Characterization of a Trace By-product of the Synthesis of the Protease Inhibitor Tipranavir (PNU-140690)

Gary E. Martin*, Bruce A. Pearlman [a], Russell H. Robins, Frank W. Crow, Wayne K. Duholke, Jane E. Guido, Chad E. Hadden, Brian D. Kaluzny, and Thomas J. Thamann

Rapid Structure Characterization Group, Pharmaceutical Development
[a] Chemical Process Research & Development, Pharmacia & Upjohn
Kalamazoo, Michigan 49001-0199
Received January 4, 1999

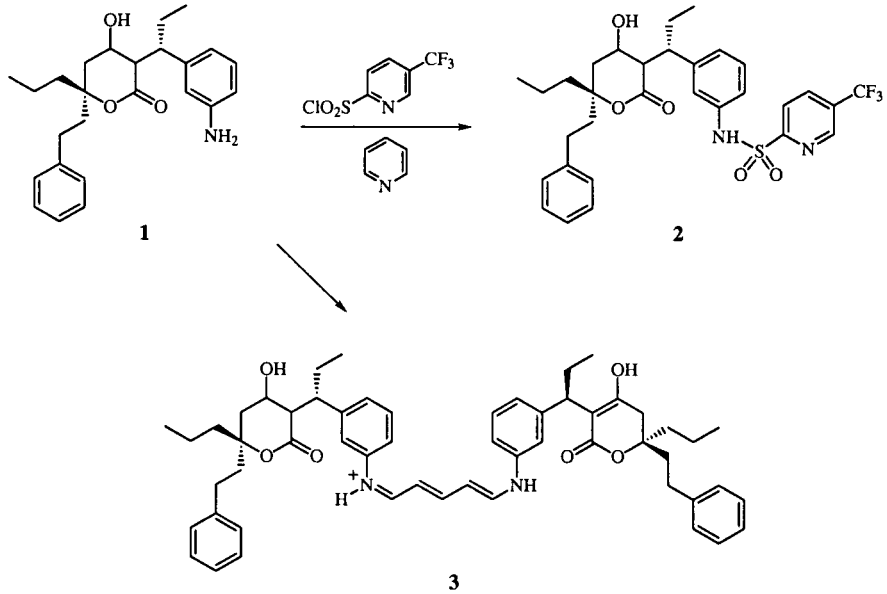
Tipranavir™ (PNU-140690) is a protease inhibitor under clinical investigation for the treatment of human acquired immunodeficiency syndrome (AIDS). During scale-up synthesis of clinical quantities of the bulk drug, a colored, transient by-product of the final coupling reaction was observed. Quantities of this colored, transient chemical species were too low (<<0.1%) for characterization by conventional spectroscopic methods. It was, however, possible to isolate sufficient material for characterization based on mass spectrometry and submicro inverse-detection gradient (SMIDG) nmr methods by methanol stripping of silica gel that had been used in purification of bulk drug. This process afforded an enriched feedstock from which small quantities of this highly colored and unstable (halflife < 18 hours in methanol and < 10 minutes in acetone) trace contaminant could be isolated by semipreparative reversed phase hplc. The impurity was identified as an unstable Zincke salt formed by the condensation of two molecules of the anilino precursor and the pyridine used as a base in the final step of the synthetic process. Following identification of this impurity, efforts were undertaken to engineer it out of the synthetic process.

J. Heterocyclic Chem., **36**, 1107 (1999).

Structural characterization of trace impurities in pharmaceuticals represents a significant analytical challenge. The characterization of impurities present at levels ≥0.1% is required by regulatory mandate. Occasionally, it is necessary to identify impurities present at lower levels (<<0.1%), such as those that impact toxicology or "pharmaceutical elegance". Ideally, following identification, such impurities can be engineered out of the synthetic process. One such elegance problem, a highly colored trace level by-product, was observed early in the development of commercial scale synthesis of tipranavir™ (**2**, PNU-140690) [1], a protease inhibitor presently undergoing clinical investigation for the treatment of human acquired immunodeficiency syndrome (AIDS) [2,3].

Historically, costly, and time-consuming large scale chromatography has been necessary to obtain sufficient quantities of trace level (<<0.1%) impurities for characterization by mass spectrometry and conventional nmr methods. Characterization is further burdened when impurities are unstable. SubMicro Inverse-Detection Gradient or SMIDG™ nmr was developed, in part, to circumvent the inherent insensitivity of nmr spectroscopic methods. SMIDG nmr probe technology greatly facilitates the characterization

Scheme I



of extremely small samples [4,5], or the characterization of transient species when the data must be collected rapidly prior to sample decomposition [6]. Using mass spectrometry in conjunction with SMIDG nmr probe technology, it was possible to characterize the colored impurity formed as a by-product of the coupling reaction in the original synthesis of tipranavir (**2**, PNU-140690) [1].

Results and Discussion.

