same as those in **1** are also observed in impurity A. However, the methyl resonance observed in **1** is not detected in impurity A. Instead, a methylene group (at δ 4.00) is observed, indicating that impurity A is a derivative of **1** with the modification at the methyl group. In addition, three additional aromatic protons (at δ 6.75, 6.94, and 7.01) and six additional aromatic carbons (at δ 116.6, 125.0, 125.5, 127.1, 128.2, and 153.9) are observed in impurity A, which are assigned to a 4-substituted amino-2-substitituted methyl phenol moiety with the aid of two-dimensional (2D) heteronuclear multiple bond correlation (HMBC) NMR data and the elemental composition established by HRMS. Therefore, the structure of impurity A is assigned as a derivative of **1**, as shown in Figure 3, which is consistent with 1D ^1H and ^{13}C and 2D COSY, heteronuclear multiple quantum correlation (HMQC), and HMBC NMR data.[7]

Structure Elucidation of Impurity B. The molecular formula of impurity B is established from the isolated sample as $C_{59}H_{70}N_{14}O_{10}$ using exact mass of $[M + H]^+$ (m/z = 1135.5431) determined by positive ESI HRMS (calcd for $C_{59}H_{71}N_{14}O_{10}$: 1135.5477). Comparison of the NMR data of impurity B with those of impurity A indicates that they are very similar except for the connectivity between the phenol and methylbenzofuran moieties. There is a methyl resonance (at δ 2.38) observed in impurity B, while the furanyl proton observed in impurity A (at δ 6.33) is absent in impurity B, indicating that in impurity B the phenol ring is connected to the furan ring, as shown in Figure 3. The structure of impurity B is

(6) *International Conference on Harmonisation (ICH) Guidelines*; Q3A(R), Impurities in New Drug Substances (Revised Guidelines), 2003.

(7) Complete ^1H and ^{13}C NMR chemical shifts assignments of impurity A are shown in Tables 1 and 2 in the Supporting Information.

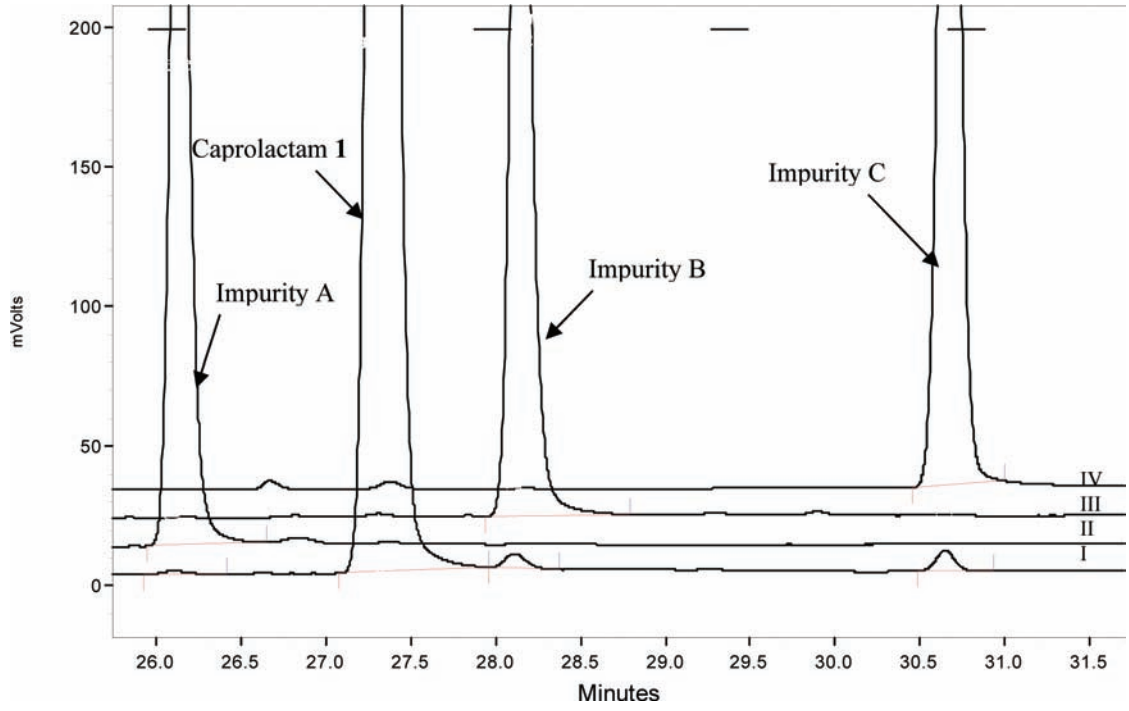


Figure 2. **Expanded analytical HPLC chromatograms of caprolactam 1 which contains impurity A, caprolactam 1, impurity B, and impurity C (I); isolated impurity A (II); isolated impurity B (III), and impurity C (IV).**

consistent with 1D ^{1}H and ^{13}C and 2D COSY, HMQC, and HMBC NMR data.[8]

