Identification of Degradants of a Complex Alkaloid Using NMR Cryoprobe Technology and ACD/Structure Elucidator

Gary E. Martin,* Chad E. Hadden, David J. Russell, Brian D. Kaluzny, Jane E. Guido, Wayne K. Duholke, Bruce A. Stiemsma, and Thomas J. Thamann

Rapid Structure Characterization Group, Global Pharmaceutical Sciences, Pharmacia Corp. Kalamazoo, Michigan 49001-0199

Ronald C. Crouch

NMR Applications Laboratory, Varian Inc., Palo Alto, California 94303

Kirill Blinov, Mikhail Elyashberg, Eduard R. Martirosian

Advanced Chemistry Development, Moscow 117513, Russian Federation

Sergey G. Molodtsov

Novosibirsk Institute of Organic Chemistry, Siberian Branch of Russian Academy of Science, Lavrentiev Avenue 9, Novosibirsk 630090, Russia

Antony J. Williams

Advanced Chemistry Development, Toronto, Ontario, M5H 3V9, Canada

Paul L. Schiff, Jr.

Department of Pharmaceutical Sciences, School of Pharmacy, University of Pittsburgh, Pittsburgh, PA 15261

Received June 6, 2002

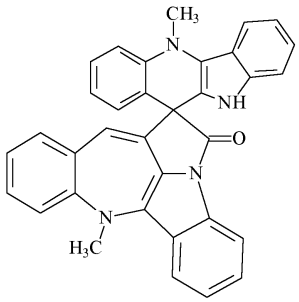
Identification of degradants of pharmaceuticals is a necessary challenge of the drug development process following the subjection of candidate molecules to a variety of physico-chemical stresses. It would be desirable to be able to conduct such studies on a minimal amount of material. As a prototypical study, the isolation and identification of degradants of a sample of the complex indoloquinoline alkaloid, cryptospirolepine, was undertaken after prolonged storage in DMSO solution using a combination of cryogenic NMR probe technology and CASE (Computer-Assisted Structure Elucidation) programs. None of the starting alkaloid remained after storage; a chromatogram of the DMSO solution demonstrated the presence of >25 components in the mixture. The two most abundant degradation products were identified as the known alkaloid cryptolepinone (~35%) and an unprecedented rearrangement product, DP-2, (~16%).

J. Heterocyclic Chem., **39**, 1241 (2002).

Introduction.

Structural characterization of degradants of pharmaceuticals is a necessary challenge of the drug development process. Candidate molecules are routinely subject to physicochemical stress challenges of various types to understand their stability and degradation behavior. Ideally, to facilitate this process, it would be desirable to be able to conduct such studies on minimal amounts of material, which imposes stringent sensitivity demands on the NMR spectrometer used in the work, mandating the use of small volume, high sensitivity NMR probes or cryogenic NMR probe technology. Mass spectrometric sensitivity is seldom an issue and should be an integral part of the structure characterization process. In addition, to expedite the process, it would also be highly desirable to be able to employ computer-assisted structure elucidation (CASE systems) to assist the investigator.

As a model study, a 2.5 mg sample of the complex spiro nonacyclic alkaloid cryptospirolepine (**1**) that had been stored in a sealed 5 mm NMR tube in d_6-DMSO for a prolonged period (~10 years) was examined to evaluate the combined utilization of cryogenic NMR probe technology and computer-assisted structure elucidation in the characterization of unknown degradants of a complex molecule. [1]



1

Computer-Assisted Structure Elucidation (CASE).

Computer-Assisted Structure Elucidation (CASE) programs are a viable means to assist even highly competent NMR spectroscopists to elucidate complex chemical structures where there are 2D NMR data in a more efficient manner. There have been a number of reports in the literature that have described expert systems [2-11] intended for this purpose. In these systems, 2D NMR data are presented in the form of atom-to-atom connectivities between atoms of the molecule, which serve as restrictions in the structure generation process in accord with the given molecular formula. Typical data sets may be comprised of: homonuclear COSY or TOCSY data; NOESY or ROESY correlation data; direct heteronuclear shift correlation spectra such as ^{1}H-^{13}C HMQC or HSQC, or more recently ^{1}H-^{15}N direct correlation spectra; and lastly long-range heteronuclear shift correlation data. The latter, until a few years ago, were exclusively restricted to statically-optimized ^{1}H-^{13}C HMBC data. More recently, long-range ^{1}H-^{15}N experiments at natural abundance have received considerable attention, forming the topic of recent, comprehensive reviews [12]. In addition, the development of numerous, accordion-optimized long-range heteronuclear shift correlation experiments has been reported and is the topic of a recent review [13]. Regardless of the ensemble of 2D NMR data acquired and utilized, the preferred approach has been based on the generation of structures based on the prediction of ^{13}C NMR shifts of candidate molecules consistent with the atom-to-atom connectivity information extracted from the various 2D NMR data sets.