The final step in the synthesis of tipranavir (**2**, PNU-140690) is shown in Scheme I and involves the sulfonylation of the aniline, **1**, with the trifluoromethylpyridinesulfonyl chloride to yield the drug, **2**, as shown in Scheme I. Work early this century by Zincke [7,8] reported the formation of highly colored, unstable salts from the reaction of aniline in the presence of pyridine under conditions similar to those used in the final step of the synthesis of tipranavir (**2**, PNU-140690) [1].

Methanol stripping of silica gel that had been used in purification of bulk drug afforded feedstock enriched in the red-colored impurity. Reversed phase semipreparative hplc on a 20 x 250 mm Kromasil C_{18} column eluted with gradient acetonitrile: aqueous trifluoroacetic acid or methanol:aqueous ammonium formate solvent systems afforded the red-colored, trace by-product. Two chromatographic passes were necessary to obtain an acceptable level of purification. Chromatography using methanol: aqueous ammonium formate gave a more stable isolate. Pooled fractions were diluted with Milli-Q water and concentrated, desalted, and reduced in chemical noise-causing components by trapping on a 10 x 250 mm Kromasil C_{18} column. The eluant from this final column was freeze-dried to yield ~0.5 mg of material that was ~80% pure (analytical hplc, 254 nm detection). Preliminary lc/ms data gave a parent ion at 850 da that could be interpreted to suggest some form of dimeric structure.

Proton nmr and preliminary 2D nmr experiments were performed on half of the isolated material; the balance of the sample was retained, frozen at -80° in case of sample decomposition. All nmr data were acquired using a Varian *INOVA* 600 nmr spectrometer equipped with a Nalorac SMIDG™-600-1.7 submicro inverse-detection gradient nmr probe using a sample prepared in 30 µl of perdeuteromethanol. The proton spectrum contained resonances consistent with the presence of *n*-propyl and β-phenethyl groups and an anisochronous methylene resonance pair arising from the δ-lactone moiety of tipranavir (**2**, PNU-140690). Resonances were also observed that were consistent with the 1,3-disubstituted phenyl moiety as well as the 1,1-disubstituted *n*-propyl linkage between it and the δ-lactone. Resonances for the 2,5-disubstituted pyridyl moiety normally observed in the ^{1}H nmr spectra of the bulk drug were not observed in the spectra of the red isolate. Rather, new resonances were observed in the aromatic/vinyl region corresponding to a doublet at 8.25 δ (J = 13 Hz), a "triplet" resonating at 7.82 δ (J = 13 Hz), and a second "triplet" resonating at 6.27 δ (J = 13 Hz).

An ^{1}H nmr spectrum was acquired with a 90 second interpulse delay to enhance integration accuracy. Using the methine proton of the 1,1-disubstituted n-propyl moiety as an integration reference, the aromatic region integrated for nine protons as expected for the β-phenethyl and 1,3-disubstituted phenyl moieties. The integrals for the other aliphatic resonances were as expected. The previously unobserved resonances at 8.25, 7.82, and 6.27 ppm, however, were anamalous, giving integrals of 1.0, 0.5 and 1.0 protons, respectively, relative to the methine of the 1,1-disubstituted *n*-propyl group. These data require the structure of the impurity to contain two aniline-derived moieties, giving a ratio of 2:1:2 for the resonances at 8.25, 7.82, and 6.27 ppm, consistent with initial assumptions derived from the preliminary lc/ms data.

High resolution mass spectrometry was performed using a Finnigan MAT 900ST magnetic sector mass spectrometer. Measurements were carried out by linear E-scan peak matching at a resolution of 9700 (m/Δm, 10% valley definition) using the protonated molecular ion from the bisindole alkaloid vincristine (825.40747 da) as a reference ion. The accurate mass was measured as 848.47287 da (less the ionizing proton), which fits to an empirical formula of $C_{55}H_{64}N_2O_6$ to within 4 ppm. This empirical formula is consistent with a structure containing two aniline-derived moieties (see **3** in Scheme I) and a C_5H_5 subunit.

The collisionally induced decomposition (CID) product ion spectra of the isolated impurity were recorded using a Finnigan TSQ-7000 mass spectrometer equipped with a µESI interface of in-house design and construction [9]. A small portion of the impurity isolate was dissolved in 1:1 acetonitrile:water containing 2% formic acid. There were several significant fragment ions observed, including ions at 456 and 430 da, each of which had corresponding ions (412 and 386, respectively) that could be accounted for by the loss of CO_2 from the δ-lactone moiety. These ions correspond to structural fragments incorporating the C_5H_5 fragment. Since all valences in the aniline represented by **1** were accounted for other than those through the amino group, these data suggested the linkage of the two anilino nitrogens *via* an unsaturated five carbon fragment as shown by **3**. Synthetically, **3** can be derived from the ring-opening of pyridine to react with two aniline molecules to afford a Zincke salt in a manner analogous to that described by Zincke [7,8]. The µESI CID product ion spectrum of **3** is shown in Figure 1.