Structure Elucidation of Impurity C. The molecular formula of impurity C is established from the isolated sample as $C_{19}H_{18}N_4O_4$ using exact mass of $[M + H]^+$ ($m/z = 367.1421$) determined by positive ESI HRMS (calcd for $C_{19}H_{19}N_4O_4$: 367.1406). There are two sets of proton−proton correlations observed in the COSY NMR experiment. One is attributed to three protons at δ 8.61, 7.99 and 9.48. They are assigned to the pyridine ring. The other three proton spin systems at δ 7.63, 7.51, and 8.19 are assigned to the phenyl ring of benzofuran moiety. Three methyl groups are observed as singlets at δ 2.27, 2.92, and 3.03, respectively. The resonance at δ 6.44 as singlet is assigned to the furan proton. The urea carbonyl carbon is observed at δ 151.8. The structure of impurity C is shown in Figure 3. The 1D and 2D NMR data are consistent with this structure.[9]

Investigation of the Root Causes of the Formation of Impurities A, B, and C. From the synthetic sequence shown in Scheme 1, it is highly likely that impurity A originated from impurity D present in the starting material **6**, as shown in Scheme 2. Impurity D contains two aniline amines which led to the formation of impurity E during the synthesis of penultimate **7**. Impurity E could not be purged from **7**, but reacted with **8** to form impurity A in the synthesis of **1**. Similarly, impurity B was likely introduced from impurity F which was also present in the starting material **6** and via impurity G during the synthesis of penultimate **7**, as illustrated in Scheme 3. Impurity C was likely formed during the synthesis of penultimate **7** and was carried over to the final step.

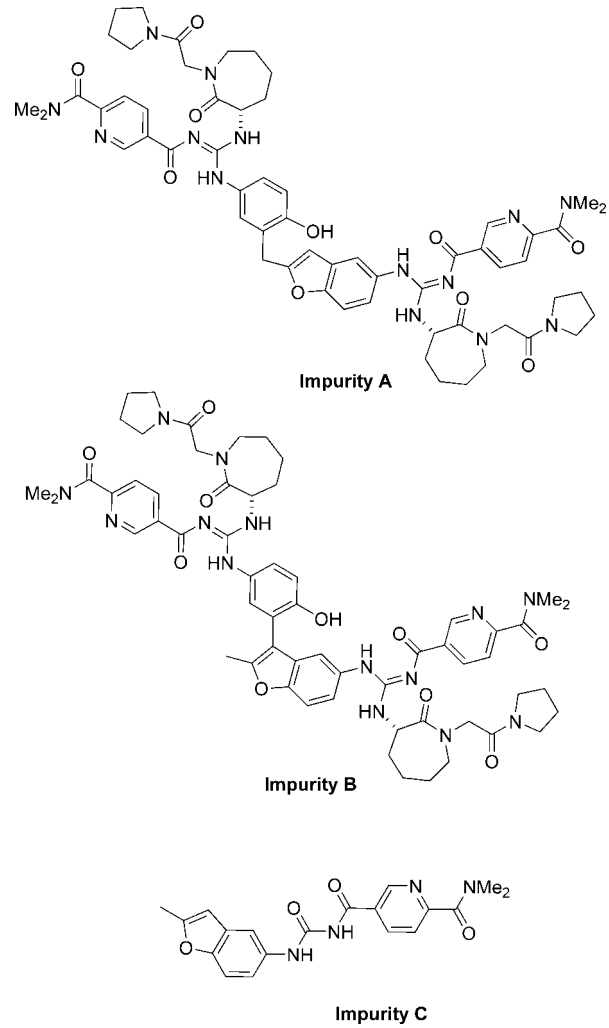
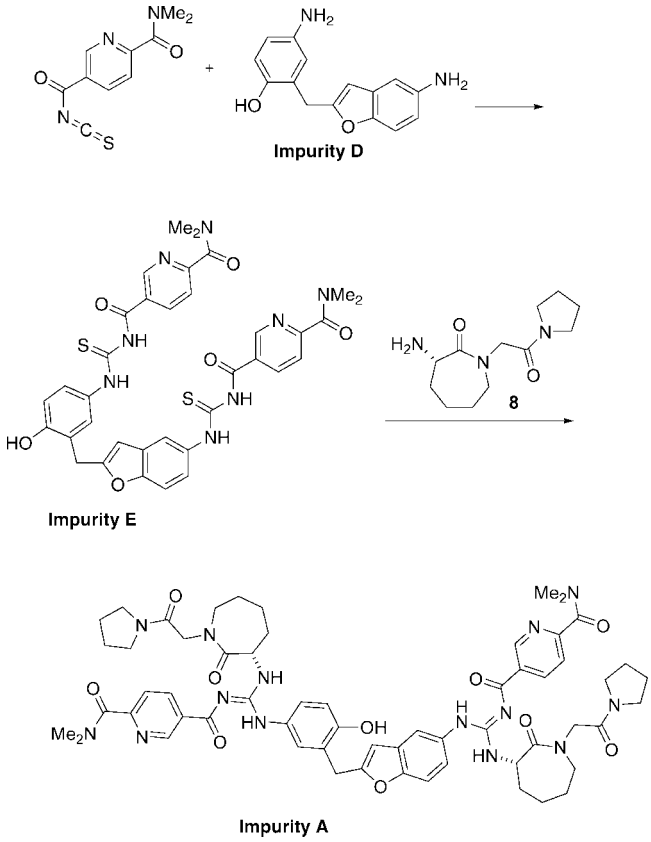