The experience gained with the development and utilization of expert systems has shown that the best results can be achieved when there is a close, and highly synergistic interaction between the spectroscopist and the computer program being used. Ideally, a competent spectroscopist will be allowed the broad possibility of using his/her experience and knowledge about properties of the molecule being studied to impose additional restrictions for the types of structures generated. Quite simply, if the spectroscopist knows that the molecule being structurally interrogated with his/her NMR experiments is a steroid, there is little reason to allow a CASE program the freedom of generating inappropriate and unrelated molecular structures. The successful implementation of such possibilities creates a "symbiotic" relationship between the human mind and the computer, ultimately synergizing the structure elucidation process. Unfortunately, most of the expert systems devised thus far [2-11] have rather restricted ability to apply *a priori* information, thus precluding the use of user-defined structural fragments during the elucidation process. Little attention was paid in the past to detecting contradictions in 2D data and/or in providing methods for their resolution. It is known that one frequent source of contradictions is, for example, the observation of correlations in COSY or long-range heteronuclear shift- correlation spectra that correspond to four or more bonds [14]. In general, CASE programs are "tuned" for shorter correlation pathways. Unusual long-range heteronuclear correlation pathways have recently been reviewed by Araya-Maturana and co-workers [15]. Different ways of eliminating contradictions have been proposed. One possible approach is to iteratively search for contradictions by multiple repetition of the structure generation process [14]. With each cycle, one correlation to a weak 2D peak is added to the structure generation process. If any correlation is across more bonds than the number set by default, the structure generator portion of the program will not produce any structures after such a correlation has been added, indicating the presence of contradictions. Another approach to overcoming this difficulty is by using a stochastic algorithm of structure generation that requires application of computer complexes on parallel processors [10]. The common drawback of the methods extant in the literature [2-11] is the inability of these programs of interacting with a diverse sub-structure base accompanied by related ^{13}C NMR sub-spectra. Our experience has shown that quite frequently important structural information about an unknown structure can be derived from such base.

The aforementioned drawbacks are largely overcome in the StrucEluc program [14]. This system is capable of using, in addition to 2D NMR spectral data, a library with 500,000 fragments that have ^{13}C NMR chemical shift assignments associated with them [15]. In those cases when the number of available 2D NMR correlations is insufficient to impose restrictions on the structure generation process (in this case the number of possible structures can be enormous and generation times unacceptable), the system searches for appropriate fragments in the library in accordance with their sub-spectra. Of the fragments found, the fragments meeting the restrictions arising from 2D spectra, are kept. The admissible combinations of good fragments are "projected" on the set of all atoms of a molecule. Consequently, the program builds and visualizes molecular connectivity diagrams (MCD); fragments, atoms, and connectivities of different length are graphically represented. At this point in the elucidation process, a capable spectroscopist has the opportunity to analyze MCD's and to make his/her revisions (specify carbon atoms hybridization set by the program, defining the possible chemical "neighborhood" with heteroatoms, *etc*.). It is envisioned that the chemist may also introduce fragments that in his/her opinion should be present in a molecule. Subsequently, both user fragments and those from the library are used for MCD creation.

Chemists commonly employ a strategy of using the assigned spectra of chemically related structures, whenever they're available, in deducing the structure of a new

compound. Generally this approach is rather successful. In order to implement this method within the StrucEluc program, algorithms enabling the automatic generation of the user's fragments library, have been incorporated. To create the latter, the chemist first enters related structures having assigned ^{13}C NMR spectra into the user database; then fragments are created according to certain rules from these structures. In this case the probability that the resulting database will contain fragments that are actually present in the unknown molecule under study is relatively high. If necessary, the program is also able to sequentially "construct" the molecule of the unknown compound from the fragments by overlapping common atoms or forming fragments sets that will be then used for structure generation under the given molecular formula. The generated structures can be verified using NMR and IR filters [15].

To provide the means for the preliminary analysis of 2D NMR data for non-contradiction, the StrucEluc program uses a heuristic algorithm that is capable in most cases (90-95%) of detecting the presence of contradictions as well as revealing the reasons for the contradiction. Then the program tries to automatically eliminate any contradictions found by lengthening those connectivies that according to the analysis results may have length contradictory to the one set by default. In the present example, we have employed the Structure Elucidator program for the identification of isolated degradants of the complex alkaloid cryptospirolepine to explore the flexibility of the program in terms of its ability to elucidate the correct structure and for the ability of a competent spectroscopist to interact with it in a convenient manner.

Chromatography:

The sample of cryptospirolepine was first interrogated by LC/MS methods to see how much of the starting alkaloid (MH^+ = 505) remained in the sample – none was found. A preparative chromatographic system was then developed for the sample, which was found to contain 26 components ranging from a major component DP-1 (~35%) to numerous components in the range of 2-3% (Figure 1). Initial effort was directed at isolating the two major components, DP-1 and DP-2 (~35 and ~16% of the total sample, respectively, based on peak area). The two major degradation products (DP-1 and DP-2) of cryptospirolepine (**1**) were isolated by reversed-phase, semi-preparative HPLC. A DMSO solution of the degraded alkaloid was diluted with mobile phase and injected onto a semi-preparative HPLC system consisting of a 21.2 x 250 mm, Kromasil C18 column with an acetonitrile-aqueous trifluoroacetic acid mobile phase. Detection was accomplished at 270 nm. Collected fractions of DP-1 and DP-2 were concentrated and desalted *via* trapping on a 10 x 250 mm, Kromasil C18 column. Eluent from the trapping column containing the degradation products was freeze-dried to yield about 1.1 mg of DP-1 at 96% purity by HPLC and about 200 μg of DP-2 at 95% purity by HPLC. NMR samples of about ~0.5 mg and ~100 μg, respectively, were used for the structural characterization effort.