A homonuclear TOCSY spectrum was recorded at 600 MHz in 23 minutes using a 24 msec mixing time (not shown), which established vicinal proton-proton connectivities throughout the molecular structure. These data also confirmed the mutual coupling of protons resonating at 8.25, 7.83, and 6.27 ppm, as was expected from the proton reference spectrum.

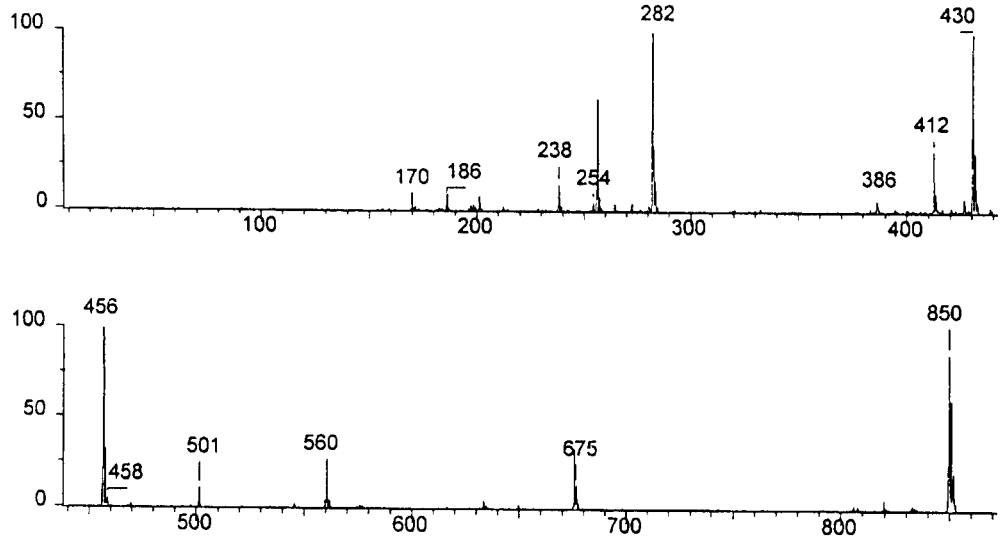


Figure 1. Collisionally Induced Decomposition (CID) product ion mass spectrum of **3** recorded using a Finnigan TSQ-7000 mass spectrometer with argon as a collision gas.

Correlations from a 300 msec ROESY spectrum (not shown) acquired from the sample in 45 minutes established the proximity of the proton resonating at 8.25 to the protons resonating at 7.41 and 7.09 ppm in the 1,3-disubstituted phenyl ring flanking the point of attachment of the anilino nitrogen.

Acquisition of a preliminary GHSQC gave the chemical shifts of the carbons directly bonded to the protons resonating at 8.25 and 6.27 ppm. The 8.25/152.7 and 6.27/106.2 ppm heteronuclear shift pairs are reasonable for a portion of the five carbon bridge linking the two anilino nitrogens as shown by **3**. The initial GHSQC spectrum of **3** was acquired in 1 hour 5 minutes with the one-bond delay optimized for a 140 Hz one-bond coupling. The data were acquired on ~250 μg of the colored isolate. These data are shown in Figure 2. It was necessary to resort to a SMIDG nmr probe both for reasons of the small size of the sample and the limited half-life of the material. The correspondingly lower concentrations inherent to the use of 3 mm or conventional 5 mm nmr probe formats would likely have precluded the successful acquisition of heteronuclear shift correlation data to verify the structure of **3** prior to degradation.

In an effort to determine the ^{13}C nmr chemical shift of the carbon directly bound to the proton resonating at 7.82 ppm, a second GHSQC spectrum was acquired in 2 hours 9 minutes, with the one-bond delay optimized for 160 Hz. The expanded aromatic region of this spectrum is shown in Figure 3. Despite the 160 Hz optimization for aromatic one-bond couplings and essentially doubled acquisition time, the experiment still failed to afford the needed chemical shift data. By the completion of data acquisition for the second GHSQC spectrum, the sample was substantially degraded. The second half of the isolated impurity sample, ~250 μg, was used to prepare a second 30 μl SMIDG nmr sample. The second SMIDG nmr sample was used to acquire the GHMBC data discussed below. In addition, it is worth noting that the chemical shift of the carbon directly bound to the proton resonating at 7.82 ppm was also determined from the 10 Hz optimized GHMBC experiment as described below.

A GHMBC experiment optimized for an assumed 10 Hz long-range coupling was acquired overnight in 12 hours 30 minutes giving the spectrum shown in Figure 4, and the aromatic expansion shown in Figure 5. As will be noted from some of the responses associated with weak contaminant peaks in the reference spectrum, there was significant degradation of the sample during the duration of the GHMBC data acquisition. The proton reference spectrum acquired prior to the acquisition of the GHMBC data is plotted above the contour plots shown in Figures 4 and 5. Pertinent long-range correlations are shown in Figure 5. The GHMBC data linked the proton resonating at 8.25 ppm to the nitrogen-bearing quaternary carbon of the 1,3-disubstituted phenyl ring as well as to the protonated carbon resonating at 125.8 ppm which is directly bound to the proton resonating at 7.82 ppm. As noted above, this resonance could not be assigned from the GHSQC data presented in Figures 2 and 3 below.

