# Identification of an Impurity in Synthetically Prepared GAZT: 4″,5″-Dehydro-GAZT



Cecilia S. Koble, William R. Hall, Virgil L. Styles, Janet L. Rideout, Doug Minick, James Cichetti, Robert L. Johnson, Ronald C. Crouch and Gary E. Martin*

Division of Organic Chemistry, Burroughs Wellcome Company, Research Triangle Park, NC 27709

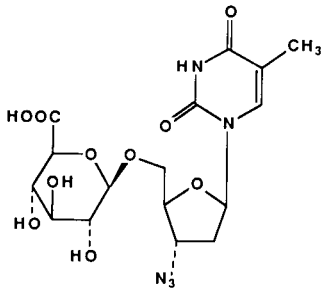
Received May 21, 1991

A minor impurity was recorded by hplc during the final purification of synthetically prepared samples of the 5′-*O*-glucuronide of azidothymidine (GAZT). Following isolation of a small quantity of this impurity by hplc, elucidation of the structure was undertaken. A combination of two-dimensional nmr methods in conjunction with infrared spectroscopy and mass spectrometry were utilized to unequivocally identify the impurity as the 4″,5″-dehydro-5′-*O*-glucuronide of AZT (DGAZT).

*J. Heterocyclic Chem.*, **28**, 1795 (1991).

## Introduction.

AZT (3′-azido-3′-deoxythymidine) is currently the only approved therapeutic agent for the treatment of acquired immune deficiency syndrome (AIDS) in humans. A major metabolite, the 5′-*O*-glucuronide of AZT (**1**), isolated from the urine of a variety of experimental animals and humans [1] was recently described.




1

Although GAZT can be prepared by the incubation of AZT with UDPGA in the presence of microsomes prepared from human cadaver liver, preparation using commercially available rabbit or bovine liver UDPGTases was unsuccessful. Thus, synthetic methods have been developed to provide necessary quantities of GAZT for metabolic studies. During the final purification of synthetically prepared GAZT, a small quantity (<5%) of an impurity was recorded by hplc. This material was isolated by analytical and semi-preparative hplc using multiple chromatographic separations with combination of the desired eluant fractions. Approximately 1 mg of material was isolated for the initial structural studies described below.

## Results and Discussion.

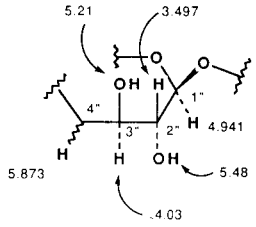
Using a thin film deposited on a non-absorbing matrix, an FT-IR study of the isolated impurity showed an absorbance at 2111 $cm^{-1}$, which confirmed the presence of the 3′-azido group. A broad absorbance from 2300-3000 $cm^{-1}$ was attributed to a strongly hydrogen-bonded carboxyl group. No specific information on the environment of the carboxyl group could be deduced from the infrared data, however, because the carbonyl stretching bands were comingled with the strong C=O stretching vibrations from the thymidine. Thus, from the infrared data it appeared that the AZT nucleus was intact and coupled to at least some portion of glucuronic acid.

To establish the molecular weight of the GAZT impurity, FAB mass spectra were obtained from samples dissolved in water using glycerol as the FAB matrix. Ions were observed at m/z values 448 and 470 and were initially assigned to $(M+H)^+$ and $(M+Na)^+$, respectively; however, no structure could be proposed for a compound of molecular weight 447 Da, which was consistent with the infrared and initial proton nmr data. A thermospray mass spectrum of the impurity did give an ion at m/z 426, which was ascribed to the protonated molecular ion of the impurity. This molecular weight was subsequently confirmed by a FAB mass spectrum (Figure 1) obtained for a sample of the impurity dissolved in $d_6$-DMSO that was used for the nmr study. The molecular weight of 425 Da for the impurity is consistent with a loss of 18 Da from GAZT. Ultimately, the ions observed at m/z 448 and 470 were assigned as $(M+Na)^+$ and $(M-H+2Na)^+$, respectively.