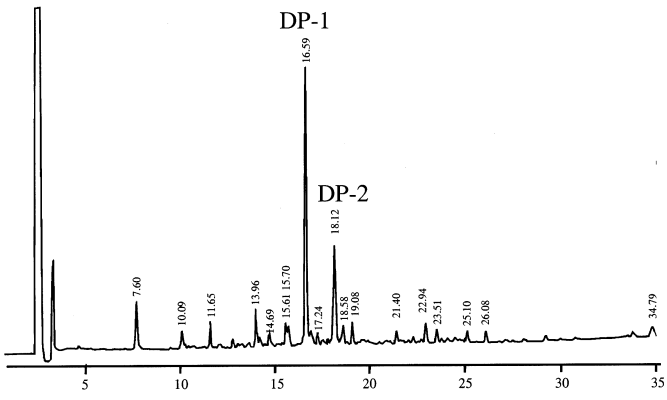
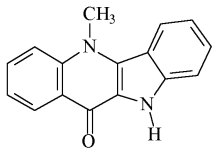


Figure 1. Chromatogram showing the distribution of peaks from a degraded sample of cryptospirolepine (**1**) after prolonged storage in d_6-DMSO. The two most abundant components, DP-1 and DP-2, were isolated in the present study. On the basis of LC/MS data, no cryptospirolepine (**1**) remained in the sample at the beginning of this study.

DP-1 – Cryptolepinone.

The isolated degradant, DP-1, which constituted the major component of the chromatogram shown in Figure 1 on the basis of peak area percent, was readily identified by mass spectrometry, giving a parent ion, MH^+ = 249, suggesting that the cryptospirolepine (**1**) had presumably been split in half during degradation. Using only GCOSY and GHSQC spectra, the structure was quickly identified (<20 min) as cryptolepinone (**2**) [17-20]. A quick (~ 1 h) GHMBC spectrum was acquired, solely to provide a few long-range connectivities, *i.e.* those from the *N*-methyls for the data to be fed into the CASE program.



2

The ACD/ StrucEluc [21] software package is composed of a number of software modules including both 1D and 2D NMR processing and NMR prediction of both ^1H and ^{13}C spectra from input chemical structures. Both NMR prediction modules offer the ability to construct user databases of structures and nuclear assignments that can be used to train the prediction algorithms. The 1D and 2D NMR and the mass spectral data were fed to the StrucEluc

program. Because of the relatively unconstrained data set (the lack of complete long-range data), the program generated a total of 208 structures as output. When the generated family of structures was sorted based on the match factor (*d*) distinguishing the deviation between the experimental and the predicted ^{13}C spectra, cryptolepinone had the lowest standard deviation (0.0). Such a deviation identifies the fact that the chemical structure already exists in the assigned literature databases. The ability to construct User Databases is useful in that it can preclude the "re-elucidation" of structures as a database grows and becomes more complete, thereby resulting in considerable time savings in the elucidation process.

The combined input of a set of GHSQC data and accurate mass and fragmentation data is a potentially useful way of establishing the identity of known compounds. The StrucEluc program produced a total of 14 structures from these input data. The output file generated, sorted on the basis of the ^{13}C average chemical shift deviation with a 5.0 ppm maximum deviation, are shown in Figure 2.

DP-2.

Mass spectrometry on the second most abundant isolate from the degraded sample gave a molecular ion, $MH^+ = 479$, which suggested a molecular formula of the DP-2 isolate as $C_{32}H_{22}N_4O$, based on a loss of 26 Da (C_2H_2) relative to well established molecular formula of cryptospirolepine (**1**) of $C_{34}H_{24}N_4O$. Significant fragment ions were observed in an MS/MS experiment at 464, 447, 435, 432, 247, 232, and 217 Da. It is interesting to note that the 232 daughter ion corresponds to cryptolepine minus a proton, suggesting a substituted cryptolepinyl moiety as a subcomponent in the structure of the degradant. This type of intuitive insight by the experienced spectroscopist should, ideally, be able to be utilized by a CASE program, which StrucEluc allows.

Physically, the sample of DP-2 dissolved in d_6-DMSO was initially reddish-orange in color. The initial proton spectrum at 500 MHz was rather broadened, suggesting that the sample was protonated from the chromatographic isolation. A small quantity of ammonia gas (headspace gas from a bottle of conc. ammonium hydroxide) was

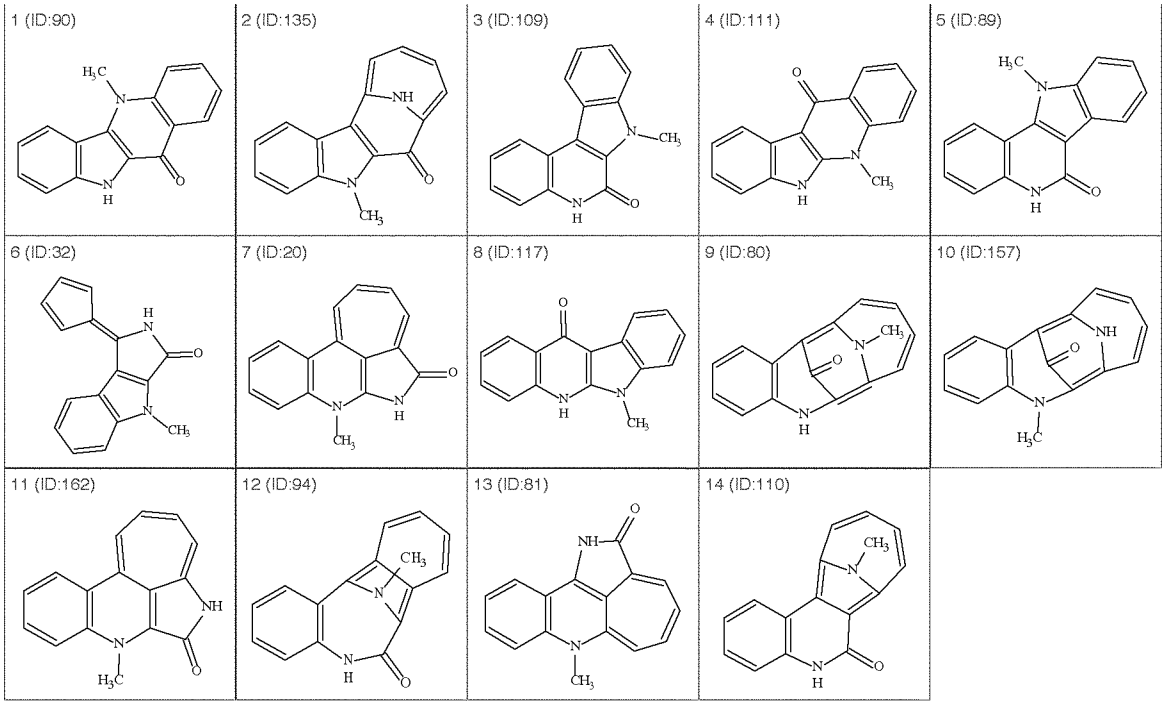
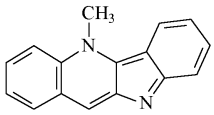