The appearance of the proton resonances of the five carbon linkage between the anilino moieties can be accounted for by averaging associated with tautomeric equilibria and the intramolecular transfer of the proton between the two anilino nitrogens. The correlation of the proton resonating at 6.27 ppm essentially with itself through the correlation

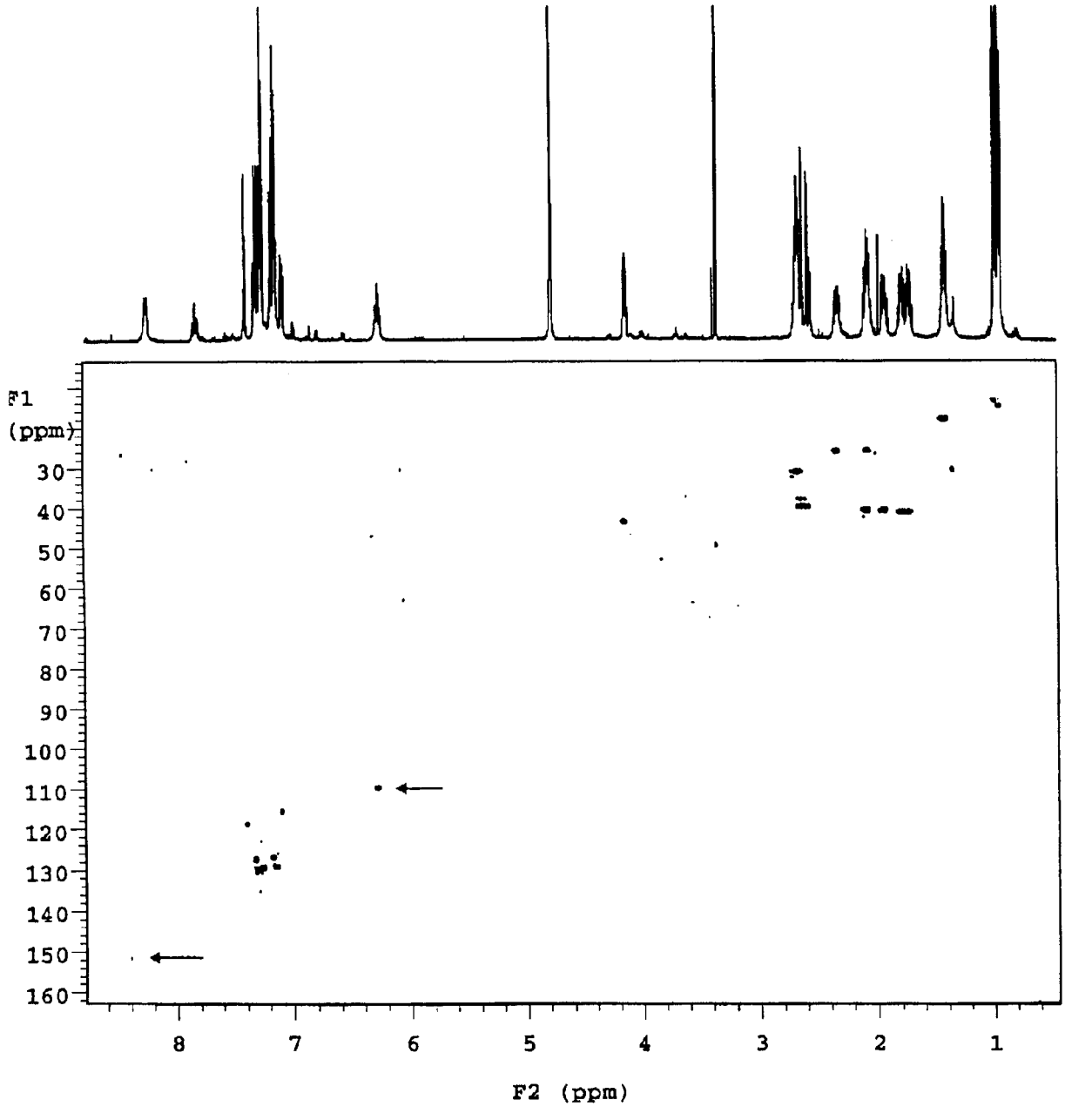


Figure 2. GHSQC spectrum of ~250 µg of **3** dissolved in 30 µl perdeuteromethanol. The data were acquired in 1 hour 5 minutes using a Varian INOVA 600 equipped with a Nalorac SMIDG-600-1.7 nmr probe. The experiment was optimized for an assumed one-bond coupling of 140 Hz.

to the carbon resonating at 106.2 ppm (see Figure 5) confirms the tautomerically averaged nature of the five carbon bridge linking the two halves of the molecule. The proton chemical shifts of **3** (in perdeuteromethanol) are also fully consistent with the data reported by Marvel, Li, and Paik (in hexadeuterodimethyl sulfoxide) [9], as shown by **4** in Scheme II.