No reason can be given for the difference in FAB mass spectra of this sample when obtained from an aqueous *versus* a DMSO solution. If the compound were the sodium salt of an acid, then ions at 448 and 470 Da should be observed. But this should also occur when either the DMSO or water solutions of the samples were used to obtain FAB mass spectra. Furthermore, the thermospray mass spectrum of the water solution should exhibit the ion at m/z 448 if the sample were the sodium salt, and it did not. This behavior has not been observed for GAZT or other glucuronide conjugates of nucleoside analogues and cannot be readily explained.

The 500 MHz proton nmr spectrum of the impurity in $d_6$-DMSO is shown in Figure 2. COSY and RCOSY spectra

of the material are shown in Figure 3 and allowed the confirmation of the structure of the 3′-azido-2′,3′-dideoxysugar. The concerted interpretation of proton-proton connectivity information in the COSY and 60 msec RCOSY spectra established the structural fragment shown by **2**.




**2**

Further the elucidation of the impurity's structure required the acquisition of heteronuclear chemical shift correlation data. Because a limited quantity of material was available, these data could only be acquired using inverse-detected methods. The HMQC spectrum [2] of the impurity is shown in Figure 4. Chemical shifts were observed that corresponded favorably to the thymidine and the 3′-azido-dideoxysugar reported for GAZT [1]. Protonated carbon chemical shifts for the structural fragment depicted by **2** were as follows: 100.4 ppm for the anomeric center, C1″; 70.7, 66.8 and 109.6 ppm for C2″, C3″ and C4″, respectively. The chemical shift observed for the C4″ resonance was considerably downfield relative to the corresponding position of GAZT, which resonated at 71.3 ppm [1]. Hence it was logical to conclude that the impurity differed structurally from GAZT beginning with the 4″-position.

Further purifications supplied an additional 2 mg of the impurity (3 mg total), which was sufficient for the overnight acquisition of an inverse-detected long-range heteronuclear chemical (HMBC) shift correlation spectrum [3]. The HMBC spectrum allowed complete identification of the impurity.

The proton resonating at 4.941 ppm, assigned as the 1″ anomeric proton, gave connectivities in the HMBC spectrum to carbons resonating at 70.7, 66.8 and 140.32 ppm. The first two carbons are the 2″- and 3″-carbon resonances, respectively, and confirmed assignments made from the COSY/RCOSY and HMQC data. The remaining connectivity to the carbon resonating at 140.32 ppm suggested bonding to a quaternary carbon. Furthermore, since the 4.941 ppm resonance is an anomeric proton, this connectivity included an oxygen atom, which in this case was the normal pyran-ether oxygen, linking the 1″- and 5″-positions of glucuronic acid. The proton at the 4″-position exhibited long-range connectivities in the HMBC spectrum to carbons resonating at 70.7, 140.3 and 163.0 ppm. The first of these connectivities was to the 2″-position. The second connectivity was to the quaternary carbon, which has already been shown to long-range correlate to the H1″ proton. The last carbon's chemical shift was reasonable for the normal 6″-carboxyl group. Given these observations, we interpreted the chemical shift of