Figure 2. Structures generated by ACD/Structure Elucidator sorted on the basis of ^{13}C average deviation (<5 ppm maximum). It is interesting that five of the fourteen structures are indoloquinoline analogs which are plausible if this had been an unknown structure.

That cryptolepinone (**2**) was the primary component of the degraded sample of cryptospirolepine (**1**) is interesting and suggests, not surprisingly, that the spiro center of the parent molecule was a labile point in the structure and prone to oxidative degradation in DMSO solution, although a mechanism to explain the formation of **2** is not readily apparent.

bubbled through the sample causing the color to shift to deep purple, consistent with the extended conjugation of cryptolepine (**3**) or a cryptolepinyl substructure as implied by the mass spectral data [22-25]. A similar ammonia gas-induced color shift was noted during the characterization of the alkaloid cryptolepicarboline, which also contains an 11-cryptolepinyl moiety in its structure [26].



3

A GCOSY spectrum (Figure 3) and a GHSQC spectrum (Figure 4) were obtained using a Varian INOVA 500 MHz three channel NMR spectrometer equipped with a Nalorac 3 mm MIDTG-500-3 gradient inverse triple resonance probe. The experiments gave a proton spectrum containing two methyl signals that could be assigned as *N*-methyls (4.56 and 5.09 ppm – typical of cryptospirolepine) and a total of sixteen protonated aromatic CH's that were subgrouped into four four-spin systems by the GCOSY data. The isolated 11-methine singlet of cryptospirolepine from the indolobenzazepine-derived portion of the molecule was absent, as was the indole NH resonance. This observation, coupled with MS/MS fragment ions and the observed color change on treatment with ammonia gas, further strengthened the hypothesis of the structure of DP-2 containing an 11-cryptolepinyl moiety. The GHSQC experiment required an overnight data acquisition (17 h) using a conventional 3 mm gradient inverse triple resonance probe, suggesting an acquisition time of at least 3-4 days to acquire a usable HMBC spectrum. A 500 msec ROESY experiment gave four *N*-methyl to aromatic methine correlations (see **4** and **6**). A phase-cycled 8 Hz optimized HMBC spectrum was recorded next at 500 MHz using a Varian 500 MHz gradient inverse 5 mm triple resonance Chili-probe™ [27]. The initial, overnight HMBC spectrum contained 32 readily assigned responses; there was considerable overlap in the region from 7.3-7.47 ppm complicating signal assignments somewhat. Then, the available data were interpreted by the authors in addition to being loaded into ACD/ StrucEluc.

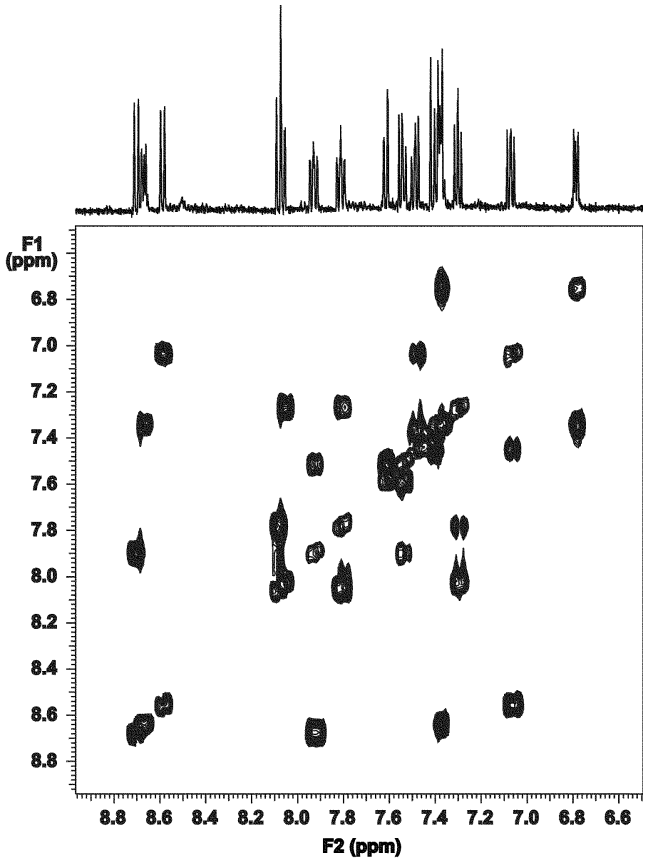
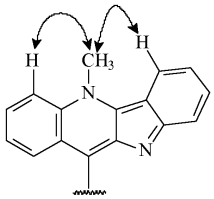


Figure 3. GCOSY spectrum of DP-2 acquired using a Varian INOVA 500 MHz three channel NMR spectrometer equipped with a Nalorac MIDTG-500-3 gradient triple resonance NMR probe.