Following the establishment of the structure of **3**, as described above, it was possible to account for virtually all of the fragment ions in the collisionally induced decomposions product ion mass spectrum shown in Scheme III.

Under acidic conditions, Marvel, Li, and Paik [10] have shown that Zincke salts such as **4** decompose to afford pyridine and the corresponding aniline. Earlier, Marvel, Caple, and Shahidi [11] demonstrated similar behavior in basic media. These observations on the instability of Zincke salts are consistent with the observed instability of **3** in solution as noted above in the introduction.

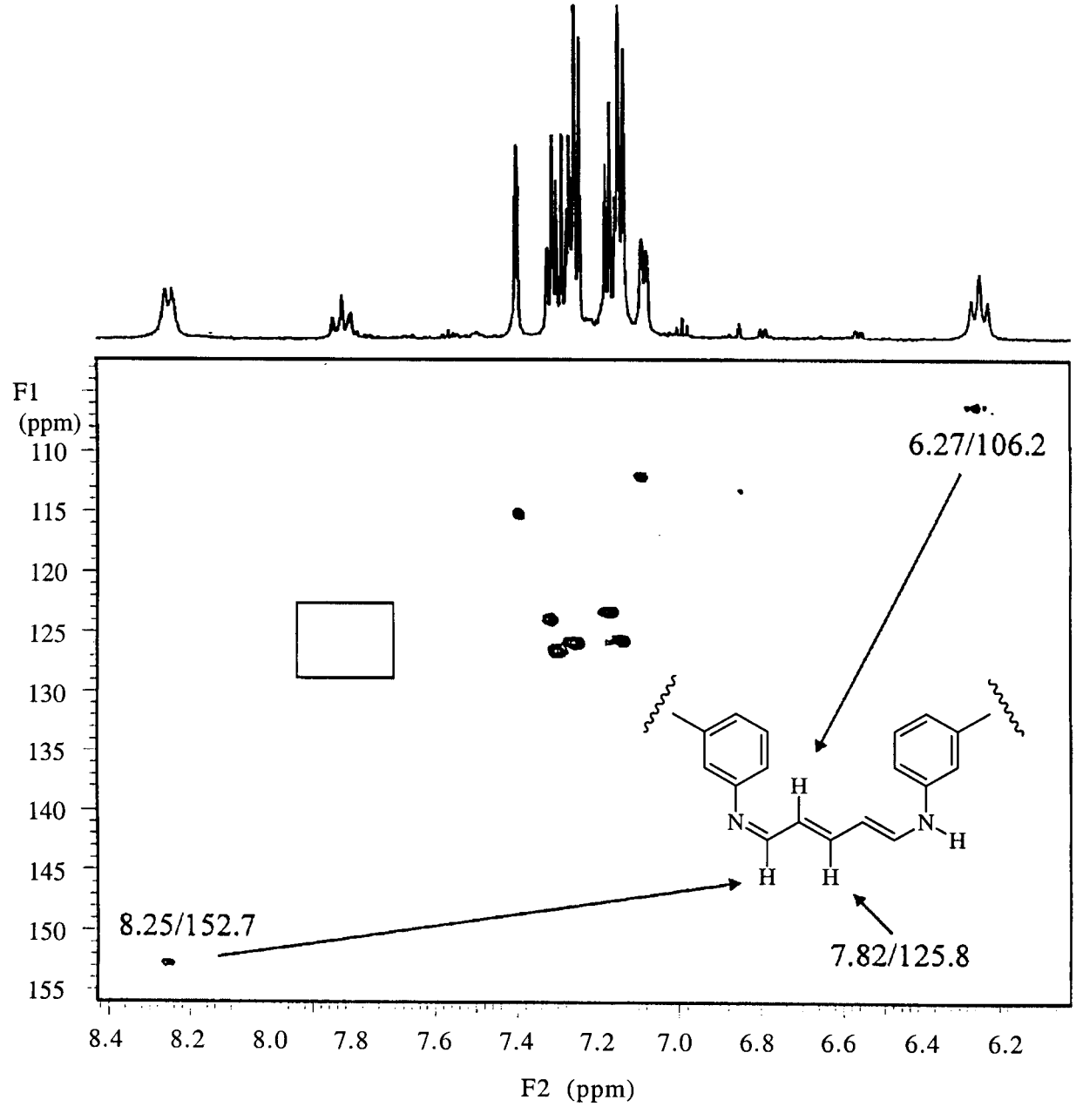


Figure 3. Expansion of the aromatic region of a GHSQC spectrum optimized for an assumed 160 Hz one-bond coupling. The data were acquired in 2 hours 9 minutes using a Varian INOVA 600 equipped with a Nalorac SMIDG-600-1.7 mm nmr probe. The sample was identical to that described above. The boxed region devoid of a response is shown to indicate the location of the 7.82/125.8 ppm direct response for the central carbon of the C_5H_5 subunit. This response was unobservable in both GHSQC spectra but was located from the 10 Hz optimized GHMBC data presented in Figures 4 and 5 below.