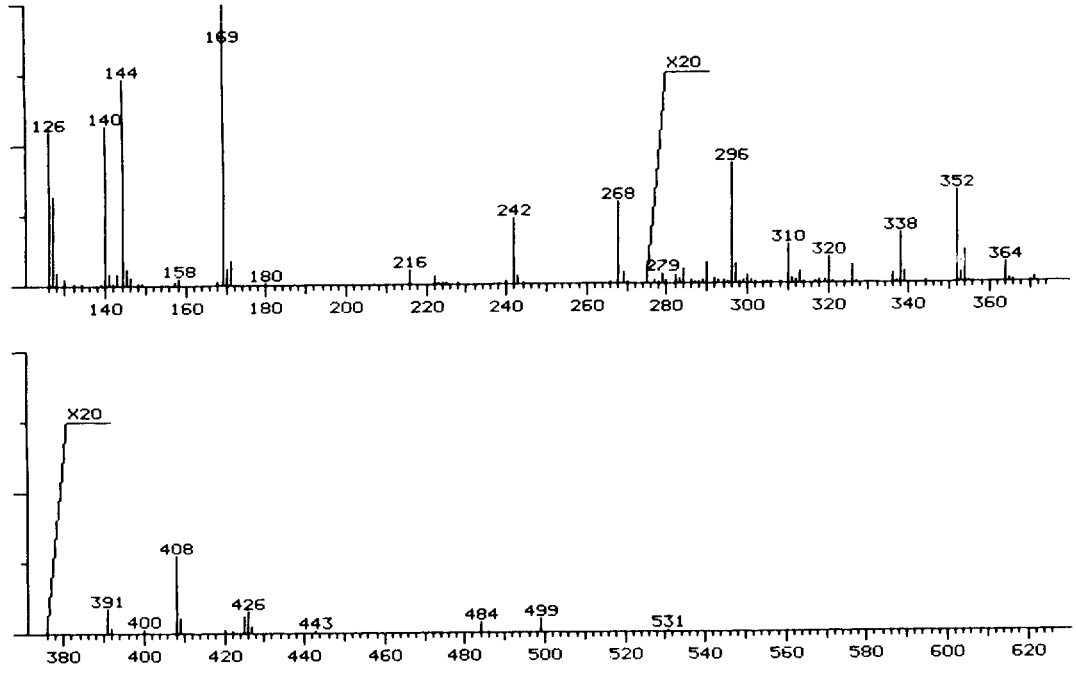
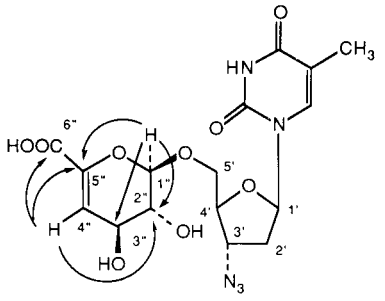



Figure 1. Thermospray LC mass spectrum (glycerol *matrix*) of the synthetic impurity iolated from the synthesis of GAZT. The sample used was taken from the $d_6$-DMSO solution used for nmr data acquisition.

the C4″ as vinylic, arising *via* dehydration, and giving a 4″,5″-dehydroglucuronic acid moiety. The connectivities just described support the presence of the 4″,5″-dehydroglucuronic acid moiety shown in structure **3**. The structure of the impurity was simply the dehydro-derivative of GAZT.




3

## Conclusions.

The structure of the impurity observed in synthetically prepared GAZT has been proven to be 4″,5″-dehydro-GAZT (DGAZT). It remains to be determined if this impurity arises during the final stages of the synthesis or perhaps during the isolation and purification of GAZT itself. Alternatively, it is also possible that the impurity could arise from dehydration during the preparation of the per-

Table 1

Comparison of the Proton and Carbon NMR Chemical Shift Data of GAZT ($d_4$-Methanol) [1] and DGAZT ($d_6$-DMSO) at 500 MHz

| Position | GAZT ($d_4$-Methanol) | | DGAZT ($d_6$-DMSO) | |
|---|---|---|---|---|
| | Proton | Carbon | Proton | Carbon |
| 4 | | 163.50 | | 163.56 |
| 6" | | 170.01 | | 162.98 |
| 2 | | | | 150.32 |
| 5" | | 75.34 | | 140.23 |
| 6 | 7.766 | 150.20 | 7.573 | 135.82 |
| 5 | | 109.53 | | 113.12 |
| 4" | 3.558 | 71.27 | 5.873 | 109.62 |
| 1" | 4.426 | 102.99 | 4.941 | 100.42 |
| 1' | 6.186 | 83.62 | 6.131 | 83.46 |
| 4' | 4.125 | 83.29 | ~4.05 [a] | 81.97 |
| 2" | 3.255 | 73.20 | 3.497 | 70.67 |
| 5'α | 4.198 | 68.70 | 4.004 | 68.88 |
| 5'β | 3.773 | | 3.775 | |
| 3" | 3.403 | 75.44 | ~4.03 [a] | 66.78 |
| 3' | 4.525 | 60.74 | 4.582 | 60.75 |
| 2'β | 2.420 | 36.27 | 2.434 | 35.83 |
| 2'α | 2.362 | | 2.254 | |
| $-CH_3$ | 1.907 | 12.04 | 1.774 | 11.88 |
| 2"-OH | | | 5.48 | |
| 3"-OH | | | 5.21 | |