Interpretation of the HMBC data quickly led to the deduction that, as suspected, the purple color of the solution in the NMR tube was due to the presence of an 11-cryptolepinyl species (**4**) contained in the structure. The initial run of ACD/ StrucEluc, performed in parallel with human data interpretation, gave an initial output of *ca.* 2000 structures, which when filtered afforded 107 11-crytolepinyl-containing structures. From the preliminary HMBC data, it was anticipated that the correct structure would necessarily locate a carbonyl resonance (167.4 ppm – typical of the 4-quinolone species contained in some *Cryptolepis* alkaloids) within 3 to 4 bonds of one of the terminal aromatic proton resonances (8.05 ppm). None of the 107 filtered structures met this requirement of the data set in a reasonable fashion – *i.e.*, none were indoloquinoline-based that could reasonably be derived from the degradation of cryptospirolepine (**1**) itself.



4

A second phase-sensitive, phase-cycled HMBC spectrum optimized for 6 Hz was acquired overnight and contained a total of 46 assignable responses, shown in Figure 5. A comparison of the heavily congested region of the conventional and phase-sensitive HMBC spectra is shown in Figure 6. As is readily noted from the comparative data presented in Figure 6, there is substantially better resolution in the highly congested region of the spectrum when phase-sensitive data are acquired. Interpretation of these data was again undertaken concurrently with the processing of the data using ACD/ StrucEluc [28]. This resulted in the output of almost 3300 structures after a generation period of almost 8 hours. After the removal of duplicate structures 355 structures remained. Examination of the elucidation results showed that structure 4 in the output table was consistent and indoloquinoline-based as expected (*vide infra)*. This structure is

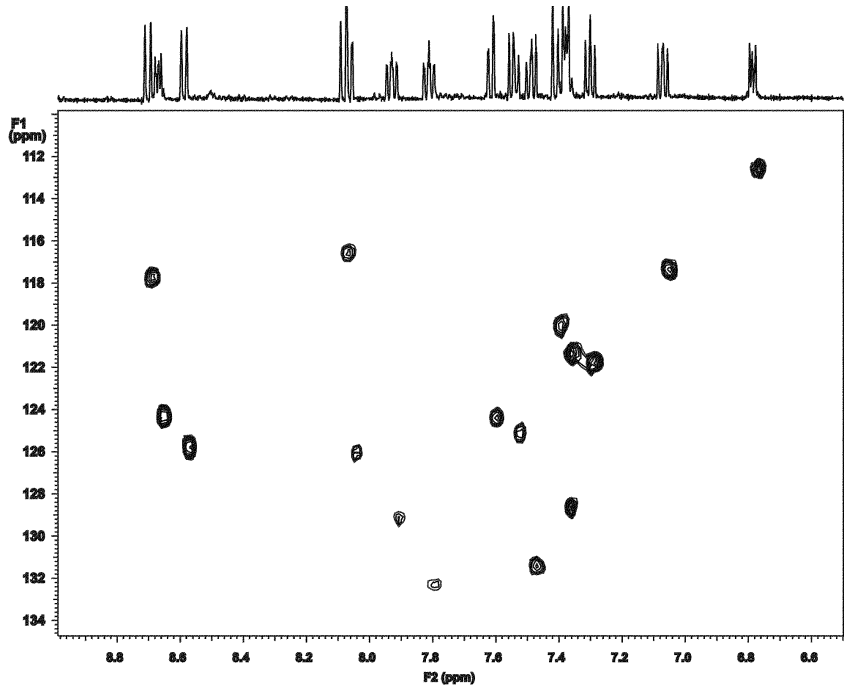


Figure 4. GHSQC spectrum of DP-2 acquired overnight using a Varian INOVA 500 MHz three channel NMR spectrometer equipped with a Nalorac MIDTG-500-3 gradient triple resonance NMR probe.

