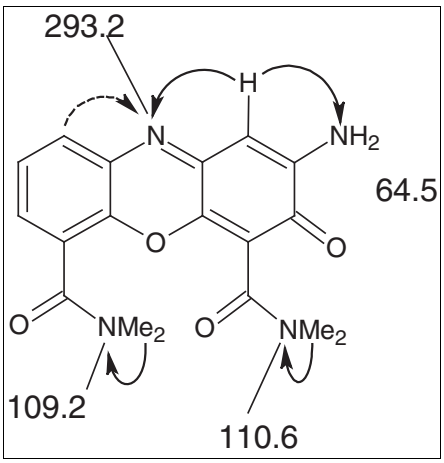
Structural Characterization of a Phenoxazine Analog Formed as a By-product of the Synthesis of a 3-Amino-*N,N*-dimethylsalicylamide

Wendy Zhong,[a] Jiong Yang,[a] Bruce D. Hilton,[a] Gary E. Martin,[a]*
Jianguo Yin,[b] Kiran Muppalla,[b] Weidong Tong,[b] and Xiaoyong Fu[b]

[a]Merck Research Laboratories, Discovery and Preclinical Sciences, Structure Elucidation Group, Summit, New Jersey 07901
[b]Merck Research Laboratories, Discovery and Preclinical Sciences, Process Chemistry, Rahway, New Jersey 07065
*E-mail: gary.martin2@merck.com

Received March 11, 2011
DOI 10.1002/jhet.981
Published online 29 October 2012 in Wiley Online Library (wileyonlinelibrary.com).



A highly colored impurity formed at low levels during the preparation of 3-amino *N,N*-dimethyl salicylamide was isolated and identified as 2-amino-*N,N,N′,N′*-tetramethyl-3-oxo-3*H*-phenoxazin-4,6-dicarboxamide using mass spectrometry and two-dimensional NMR methods that included long-range ^{1}H-^{15}N heteronuclear chemical shift correlation data to assemble the proton-deficient "right-half" of the 2-aminophenoxazin-3-one nucleus. The structure was independently confirmed using computer-assisted structure elucidation methods and by synthesis.

J. Heterocyclic Chem., **49**, 1163 (2012).

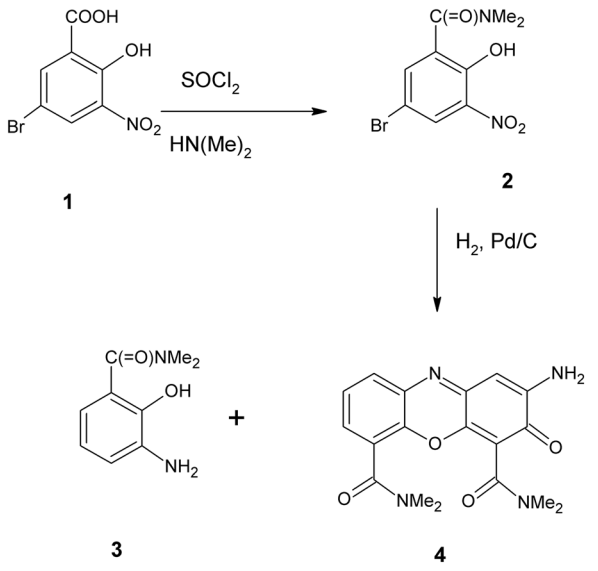
INTRODUCTION

By-products of chemical reactions are problematic in any chemical synthesis. In the case of pharmaceuticals, low-level impurities formed by side reactions can constitute a diverse array of problems that range from issues of pharmaceutical elegance when colored impurities form in the final drug substance to unwanted pharmacologic activity. For these reasons, it is often desirable to investigate the structures of by-products of reactions used to prepare chemical synthons for research programs. We recently had need of 3-amino-*N, N*-dimethylsalicylamide, **3**, as a starting material for a reaction sequence to make compounds as potent and selective CXCR2 chemokine receptor antagonists [1–4]. When the salicylamide analog was prepared (Scheme 1) an HPLC chromatogram showed several low-level impurities to be present. One of the major impurities was compound 4. A decision was made to isolate and characterize this impurity to help to understand its origin and fate along the synthesis.