[a] These chemical shifts are approximate since the involved resonances were overlapped.

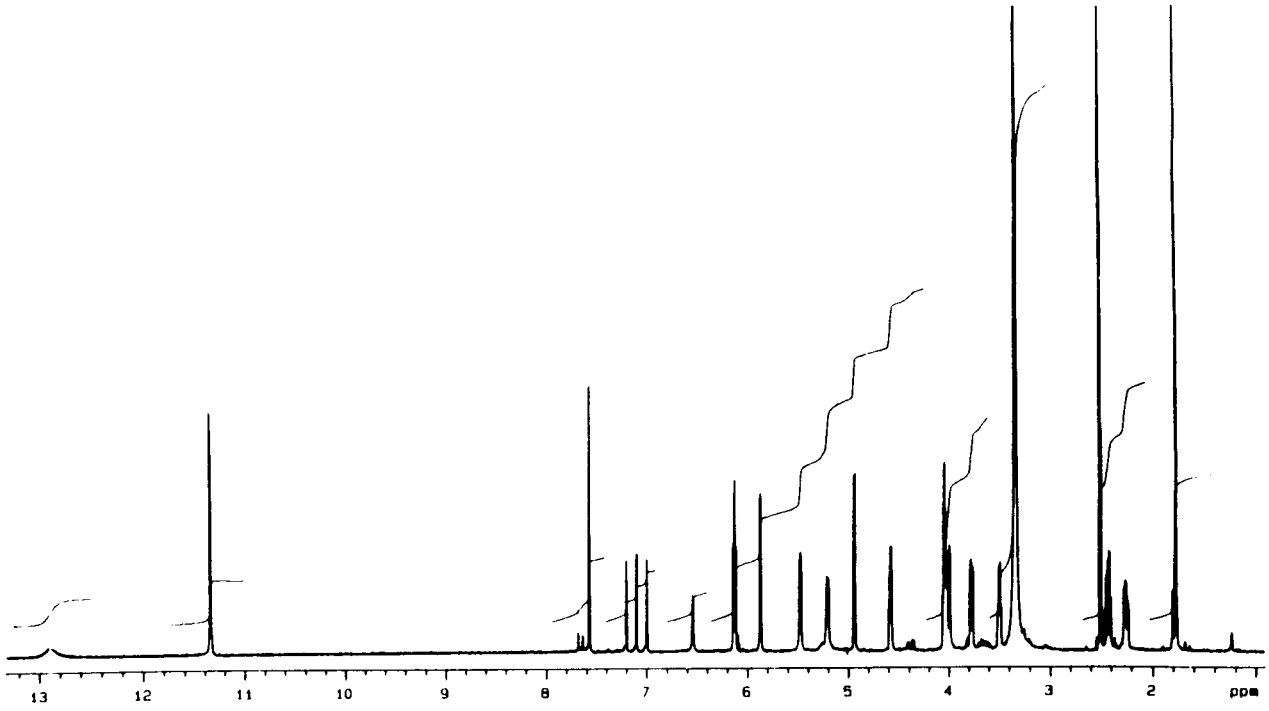



Figure 2. Proton nmr spectrum of DGAZT isolated as an impurity from synthetically prepared GAZT. The spectrum was recorded at an observation frequency of 500 MHz in $d_6$-DMSO.