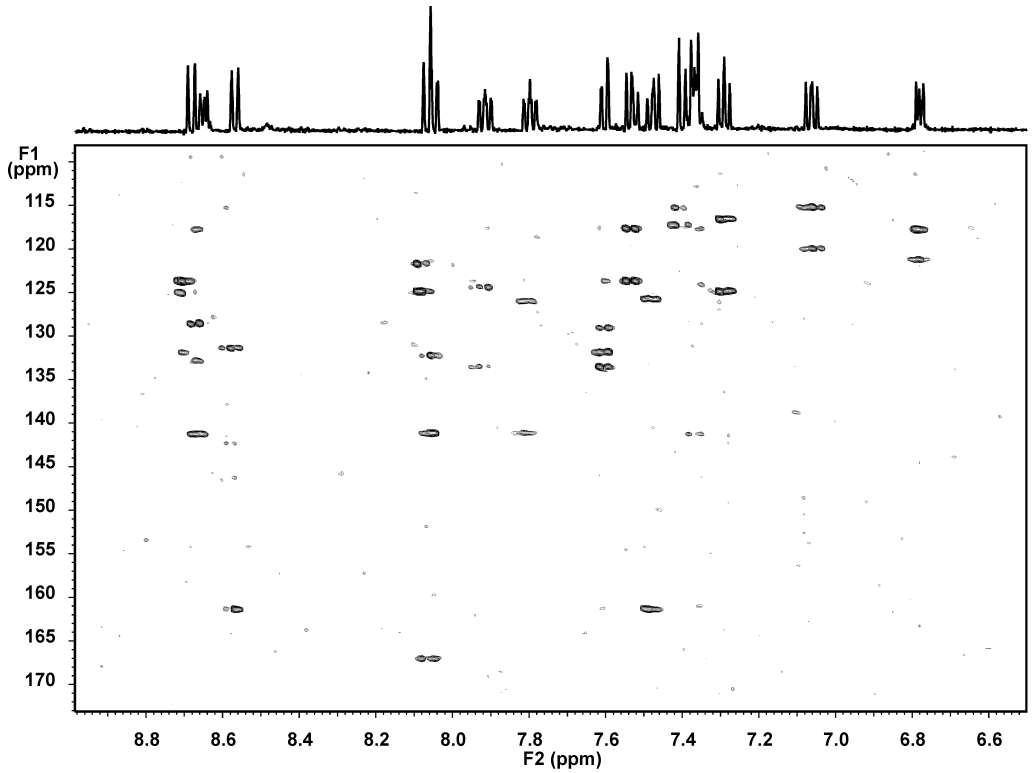


Figure 5. Phase-sensitive HMBC spectrum of the aromatic region of an ~100 mg of DP-2 in 150 mL d_6-DMSO in a 3 mm NMR tube acquired overnight using a Varian INOVA 500 MHz three channel NMR spectrometer equipped with a 5 mm gradient inverse triple resonance cryogenic NMR probe.

5,5'-dimethyl-5'*H*-10,11'-biindolo[3,2-*b*]quinolin-11(5*H*)-one. Structures were sorted on the basis of the agreement between the calculated ^{13}C chemical shift (d_A) and the "observed" ^{13}C shifts taken from the HMBC spectrum.

When the 11-cryptolepinyl fragment was included as one of the user fragments the elucidator produced 1268 structures in less than 10 minutes. This filtered to 111 structures after the removal of duplicates. The structure with the best combination of ^{13}C and ^{1}H deviations was displayed at position 2 (Figure 7) and was again 5,5'-dimethyl-5'*H*-10,11'-biindolo[3,2-*b*]quinolin-11(5*H*)-one (**7**) consistent with the expected structure. It is also worth noting that this structure has the best agreement when considered on the basis of calculated *vs.* observed ^{1}H chemical shift data.

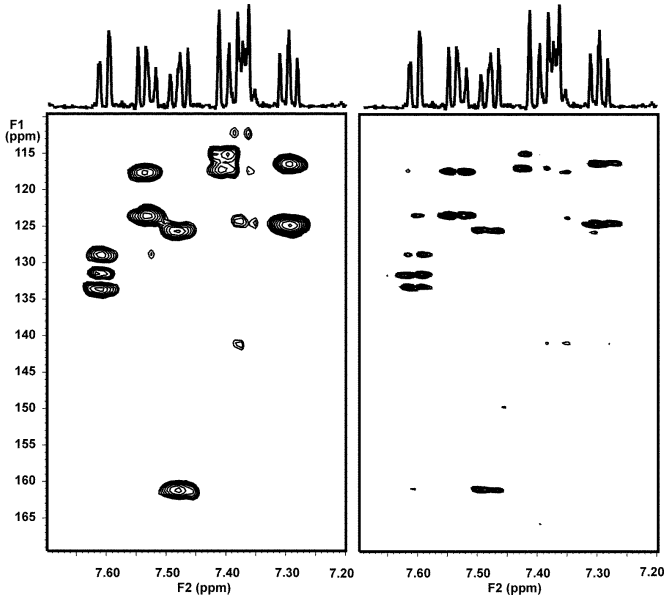


Figure 6. Comparison of the congested region of the long-range ^{1}H-^{13}C heteronuclear shift correlation spectra of DP-2. The conventional, phase-cycled HMBC data are shown on the left; phase-cycled, phase-sensitive HMBC data are shown on the right. Processing was identical but the data on the left were acquired using 200 increments of the evolution time, ni, while 224 hypercomplex increments were used to acquire the data shown in the right panel. All of the long-range correlations are well resolved in the phase-sensitive data, facilitating interpretation. In addition, some weak responses are observed in the phase-sensitive data that were not observed in the conventional, phase-cycled experiment.

A long-range ^{1}H-^{15}N CIGAR-HMBC experiment was also performed with an optimization from 3-6 Hz [13, 29-31] using the Chili-probe over a long weekend (~72 h). Long-range correlations were observed for the two *N*-methyl groups - 4.56/109.7 and 5.09/158.4 ppm. The former was consistent with the *N*-methyl chemical shift of the *N*-methyl contained in the indolobenzazepine-derived portion of cryptospirolepine, and cryptolepinone. The latter, in contrast, didn't agree well at all, but rather was in almost exact agreement with the corresponding *N*-methyl resonance of cryptolepine. One other long-range correlation was observed in the CIGAR-HMBC spectrum, that being a $^{5}J_{NH}$ correlation from the H1 resonance of the 11-cryptolepinyl moiety to the N10 nitrogen resonance (exo N=C) at 230 ppm. In comparison, the corresponding ^{15}N resonance of unsubstituted cryptolepine is observed at 207.8 ppm. The ^{15}N chemical shift data and observed long-range ^{1}H-^{15}N correlations were also fed as constraints to StrucEluc [32].