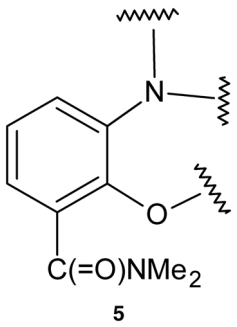
RESULTS AND DISCUSSION

The deep reddish color of the isolated impurity in the solution suggested it to be highly conjugated. The molecular formula of the impurity was determined using a Thermofisher LTQ Orbitrap™ mass spectrometer equipped with a Surveyor HPLC system. The measured accurate mass of the protonated unknown impurity was *m/z* 355.14069, which corresponded to an empirical formula of $C_{18}H_{19}N_4O_4$ $[M+H]^+$, with an error of 1.70 ppm between the measured accurate mass and calculated exact mass. The molecular formula also suggests that the impurity has highly ordered double bond and ring character. A sodium adduct was also observed at *m/z* 377.12277, with an error of 1.98 ppm.

Scheme 1



and the one valence of the oxygen were not satisfied. The remainder of the empirical formula was $C_6H_9N_2O_2$. Seven protons are accounted for by an aromatic methine (6.52/99.9 ppm $^1H/^{13}C$) and two *N*-methyl groups. Dimeric structures containing N N or N N linkages could be ruled out as structural possibilities based on the presence of only the single remaining sp^2 methine based on the HSQCAD spectrum of the isolate.



5

Multistage MS/MS experiments were then performed to obtain additional structural information. The major fragment ion at *m/z* 310 arose from the loss of a dimethylamino group. The MS^3 experiment performed on the *m/z* 310 fragment ion resulted in a major fragment ion at *m/z* 267, which is the result of the loss of the second NMe_2 group. These data suggest that the impurity molecule contains two NMe_2 groups. The MS^4 experiment of the fragment ion at *m/z* 267 gave the major fragment ion at *m/z* 211, which can be attributed to the sequential loss of two CO groups. Hence, the data indicated the presence of two *N,N*-dimethylcarboxamido moieties in this unknown impurity. The MS^5 experiment of the *m/z* 211 fragment ion gave a major fragment ion at *m/z* 183, which is the result of another CO loss from *m/z* 211. The MS data required that at least three CO moieties were present in the molecule.

The proton NMR spectrum at 600 MHz in d_4-MeOD was quite simple, containing aromatic resonances for three coupled, contiguous aromatic protons ($^1H/^{13}C$ HSQCAD; 7.80/130.4 dd; 7.50/127.2, app. triplet; 7.46/128.7, dd), a singlet at 6.53/99.9 ppm, and four methyl singlets at 3.17/35.4, 3.15/35.0, 2.96/38.3, and 2.90/39.0 that can be attributed to *N*-methyl groups on the basis of the proton and carbon chemical shifts. There was a very broad resonance in the aliphatic region of the spectrum but it could not be accurately integrated to determine how many protons it represented.

The ^1H-^{13}C HSQCAD and GHMBCAD spectra allowed the facile assembly of the portion of the molecule that was clearly derived from the desired 3-amino-*N,N*-dimethylsalicylamide as shown by **5**. As there were no proton resonances that could be attributed to an aniline amino group or a phenol, the two valences of the nitrogen

The ^{13}C one-dimensional spectrum of the impurity isolate contained 24 nonsolvent resonances, six of which could be assigned to a secondary, low-level impurity in the sample on the basis of GHMBC correlations. Eliminating the extraneous resonances in the ^{13}C spectrum from consideration left 18 carbons in agreement with the molecular formula determined by accurate mass measurements. Nine carbons can be assigned to **5**. The long-range ^1H-^{13}C GHMBC spectrum established correlations from the four *N*-methyl singlets to two carbonyl resonances at 166.1 and 168.6 ppm. The latter was assigned as the *N,N*-dimethyl amide carbonyl contained in **5** based on a long-range correlation from the aromatic proton resonating at 7.46 ppm. The only correlations to the carbonyl resonating at 166.1 are from the N-methyl groups. This latter *N,N*-dimethyl amide and its attached carbonyl accounts for three more carbons, giving a total of twelve. One additional carbon was accounted for from the CH pair resonating at 99.9/6.52 ppm in the HSQCAD spectrum. The remaining five ^{13}C resonances observed at 114.1, 146.0, 148.9, 149.5, and 178.9 ppm were all nonprotonated resonances.