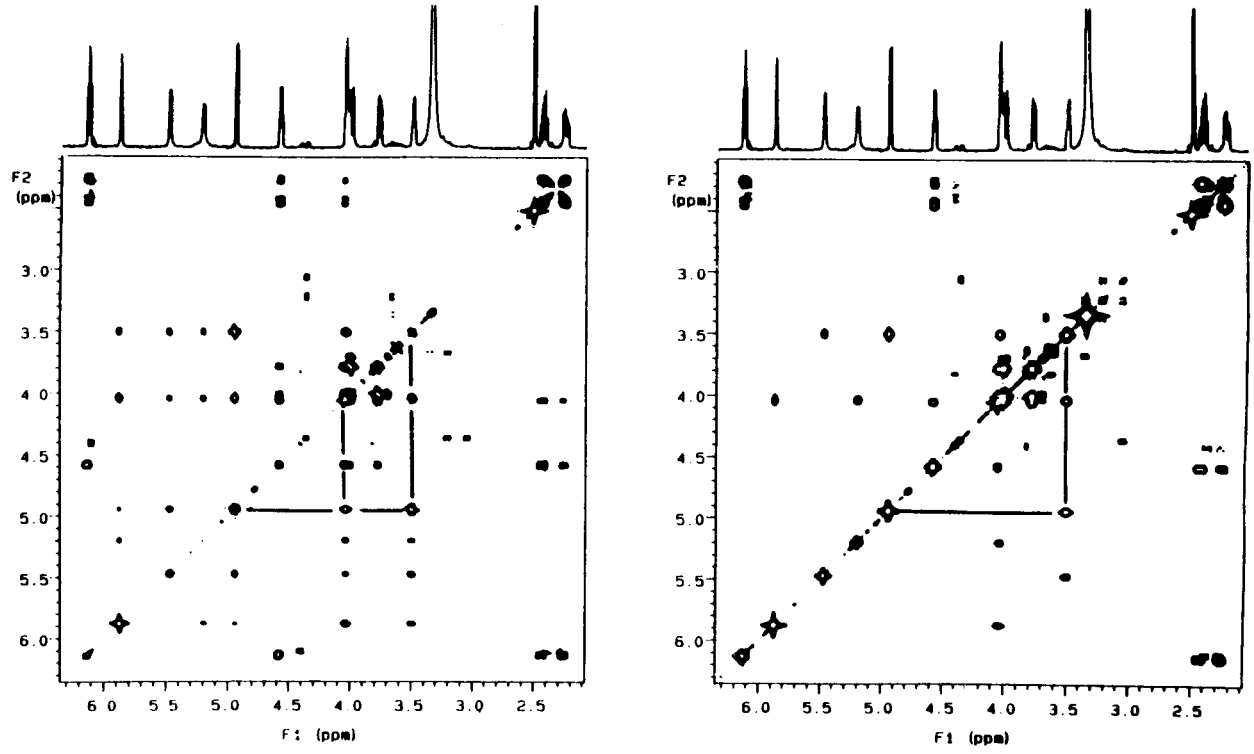



Figure 3. COSY and RCOSY spectra of DGAZT. The COSY spectrum is shown to the right; the RCOSY spectrum is shown to the left and was acquired with a mixing (relay) time of 60 msec.

acetylated 1-bromoglucosyl synthon used in the synthesis. The corresponding dehydrated peracetyl impurity could then be carried along through the synthesis ultimately giving DGAZT. Proton and carbon chemical shift data for GAZT [1] and **3** are shown comparatively in Table 1.

## EXPERIMENTAL

DGAZT was isolated from synthetic GAZT by preparative high-pressure chromatography on a Hamilton PRP-1 column (21.5 x 250 mm, 10$\mu$ particle) eluted with 8% acetonitrile in water at 10 ml/min. Analytical chromatography was performed using a Supelco LC-8 column (4.6 x 150 mm, 5$\mu$ particle) eluted with 15% methanol in water (0.1% trifluoroacetic acid and 0.1% triethylamine by volume in each solvent) at 1 ml/min.