Finally, there are several ir spectral features of the impurity, **3**, that vary from the data for the bulk drug, Tripranavir (**2**). There is an increase in the intensity of the ir envelope at 1636 cm^{-1} observed in the ir spectrum of the impurity consistent with the increased double bond character of the impurity. There is an intense ir band in the impurity spectrum at 1547 cm^{-1} consistent with C=N stretching that is not present in the spectrum of **2**.

Relative to the parent drug, there is also an increase in the intensity of the broad manifold of vibrations centered in the 1170 cm^{-1} consistent with and increase in C-O stretching and C-OH bending of alcohol groups that can be accounted for by the components of the second δ-lactone moiety contained in the impurity structure relative to the parent drug. Finally, the impurity, **2**, exhibited a maximal absorbance at 487 nm, which is

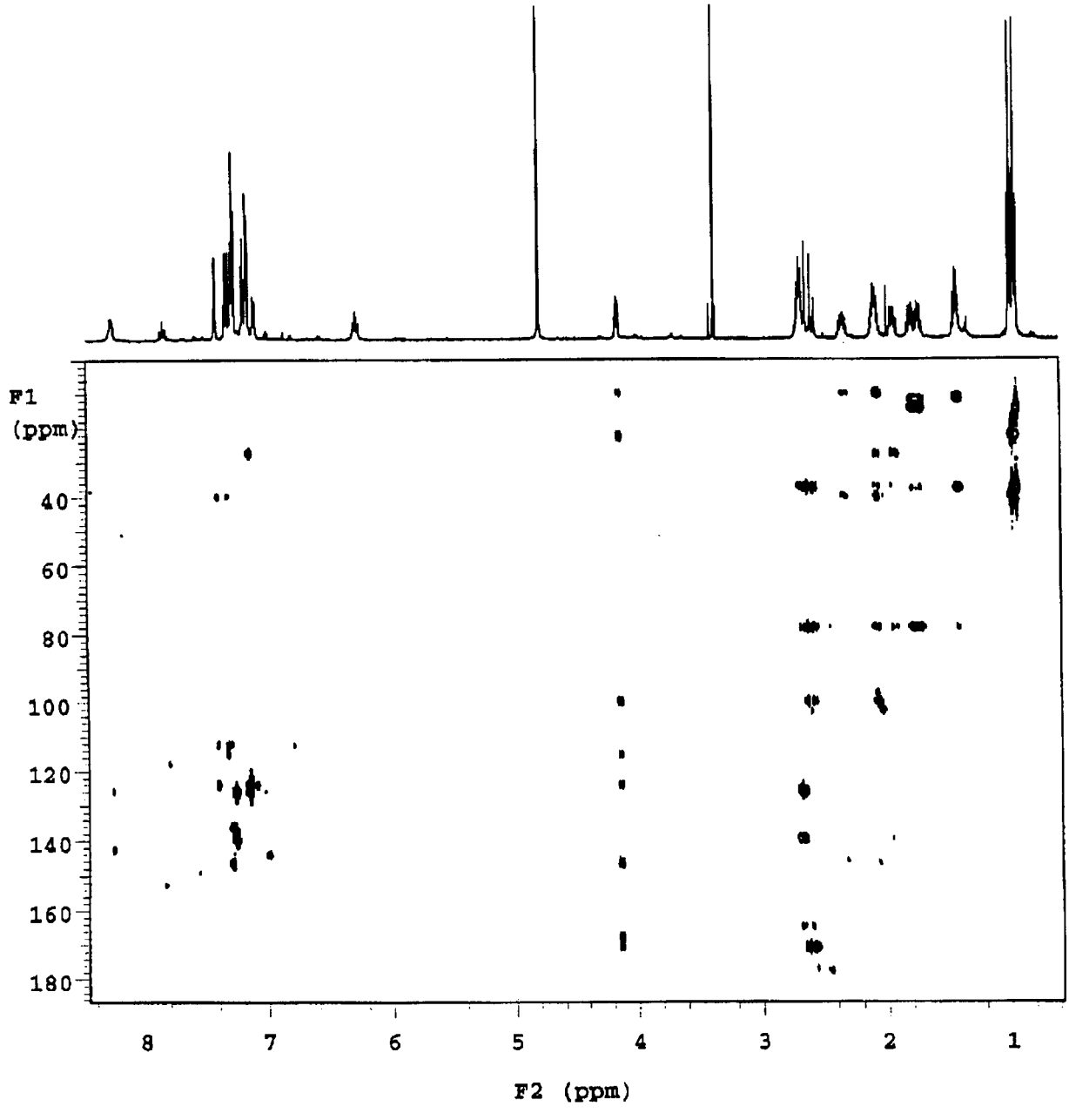


Figure 4. GHMBC experiment optimized for an assumed 10 Hz long-range coupling. The data were acquired in 12 hours 30 minutes using a sample prepared from approximately 250 µg of **3** dissolved in 30 µl perdeuteromethanol. The expansion of the aromatic region is shown in Figure 5.

also consistent with the data for the protonated form of various Zincke salts [10].