The proton singlet resonating at 6.52 ppm exhibited long-range ^1H-^{13}C correlations in the GHMBCAD spectrum to the quaternary carbons resonating at 146.0, 148.9, and 178.9 ppm. The first two resonances, and the carbon resonating at 149.5, are suggestive of quaternary aromatic carbons bearing a nitrogen or oxygen; the latter carbon, resonating at 178.9 ppm, is consistent with a carbonyl. The carbon resonating at 114.1 ppm is a quaternary sp^2 carbon with no coordination to nitrogen or oxygen and likely with oxygen and/or nitrogen atom(s) positioned two bonds (β) distant to be consistent with the upfield shift of this quaternary carbon. To further supplement the long-

range proton–carbon correlations observed in the ^{1}H-^{13}C GHMBCAD spectra, a 6 Hz optimized ^{1}H-^{15}N GHMBCAD spectrum was acquired, which proved to be the key to the elucidation of the impurity structure. The proton singlet resonating at 6.53 ppm was long-range correlated to nitrogens resonating at 293 and 64 ppm. The ^{15}N shift of the former is suggestive of a C N nitrogen and the latter can be attributed to an aromatic NH_2 group. The latter could be associated with the broad resonance in the aliphatic region of the proton spectrum.

An H/D exchange experiment was conducted to determine how many exchangeable protons were present in the molecule. The measured accurate mass of the deuterated molecule was *m/z* 358.15765, which corresponds to an empirical formula of $C_{18}H_{16}D_3N_4O_4$ $[M + D]^+$, with an error of −3.51 ppm between the measured accurate mass and calculated exact mass. The empirical formula of the deuterated molecule incorporated three deuterium atoms. The deuterium exchange data suggested that two exchangeable protons were replaced by deuterium and the third deuterium is the ionizing deuteron, which confirmed the incorporation of two deuterium atoms in the structure. Combined H/D exchange MS and NMR data provided strong confidence that a primary amine moiety was present in the molecule.

When the long-range ^{1}H-^{13}C and ^{1}H-^{15}N GHMBC data were integrated with substructure **5**, a structure based on a phenoxazine nucleus emerges that is consistent with all the observed ^{1}H, ^{13}C, and ^{15}N chemical data as well as all long-range ^{1}H-^{13}C and ^{1}H-^{15}N correlations. The upfield shift of the 6.53/99.9 ppm $^{1}H/^{13}C$ resonant pair is readily explained by nitrogens flanking this carbon. The observed long-range correlations and the ^{13}C and ^{15}N chemical shift assignments are shown in Figure 1. The balance of the carbon resonance assignments for the tri-substituted phenyl ring of the phenoxazine nucleus are consistent with ^{13}C chemical shift calculations performed using neural network methods with ACD's CNMR program, v. 11.01. Dashed arrows denote weak correlations; double-headed arrows denote mutually long-range coupled protonated carbons; solid and dashed black arrows denote ^{1}H-^{15}N long-range correlations.

The proposed fragmentation pathways for the protonated (deuterated) impurity derived from the multistage MS/MS experiments are shown in Figure 2. The fragmentation patterns of the deuterated molecule are very similar to that of the protonated molecule; each fragment ion in the deuterated molecule was shifted two daltons higher relative to the corresponding ion of the protonated molecule. The MS data suggested that the primary anilino group was present in each fragment ion and was not readily lost. This observation provides an additional confirmation that the impurity structure contains an aryl amino group that cannot easily fragment during low energy activation processes.