The FT-IR spectra were recorded using an Analect Fourier transform infrared instrument equipped with a narrow-band high-sensitivity MCT detector. An infrared reflectance spectrum of DGAZT was recorded at a resolution of 2 wavenumbers.

Fast atom bombardment (FAB) mass spectra were obtained using a VG 70SQ mass spectrometer equipped with a VG 11-250J data system. Data were analyzed using a Kratos Mach 3 work station. The FAB spectra were obtained using the following conditions: accelerating voltage 7 KV; FAB gun voltage 7 KV; FAB gun current 1 milliamp; FAB gas - xenon; FAB matrix - glycerol.

Thermospray lc/ms was run using a Nermag R 10-10 quadrupole mass spectrometer equipped with a Vestec Thermospray interface from Nermag. A Waters U6K injector was used to introduce 10 $\mu$l of a methanol solution of DGAZT under the following conditions: mobile phase 90% 0.05*M* ammonium acetate, 10% acetonitrile; flow rate 1 ml/min; interface tip temperature 230°; source block temperature 236°. The thermospray lc/ms spectra were analyzed using the Nermag R 10-10 data system.

The nmr spectra recorded for DGAZT were obtained by dissolving the available material in 0.8 ml of 100% $d_6$-DMSO (Merck) and filtering through glass wool into a 5 mm nmr tube. All spectra were recorded using a Varian VXR-500S spectrometer operating at an observation frequency of 499.843 MHz for proton observation and equipped with a Z-Spec 5 mm inverse-detection probe model ID500 obtained from Nalorac Cryogenics Corp, Martinez, CA.

The proton reference spectrum was recorded using a spectral width of 3671 Hz digitized with 32K points, giving an acquisition time of 2.37 sec with an interpulse delay of 1 sec. The 90° proton pulse width was 14.5 $\mu$sec. A carbon spectrum was recorded at 125.697 MHz digitized by 32K points with a spectral width of 25 KHz, giving an acquisition time of 1 sec followed by a 1 sec interpulse delay. A total of 33972 transients were accumulated. The COSY spectrum was recorded as 1024 x 512 points with 4 transients/$t_1$ increment. The data were zero-filled to 2048 x 2048 points during processing and were symmetrized prior to plotting. The RCOSY spectrum was acquired and processed in the same fashion and differed only by a 60 msec fixed delay used to relay magnetization. The HMQC spectrum was acquired using the se-