Conclusion.

In conclusion, it was possible to fully characterize a highly colored, trace impurity present at <<0.1% in the initial commercial scale synthesis of tipranavir (**2**, PNU-140690). Once the structure of the impurity was determined, efforts to preclude the formation of **3** during the synthetic production of bulk quantities of tipranavir (**2**, PNU-140690) were undertaken. The utilization of 600 MHz SMIDG nmr probe technology was pivotal in our ability to acquire the necessary homo- and heteronuclear nmr shift correlation data to establish the structure of **3** because of the inherent instability of this molecule in solution. It is likely that an attempt to identify the structure of this impurity using conventional nmr methodology would have been difficult at best.

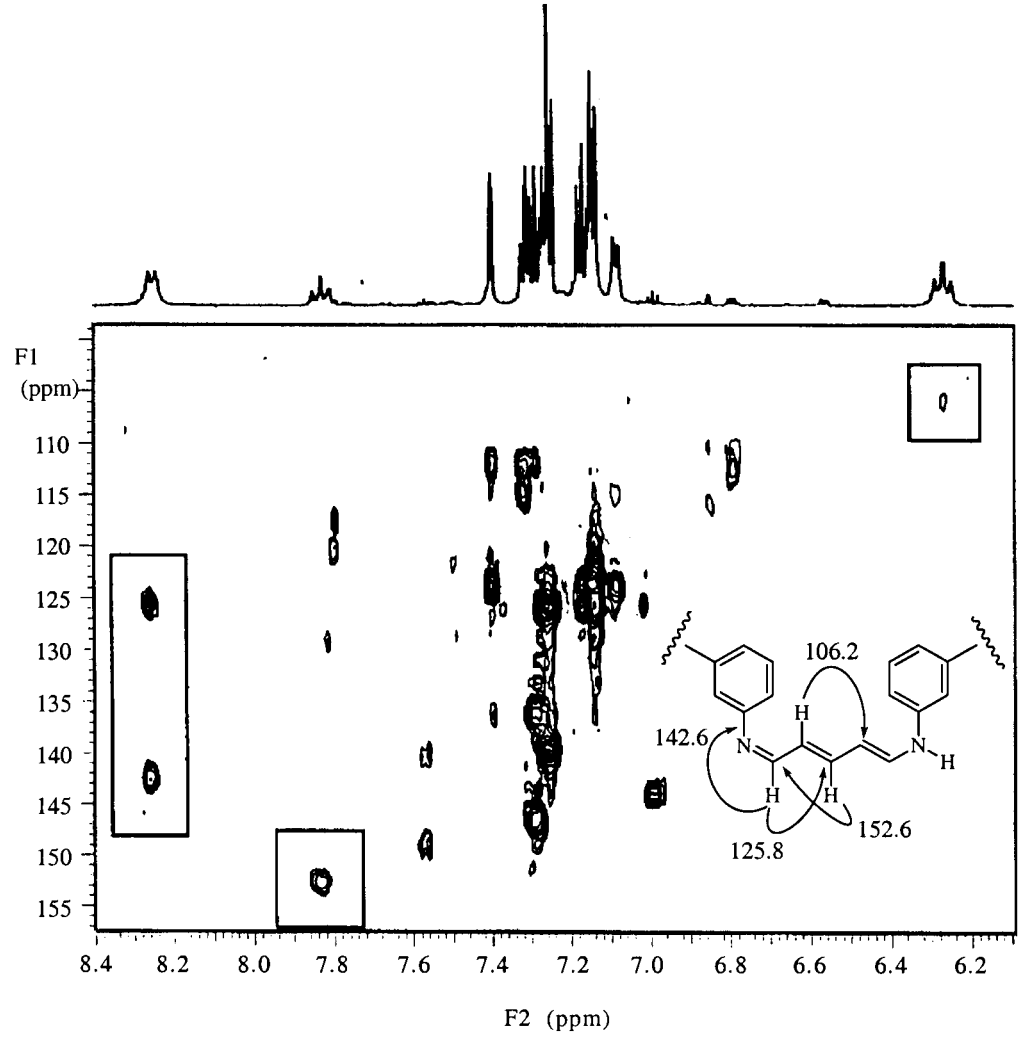
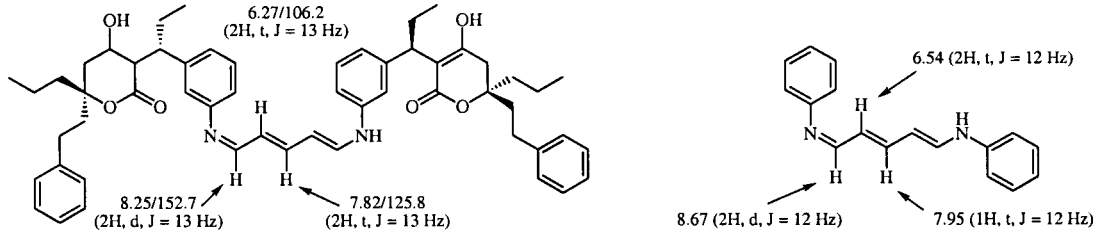