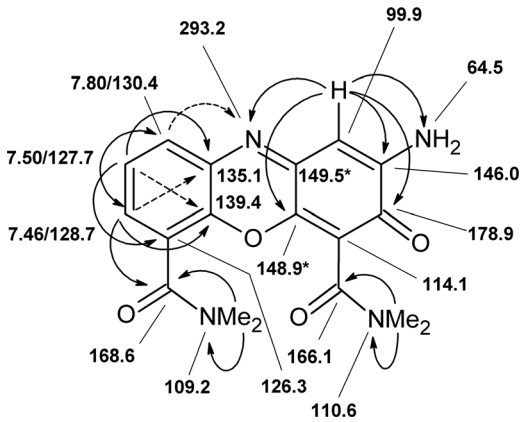


Figure 1. Structure of **4** showing ^{1}H, ^{13}C, and ^{15}N chemical shift assignments and long-range ^{1}H-^{13}C and ^{1}H-^{15}N long-range heteronuclear shift correlations. ^{1}H-^{13}C and ^{1}H-^{15}N long-range correlations are designated by black arrows. Double-headed arrows denote mutually long-range coupled resonances. Weak correlations are denoted by dashed arrows.

COMPUTER-ASSISTED STRUCTURE ELUCIDATION

Although determining the structure of the phenoxazinone, **4**, did not require computer-assisted structure elucidation (CASE) methodology, it is still instructive to input the spectral data into a CASE program, such as ACD's Structure Elucidator™ program, v11.01, to examine the output for alternative structures consistent with the spectral data. This approach is germane in the present case since there were two carbons with no long-range correlations (resonating at 114.1 and at 149.5 ppm) and because of the paucity of correlations from the right half of the molecule. All the data collected for **4** were used as input to the program. In a case where multiple atoms are incorporated in a structure that does not have any 2D NMR correlations to them, it is crucial that unambiguous structure elements be included in the constraints used for a CASE program insofar as possible. Constraints imposed on this CASE computation run were derived directly from the data discussed above included: (1) the two *N,N*-dimethyl amide groups (including their respective carbonyl groups) were constructed by hand; (2) a fragment that comprises the three contiguous, mutually coupled aromatic protons and their respective carbons; and (3) the NH_2 group was assembled as implied by the ^{1}H-^{15}N GHMBCAD data and by the deuterium exchange MS data. These constraints were unambiguously derived from the ensemble of mass spectrometric and NMR spectroscopy experiments performed. For the initial CASE computation, the requirement was imposed that the carbon resonating at 178.9 ppm had to be considered as a carbonyl. Although reasonable, this assignment was not clear-cut. When the program was launched, the calculation took 85 min and generated 62 structures after filtering, with nine structures remaining in the output table after the removal of duplicates.