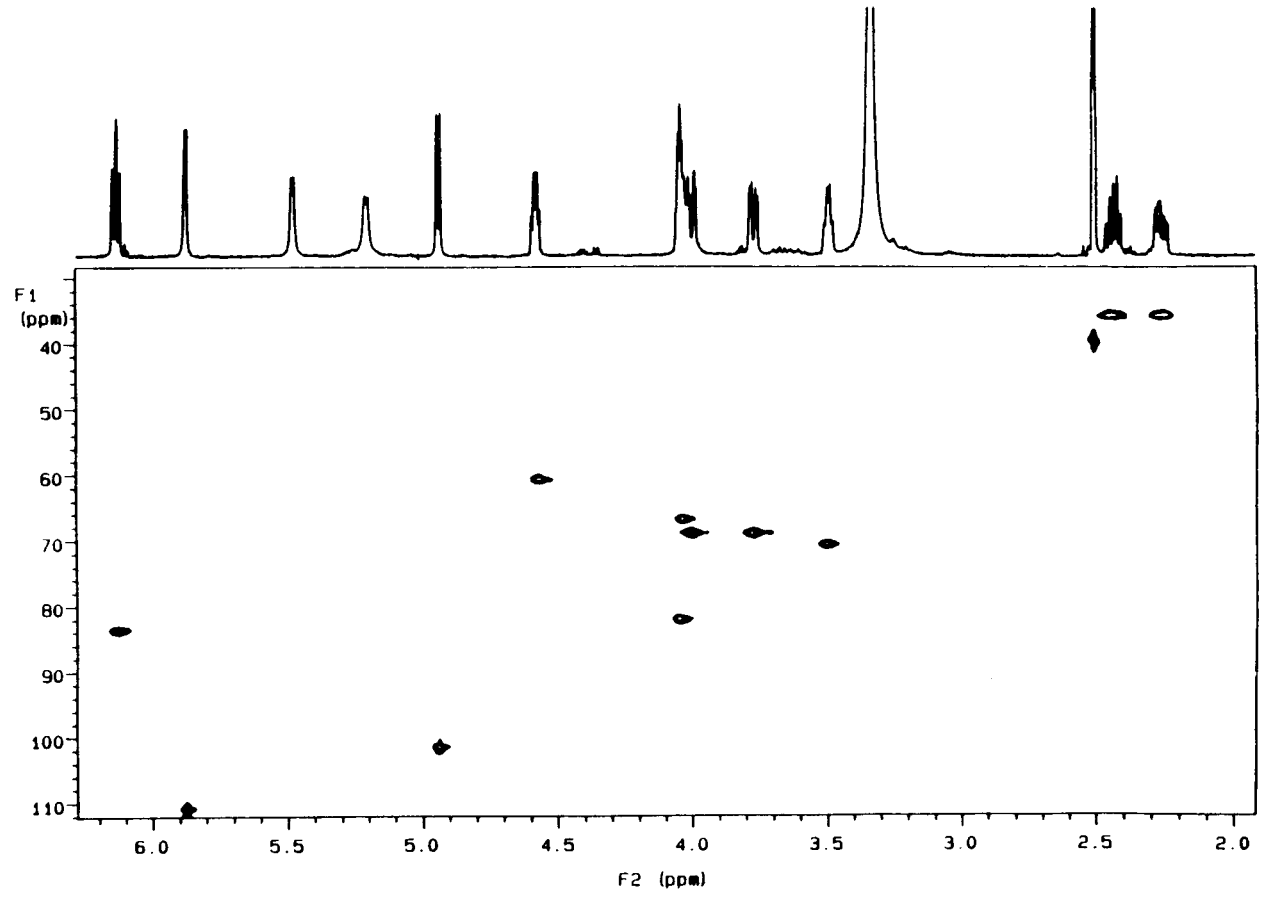



Figure 4. Heteronuclear chemical shift correlation (*via* $^{1}J_{CH}$) spectrum of DGAZT in $d_6$-DMSO recorded at 500 MHz using the HMQC pulse sequence of Bax and Subramanian [2].

quence of Bax and Subramanian [2] on a 1 mg sample of material. The data were taken as 2048 x 256 points with 128 transients/increment. Spectral widths were 3617 and 25139 Hz in $F_2$ and $F_1$, respectively. The 90° proton pulse width was 14.5 μsec; the 90° $^{13}C$ pulse width was 13.7 μsec. Broadband $^{13}C$ decoupling during acquisition was applied using WALTZ with $\gamma H_2/2\pi$ = 27.2 KHz. The one-bond heteronuclear coupling was optimized for an assumed average 140 Hz. The interpulse delay was 1 sec, and the null interval following the BIRD pulse at the beginning of the sequence was 300 msec. Data accumulation required 13.7 hours. The HMBC spectrum of DGAZT was recorded using the sequence of Bax and Summers [3] on a 3 mg sample of material. The parameters were largely the same as those used for the HMQC spectrum. The long-range delay was optimized for 8 Hz (62.5 msec). The interpulse delay was set to 1.8 sec and a total of 128 transients were accumulated/increment giving a total acquisition time of 17 hours.

## REFERENCES AND NOTES

[1] S. S. Good, C. S. Koble, R. Crouch, R. L. Johnson, J. L. Rideout and P. de Miranda, *Drug Metab. Dispos.*, **18**, 321 (1990).

[2] A. Bax and S. Subramanian, *J. Magn. Reson.*, **67**, 565 (1986).

[3] A. Bax and M. F. Summers, *J. Am. Chem. Soc.*, **108**, 2093 (1986).