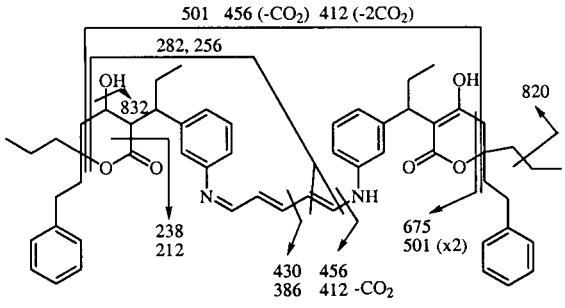
Figure 5. Expansion of the aromatic region of the 10 Hz optimized GHMBC experiment shown in Figure 4. The long-range couplings from the proton resonating at 8.25 ppm correlate this resonance to the quaternary carbon of the 1,3-disubstituted ring as well as to the central carbon of the five carbon bridge between the two halves of the molecule, which resonates at 125.8 ppm. The correlation from the proton resonating at 7.82 ppm redundantly confirms its position relative to the imino carbon resonating at 152.6 ppm. As will be noted from the contour plot, significant degradation occurred during the overnight acquisition of the GHMBC data. The correlations to the weak impurity peaks in the reference spectrum at ~6.8 ppm near 110 ppm in F_1 are an example (the reference spectrum was acquired prior to the overnight acquisition, these impurity peaks were significantly larger after 2D nmr data acquisition). Despite the degradation of the sample, the data still allowed the successful confirmation of the structure.PP

Scheme II



3

Scheme III



REFERENCES AND NOTES

[1] T. M. Judge, G. Phillips, J. K. Morris, K. D. Lovasz, K. R. Romines, G. P. Luke, J. Tulinsky, J. M. Tustin, R. A. Chrusciel, L. A. Dolak, S. A. Mizsak, W. Watt, J. Morris, S. L. Vander Velde, J. W. Strohbach, and R. B. Gammill, *J. Am. Chem. Soc.*, **119**, 3627 (1997).

[2] S. Thaisrivongs, H. I. Skulnick, S. R. Turner, J. W. Strohbach, R. A. Tommasi, P. D. Johnson, P. A. Aristoff, T. M. Judge, R. B. Gammill, J. K. Morris. K. R. Romines, R. A. Crusciel, R. R. Hinshaw, K.-T. Chong, W. G. Tarpley, S. M. Pope, D. E. Slade, J. C. Lynn, M.-M. Horng, P. K. Tomich, E. P. Seest, L. A. Dolak, W. J. Howe, G. H. Howard, F. J. Schwene, L. N. Toth, G. E. Padbury, G. J. Wilson, L. Shiou, G. L. Zipp, K. F. Wilkinson, B. D. Rush, M. J. Ruwart, K. A. Koeplinger, Z. Zhao, S. Cole, R. M. Zaya, T. J. Kauk, M. N. Janakiraman, and K. D. Watenpaugh, *J. Med. Chem.*, **39**, 4349 (1996).

[3] S. R. Turner, J. W. Strolibach, R. A. Tommasi, P. A. Aristoff, P. D. Johnson, H. I. Skulnick, L. A. Dolak, E. P. Seest, P. K. Tomich, M. J. Bohanon, M.-M. Horng, J. C. Lynn, K.-T. Chong, R. R. Hinshaw, K. D. Watenpaugh, M. N. Janakiraman, and S. Thaisrivongs, *J. Med. Chem.*, **41**, 3407 (1998).

[4] G. E. Martin, R. C. Crouch, and A. P. Zens, *Magn. Reson. Chem.*, **36**, 551 (1998).

[5] G. E. Martin. J. E. Guido, R. H. Robins, M. H. M. Sharaf, P. L. Schiff, Jr., and A. N. Tackie, *J. Nat. Prod.*, **61**, 555 (1998).

[6] G. E. Martin, C. E. Hadden, A. N. Tackle, M. H. M. Sharaf, and P. L. Schiff, Jr., *Magn. Reson. Chem.*, **37**, 529 (1999).

[7] T. Zincke, *Liebigs Ann. Chem.*, **330**, 361 (1903).

[8] T. Zincke, *Liebigs Ann. Chem.*, **333**, 296 (1904).

[9] R. H. Robins, U. S. Patent No. 5,235,186; assigned to Finnigan, Inc., San Jose, CA.

[10] E. N. Marvell, T. H. Li, and C. Paik, *Tetrahedron Letters*, 2089 (1973).

[11] E. N. Marvell, G. Caple, and I. Shahidi, *J. Am. Chem. Soc.*, **92**, 5641 (1970).