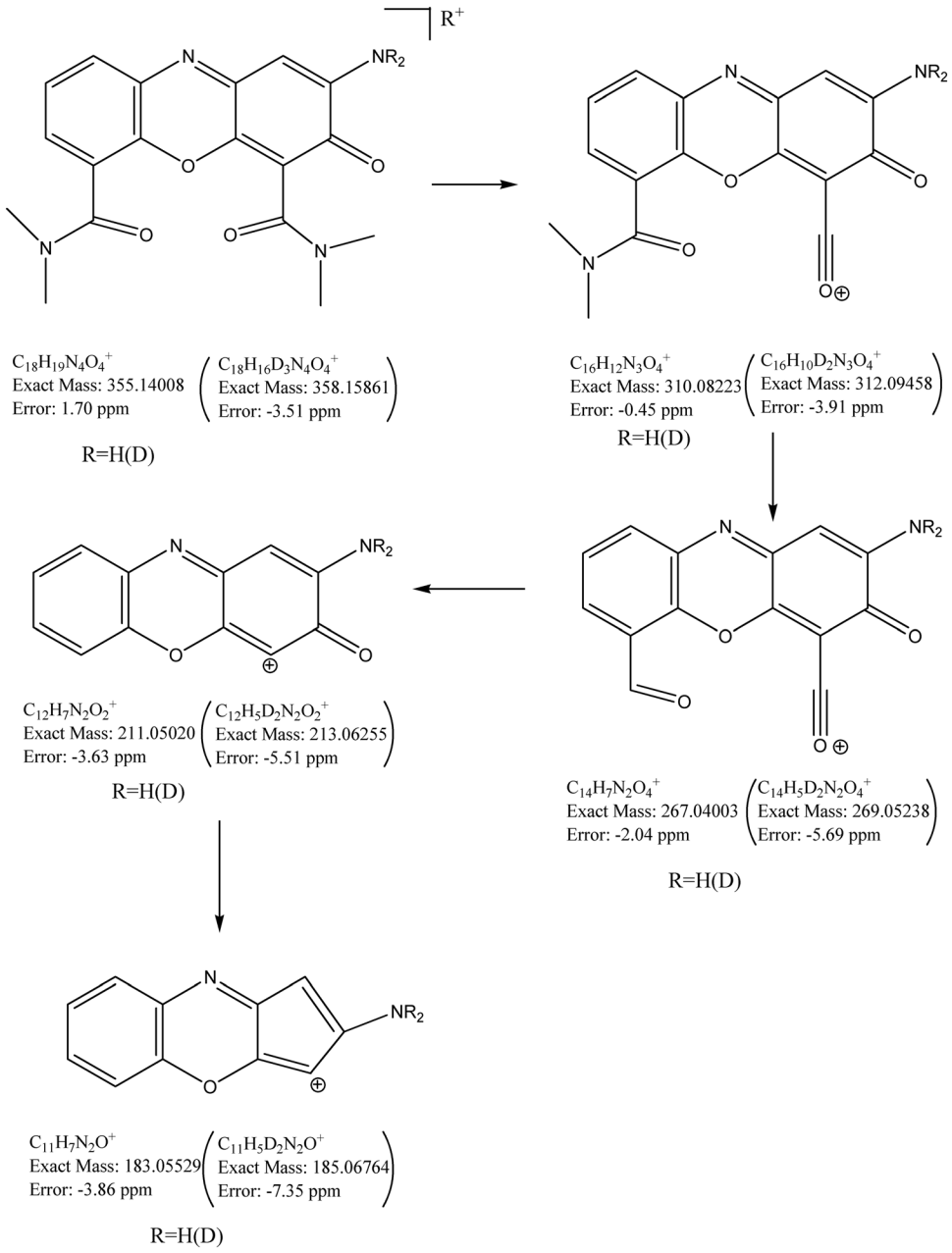


Figure 2. The proposed fragmentation pathways for protonated (deuterated) **4** are shown above.

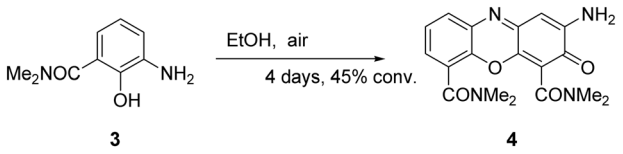
In contrast, when the imposed carbonyl constraint for the 178.9 ppm ^{13}C resonance was removed, the calculation took overnight, the program generating ~ 3.7 million structures before filtration. The results of the longer run, however, confirmed the results obtained with the carbonyl constraint included; no reasonable new structures were produced by the program.

The nine structures remains after filtering to remove duplicate structures are shown in Figure 3. The structure in the first position of the output table, which was sorted on the basis of the ^{13}C neural net chemical shift calculation match factor relative to the experimental ^{13}C data, exhibited only a slightly better match factor of $d_N(^{13}C) = 3.352$ relative to the match factor of $d_N(^{13}C) = 3.381$ for the correct structure in position two of the output table. There were no plausible alternative structures generated by the program that could account for all the data.

SYNTHESIS

It is known that electron-rich 2-aminophenol derivatives can undergo catalytic oxidative cyclocondensation to form 2-amino-3*H*-phenoxazin-3-one derivatives [5]. Formation of the low-level impurity **4** from compound **3** could occur

via an oxidative (air) cyclocondensation mechanism. When a solution of compound **3** in ethanol was stirred at room temperature for more than four days in a flask open to air, about half of the material decomposed, resulting in a complex mixture containing about 15% of compound **4**.