In the absence of the 11-cryptolepinyl fragment the elucidator produced about 4700 structures over about 15 hours. Removal of duplicate structures gave 334 final structures with the 5,5'-dimethyl-5'*H*-10,11'-biindolo[3,2-*b*]quinolin-11(5*H*)-one (**7**) again positioned second on the basis of ^{13}C shift and first based on ^{1}H shift data. When the 11-cryptolepinyl fragment and the ^{15}N correlation data were used 111 structures were generated in 20 minutes with 5,5'-dimethyl-5'*H*-10,11'-biindolo[3,2-*b*]quinolin-11(5*H*)-one (**7**) again located at position 2 in the table of structures sorted on the basis of chemical shift. A comparison of the

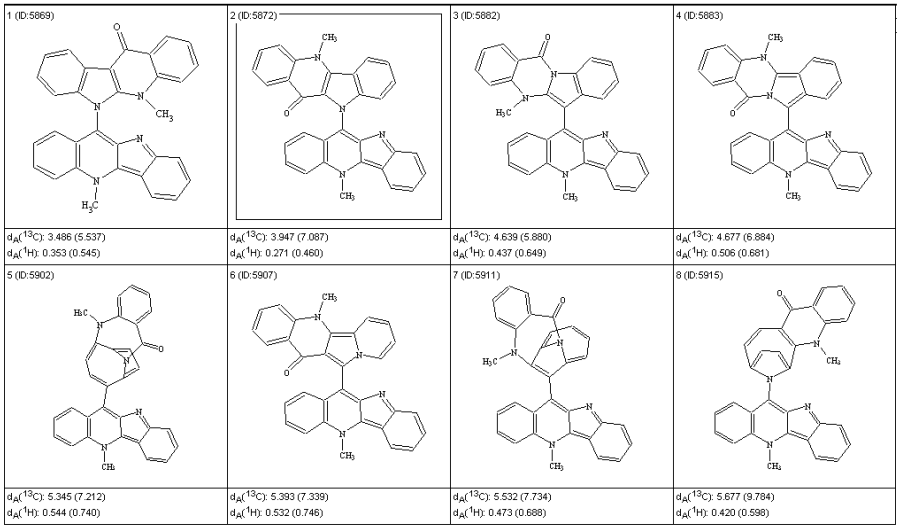


Figure 7. When the 11-cryptolepinyl fragment was included as one of the user fragments the Structure Elucidator produced 1268 structures in less than 10 minutes, which were filtered to 111 after the removal of duplicate structures. The structure with the best combination of ^{13}C and ^{1}H deviations was displayed at position 2 and again was the 5,5'-dimethyl-5'*H*-10,11'-biindolo[3,2-*b*]quinolin-11(5*H*)-one, consistent with the expected structure.

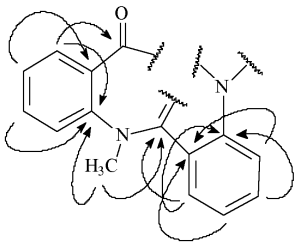
Table 1

Parameters used and results output from StrucEluc calculation.

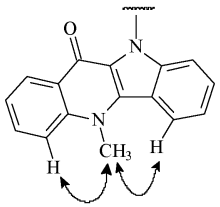
Cryptolepine fragment usage	N–H HMBC usage	Number of connectivities	Number of structures (generated/after duplicates removal)	Generation time	DP-2 position (by CNMR / by HNMR)
Yes	Yes	4	2158/111	20 min	2/1
Yes	No	2	1268/111	10 min	2/1
No	Yes	576	4683/334	15 h	2/1
No	No	288	3282/355	8 h	2/1

computation times and results obtained using Structure Elucidator are collected in Table 1.

The long-range data were interpreted manually to give the structural fragment shown by **5**. Key correlations observed in the HMBC spectrum are shown on the structure. ROESY correlations are denoted by double-headed arrows. Valence requirements and the empirical formula were satisfied by one additional sp^2 carbon, linking the carbonyl, the nitrogen, and the *N*-methyl containing bridge to give a second indoloquinoline moiety in place of the starting indolobenzazepine-containing structural unit of the starting cryptospirolepine to give **6**.

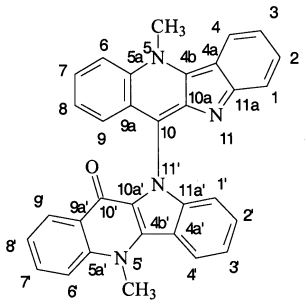


5



6

Linking the *N*-cryptolepinonyl substructure represented by **6** to the 11-cryptolepinyl substructure shown by **4** established the final structure of DP-2 as 5,5'-dimethyl-5'*H*-10,11'-biindolo[3,2-*b*]quinolin-11(5*H*)-one, which is shown by **7**.



7

Complete ^1H and ^{13}C, and partial ^{15}N chemical shift assignments and observed long-range heteronuclear couplings are summarized in Table 2.

It is interesting to note that the spiro-center of cryptospirolepine (**1**) was at the focal point of the degradation chemistry thus far examined. It will be interesting to see if this trend continues for the remainder of the numerous compounds represented in the chromatogram in Figure 1. The formation of cryptolepinone (**2**) can be assumed to be