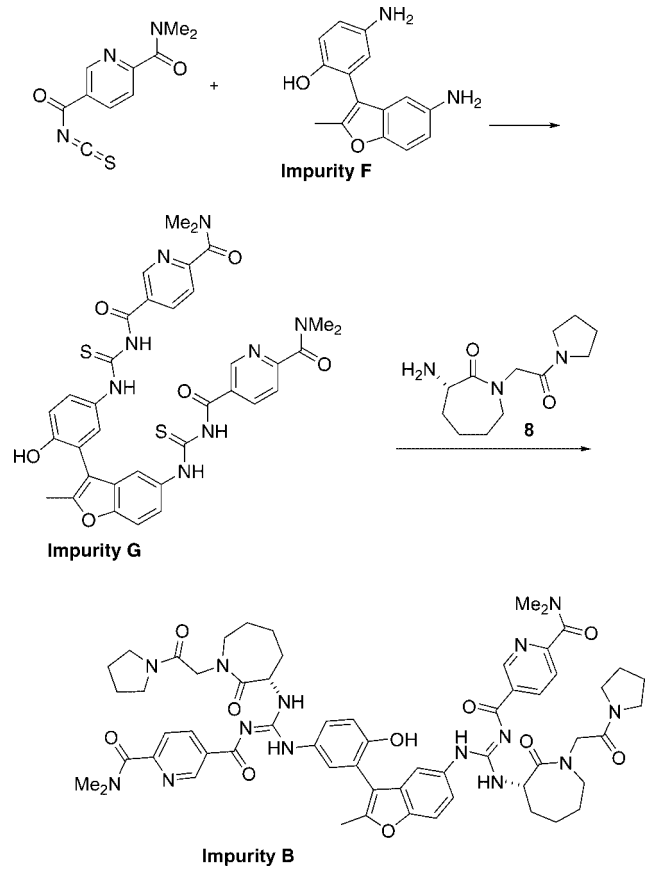
Figure 3. **Structures of impurities A−C.**

(8) Complete ^{1}H and ^{13}C NMR chemical shifts assignments of impurity B are shown in Tables 3 and 4 in the Supporting Information.
(9) Complete ^{1}H and ^{13}C NMR chemical shifts assignments of impurity C are shown in Tables 5 and 6 in the Supporting Information.

Scheme 2. **Proposed formation of impurity A during the synthesis of 1**



Scheme 3. **Proposed formation of impurity B during the synthesis of 1**



In order to confirm this hypothesis, a batch of penultimate **7** was analyzed by LC/MS to search for impurities C, E, and G. Three peaks with the expected MW of 724 for impurities E and G and 366 for impurity C were observed.[10] Impurity C detected in penultimate **7** was further confirmed by comparing its HPLC retention time with that of an authentic sample which was isolated and characterized from a batch of **1**. A batch of starting material **6** was also investigated by LC/MS. As expected, two impurity peaks with the same MW of 254, which are highly likely corresponding to impurities D and F, were detected in the starting material **6**.[11]