We then performed the synthesis of the compound 4 using conditions described in Ref. [1] as shown in Scheme 2. The reaction gave the desired compound spectroscopically identical to **4** in 30% isolated yield.

EXPERIMENTAL

Preparative chromatography. A sample of 5 g of 3-amino-*N,N*-dimethylsalicylamide (**3**) containing the impurity (**4**) was processed using a Phenomenex Synergi Polar-RP, 250 × 15 mm^2, 4 μ, column and eluted with 40/60 MeOH/0.1% TFA in water (flow rate: 8 mL/min; run time:12.2 min). Three fractions were

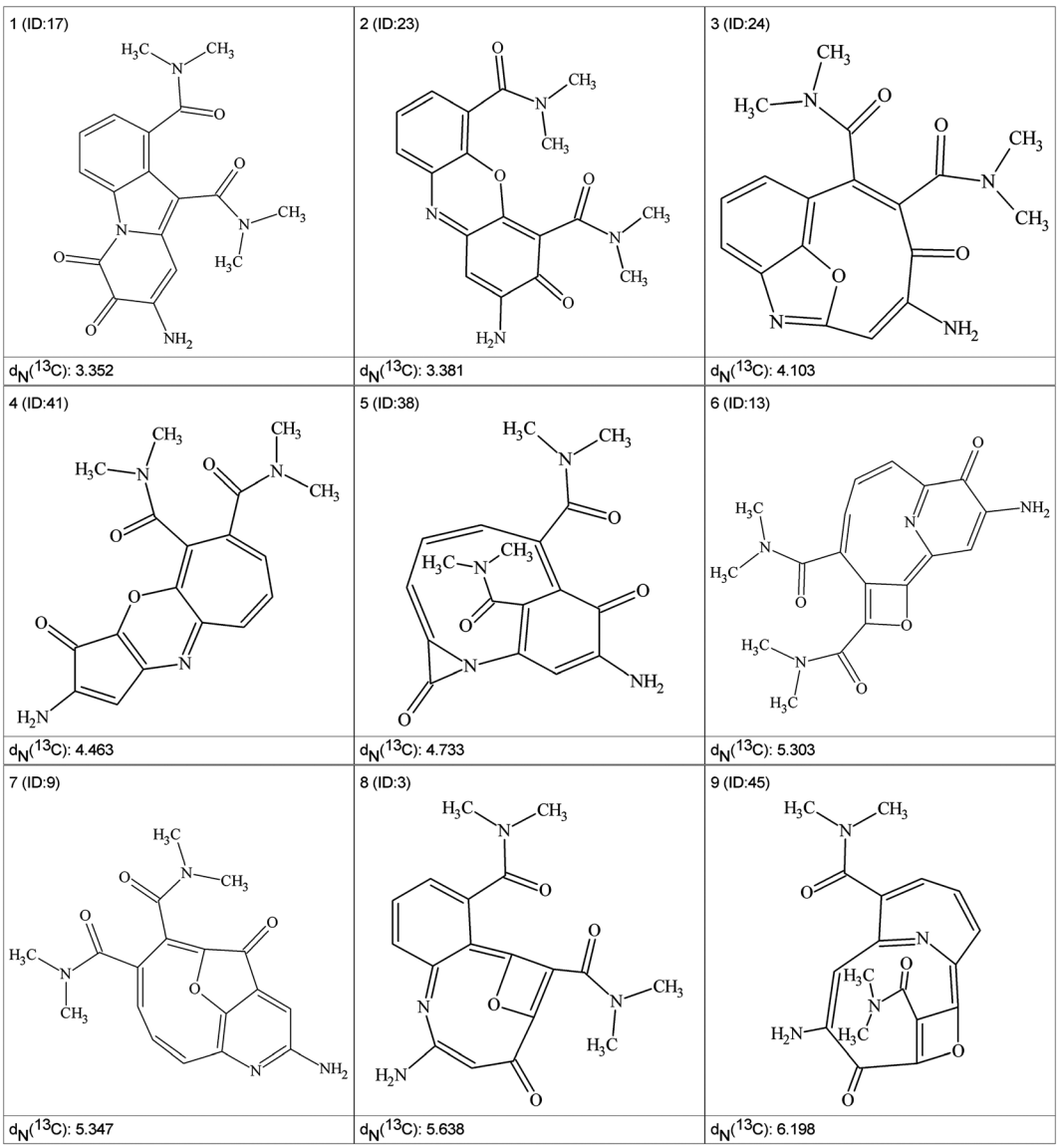
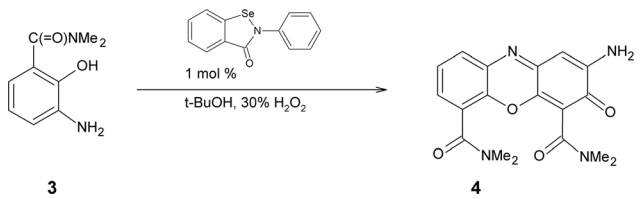


Figure 3. The nine structures generated by the Structure Elucidator v11.01 CASE program are shown. A total of 62 structures were generated and stored after filtration in 85 min of computation time. Nine structures were stored after the removal of duplicates. The correct structure, **4**, which is also fully consistent with the chemistry involved, is shown in the second position with a $d_N(^{13}C) = 3.381$, which is only very slightly less than the structure in the first position with $d_N(^{13}C) = 3.352$. When the constraint of requiring the ^{13}C resonance at 178.9 ppm to be considered as a carbonyl was removed, the program generated ~ 3.7 million structures in an overnight calculation before filtration. There were no structures generated by removing the constraint that are consistent with the data that are not included in the nine structures above.

Scheme 2



3 **4**

collected containing the *m/z* 354 target impurity in 13, 40, and 67%, respectively. The fraction containing 67% of the target impurity was concentrated five-fold before further purification by preparative HPLC using a YMC Pack pro C_{18} 10 × 150 mm^2 3 μ column, with a flow rate of 3.2 mL/min using a gradient 24/76 of MeOH 1% TFA in water for 17 min then ramped to 80/20 in 1 min and back to 24/76 in 1 min. The total run time was 22 min with elution monitored at 320 nm. The fractions were dried using a rotovap followed by lyophylization. Stability data showed that the isolated impurity was stable in methanol but unstable in dimethyl sulfoxide.

NMR spectroscopy. All NMR experiments were performed using a sample prepared by dissolving ~ 0.7 mg of the isolated reaction by-product in 200 μL methanol-d_4 (Cambridge Isotope Laboratories), after which it was transferred to a 3 mm NMR tube (Wilmad) using a flexible Teflon™ needle and a gas-tight syringe (Hamilton). The NMR data were acquired using a 600 MHz Varian three channel NMR spectrometer equipped with a 5 mm gradient triple resonance Varian ColdProbe™ operating with the RF coil at 20 K. The sample temperature was regulated at 25°C. The NMR experiments consistent of a proton reference spectrum, GCOSY, ^{1}H-^{13}C HSQCAD, 8 Hz optimized ^{1}H-^{13}C GHMBCAD, and a 6 Hz optimized ^{1}H-^{15}N GHMBCAD spectrum. All experiments were performed using experiments from the standard pulse program library. A ^{13}C reference spectrum was also acquired using the X-coil of the ColdProbe. The NMR data were processed off line using ACD SpecManager™ and Structure Elucidator™ both v11.01.

Mass spectrometry. LC/MS experiments were conducted on the sample using a Thermo LTQ-Orbitrap instrument equipped with a Surveyor HPLC system. Multistage MS/MS experiments were conducted in the linear trap portion of the instrument by direct infusion at a flow rate of 5 μL/min, with the helium as the collision gas. LC/MS experiments were performed under the gradient conditions: 20% B to 100% B in 15 min at the flow rate of 200 μL/min. Mobile phase compositions were: A: 95% HPLC grade H_2O (10 mM NH_4OAc)/5% ACN, and B: 5% H_2O (10 mM NH_4OAc)/95% ACN. The sample was eluted using an XBridge C18 HPLC column, 2.1 × 150 mm^2, with 3.5 μm particle size and 5 μL injection volume. H/D exchange experiments were conducted *via* direct infusion and the sample was dissolved in ACN/D_2O with the concentration of 0.5 mg/mL. The sample was introduced through ESI source at a flow rate of 5 mL/min.