The Impact of the Established Impurity Profile on Process Research and Development. In a subsequent campaign to prepare a GMP batch (1 kg) of **1** to support a clinical study using the same synthetic scheme, impurities A, B, and C were observed at levels of 0.17, 0.55, and 0.46 AP, respectively. The levels of impurities A and B in this GMP batch were similar to those in the toxicological batch. The level of impurity C was significantly higher than the qualified threshold of 0.20 AP in the toxicological batch. A rework was necessary to reduce the level of impurity C in this GMP batch. With the knowledge gained through the course of the impurity isolation and structure elucidation, a rework protocol was developed based on the low solubility of impurity C in organic solvents. EtOH was chosen as the solvent for rework since it had already been used in the final crystallization of **1**. The GMP batch was dissolved in EtOH and filtered, removing most of the insoluble impurity C. The filtrate was treated with heptane to induce crystallization of **1**. The level of impurity C in the final API **1** was reduced from 0.46 AP to 0.14 AP.[5]

Conclusions

The structures of three critical impurities A−C and their origins of formation during the preparation of the API **1** were identified at an early stage of development. The structural knowledge led to the development of a rework protocol that purged out impurity C to an acceptable level for releasing a GMP batch. Further analysis revealed that impurity C was formed during the synthesis of penultimate **7** and carried over to the API step, which suggested that impurity C could be controlled at the penultimate step in future development. Similar solubilities and HPLC retention times compared to those of the API, in addition to the established structures of impurities A and B suggest that they have similar physicochemical properties as those of the API **1**, although they are pseudodimers. This could, presumably, be the reason why preliminary efforts to purge impurities A and B at the API step were not effective. However, from the knowledge about their root cause of formation, it is apparent that they should be controlled by setting appropriate specifications for their precursors, impurities D and F, in the starting material **6** if this synthetic route is selected in future development.

(10) LC/MS chromatograms of impurities C, E, and G and mass spectra of E and G are shown in Figures s2 and s3 in the Supporting Information.

(11) LC/MS chromatogram and mass spectra of impurities D and F are shown in Figure s4 in the Supporting Information.

Experimental Section

Analytical HPLC Conditions. Analytical HPLC analysis was conducted using a Phenomenex LUNA C18(2), 150 mm × 4.6 mm-i.d., 5 μm particle size column (Phenomenex, Inc., U.S.A.) at ambient temperature. The mobile phases consisted of A: 0.5 mL trifluoroacetic acid in 1000 mL water; B: 0.5 mL trifluoroacetic acid in 1000 mL acetonitrile. Program gradient elution (time (min)/% B = 0/16, 3/16, 40/53, 41/16, 51/16) was used with UV detection at 220 nm and a flow rate of 1.0 mL/min.

Preparative HPLC Conditions. Preparative scale separation were achieved using a YMC Pack Pro C18, 250 mm × 20 mm-i.d., 5 μm particle size column (Waters Corp. U.S.A.) at ambient temperature. The mobile phase consisted of water (A) and acetonitrile (B). The flow rate was set at 18 mL/min. The UV detection wavelength was selected at 220 nm. A linear gradient method of 40−70% B from 0 to 15 min was used as first-stage preparative separation. A second-stage preparative HPLC separation was conducted using a shallower linear gradient method of 30−50% B in a period of 0−30 min to further purify impurity A. Fractions of interest were pooled and concentrated via rotary evaporation. The concentrated fraction containing impurity B was extracted with CH_2Cl_2. The organic phase was dried over Na_2SO_4. The solvent was rotary evaporated to dryness to afford impurity B as an off-white powder. The concentrated fraction containing impurities A and C was placed in a refrigerator. Impurity C was precipitated and isolated by filtration and air drying. The filtrate was extracted with CH_2Cl_2 and subjected to second-stage preparative HPLC separation to isolate impurity A.

NMR Experimental Conditions. The NMR data were obtained on a Bruker Avance DRX 400 MHz spectrometer. Multidimensional spectra with ^{1}H detection were acquired with the operating frequency of 400.13 MHz at 25 °C using a Nalorac multinuclear 3-mm inverse probe with a *Z*-gradient. ^{13}C NMR spectrum was acquired with the operating frequency of 100.6 MHz at 25 °C using a Nalorac 3-mm dual probe.

Acknowledgment

We thank Dr. B. L. Kleintop for acquiring HRMS data for impurities A−C and Drs. P. Manchand, W.-S. Li, and D. A. Conlon, Process Research & Development Department for valuable discussions and contributions, respectively.

Supporting Information Available

Tables of ^{1}H and ^{13}C NMR assignments for impurities A−C and LC/UV/MS chromatograms and mass spectra of impurities D−G. This material is available free of charge via the Internet at http://pubs.acs.org

Received for review March 24, 2008.

OP800067V