Synthesis. To a mixture of 10 g of 3-amino-*N,N*-dimethylsalicylamide (**3**) and selenium catalyst (Ebselen, 100 mg; 1 mol %) in 150 mL of t-butanol was added 25 mL of 30% hydrogen peroxide. The reaction mixture was stirred at room temperature for 20 h. The precipitated solid was collected by filtration and recrystallized from 100 mL of methanol to give 2.9 g (30% yield) of red solids in >99% purity that matched the chromatographic retention time, MS, and NMR data contained in this report for **4**.

CONCLUSIONS

The structure of a highly colored oxidative cyclization by-product of the synthesis of 3-amino-*N,N*-salicylamide (**3**) was elucidated using an ensemble of 2D heteronuclear chemical shift correlation data, in which the long-range ^{1}H-^{15}N GHMBC data were pivotal. Without the latter data, it is quite likely that it would have been impossible to correctly elucidate the structure of this by-product of the synthesis as **4** because of the proton-deficient nature of the right half of the phenoxazine core of the molecule. Accurate mass measurements, MS/MS, and deuterium exchange studies afforded data that were consistent with the structure derived from the NMR data. When the spectroscopic data were used as input for the ACD Structure Elucidator v11.01 CASE program, the program generated the spectroscopist-determined structure, which was in the second position of the output table of nine structures. None of the other eight structures generated by the program were reasonable, given the nature of the synthetic route thereby serving to further validate the elucidated structure. In parallel, the structure of the identified impurity was also confirmed by the synthesis of an authentic sample. Armed with knowledge of the structure of the unwanted by-product of the reaction, steps can be taken to engineer this by-product out of the reaction by modifying reaction conditions, reaction times, solvents, or by changing reagents.

REFERENCES AND NOTES

[1] Dwyer, M. P.; Yu, Y.; Chao, J.; Cynthia, A.; Chao, J.; Biju, P.; Girijavallahan, V.; Viyyoor, R.; Rindgen, D.; Bond, R.; Mayer-Ezel, R.; Jakway, J.; Hipkin, W. R.; Fossetta, J.; Gonsiorek, W.; Biao, H.; Fan, X.; Terminelli, C.; Fine, J.; Lundell, D.; Merritt, J. R.; Rokosz, L. L.; Kaiser, B.; Li, G.; Wang, W.; Stauffer, T.; Ozgur, L.; Baldwin, J.; Taveras. A. G. J Med Chem 2006, 49, 7603.

[2] Biju, P.; Taveras, A.; Yu, Y.; Zheng, J.; Hipkin, R. W.; Fossetta, J.; Fan, X.; Fine, J.; Lundell D. Bioorg Med Chem Lett 2009, 19, 1434.

[3] Liu, S.; Liu, Y.; Wang, H.; Ding, Y.; Wua, H.; Dong, J.; Wong, A.; Chen, S.; Li, G.; Chan, M.; Sawyer, N.; Gervais, F.; Henault, M.; Kargman, S.; Bedard, L.; Han, Y.; Friesen, R.; Lobell, R.; Stout, D. Bioorg Med Chem Lett 2009, 19, 5741.

[4] Chao, J.; Taveras, A.; Chao, J.; Aki, C.; Dwyer, M.; Yu, Y.; Biju, P.; Rindgen, D.; Jakway, J.; Hipkin, W.; Fosetta, J.; Fan, X.; Lundell, D.; Fine, J.; Minnicozzi, M.; Phillipsa, J.; Merrittb, J. R.; Bioorg Med Chem Lett 2007, 17, 3778.

[5] Giurg, M.; Piekielska, K.; Gebala, M.; Ditkowski, B.; Wolanski, M.; Peczynska-Czoch, W.; Mlochowski, J. Synth Commun 2007, 1779.