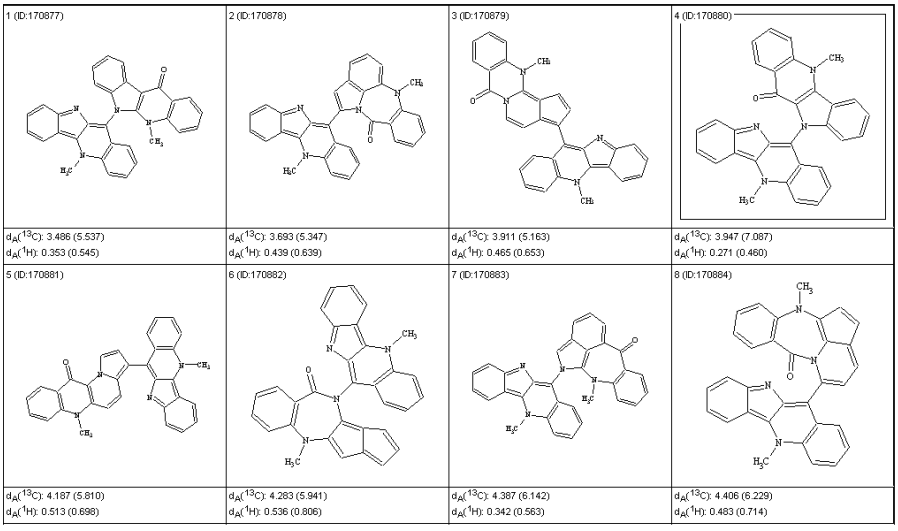


Figure 8. Structures generated by the second Structure Elucidator run using phase-sensitive, phase-cycled HMBC data. The 8 structures shown, from a total of 355 generated, were those with a ^{13}C average chemical shift deviation of < 10.0 ppm. The fourth structure agrees with the proposed structure of the DP-2 degradant of cryptospirolepine, as shown by **7**.

Table 2

^{1}H, ^{13}C, and ^{15}N Chemical Shifts, Long-Range Heteronuclear and ROESY Correlations observed for 5,5'-dimethyl-5'*H*-10,11'-biindolo[3,2*b*]quin-olin-11(5*H*)-one (7).

Position	d^{1}H	d^{13}C	d^{15}N	Long-range ^{1}H-^{13}C or ^{1}H-^{15}N Correlations (w = weak; vw = very weak; nd = not determined)	Important ROESY
1	7.40	120.1	---	C4a, C3	
2	7.47	131.5	---	C4, C11a	
3	7.06	117.4	---	C1, C4a	
4	8.58	125.8	---	C3, C4b(w), C11a	
4a	---	115.5	---	---	
4b	---	142.2	---	---	
5	---	---	158.6	---	
5-Me	5.09	39.8	---	C4b,C5a, N5	H4, H6
5a	---	133.6	---	---	
6	8.70	117.9	---	C8, C9a	
7	7.92	129.3	---	C5a, C9	
8	7.53	125.2	---	C6, C7, C9a	
9	7.61	124.4	---	C5a, C7, C10, N11 ($^5J_{NH}$)	
9a	---	123.8	---	---	
10	---	133.3	---	---	
10a	---	136.9	---	---	
11	---	---	230.0	---	
11a	---	161.4	---	---	
1'	6.77	112.6	---	C2', C3', C4a'	
2'	7.36	121.4	---	C4', C11a'	
3'	7.38	128.6	---	C1', C4a'	
4'	8.66	124.4	---	C2', C3', C4a', C4b', C11a'	
4a'	---	117.7	---	---	
4b'	---	132.9	---	---	
5'	---	---	109.4	---	
5'-Me	4.56	37.0	---	C4a', C5a', N5'	H4', H6'
5a'	---	141.4	---	---	
6'	8.08	116.6	---	C8', C9a'	
7'	7.80	132.2	---	C5a', C9'	
8'	7.29	121.8	---	C6', C7', C9a'	
9'	8.05	126.1	---	C5a', C7', C10', C10a'(vw)	
9a'	---	124.9	---	---	
10'	---	167.1	---	---	
10a'		130.0		---	
11'	---	---	nd	---	
11a'	---	141.4	---	---	

a relatively straight-forward oxidative degradation. In contrast, however, there is no readily obvious chemical pathway to explain the much more complex conversion of cryptospirolepine (**1**) to DP-2 (**7**).

Mechanistic considerations that must necessarily be implicit for such a conversion to have occurred are still being considered. The structures of additional degradants represented by the chromatogram shown in Figure 1 is ongoing and will form the basis for future reports. Of necessity, the rigorous structural determination of samples still considerably smaller than DP-2 will mandate the utilization of NMR cryoprobe technology.

Conclusions.

The significant improvement in sensitivity offered by cryogenic NMR probes can be applied to the elucidation of novel, unknown structures. The added sensitivity can be used to reduce data acquisition times for experimental data that are obligatory and can make feasible the acquisition of data such as long-range ^{1}H-^{15}N 2D heteronuclear shift correlations that would be inaccessible in practical periods of time when using conventional NMR probe technology. These data can also be used to advantage with CASE (Computer-Assisted Structure Elucidation) programs such as ACD/StrucEluc to aid structural chemists in the characterization of unknowns. As shown in this example, a group of spectroscopists with intimate knowledge of the structure elucidation of this family of indoloquinoline alkaloids still required about three days of interpretation time once all of the data were in hand. In contrast, using StrucEluc, under the worst-case scenario, plausible structural hypotheses would be available to the chemist for his/her consideration overnight once all of the data were in hand. It should also be noted that a CASE program such as StrucEluc could be used while data acquisition is on-going, allowing the investigator to refine his/her thinking "on the fly". This approach should offer advantages in the identification of natural products or the identification of degradation products of unknown types arising from compounds of known structure.

EXPERIMENTAL

The original sample of 2.5 mg of cryptospirolepine (**1**) was prepared in ~500 µL of deuterodimethylsulfoxide, 99.96% D and then sealed in a standard 5 mm NMR tube. On initial preparation, the sample was a red-orange color. After the elucidation of the structure of cryptospirolepine [1] the sealed NMR tube was stored at ambient temperature for ~10 years with no special precaution taken. On prolonged storage, the sample darkened to a deep brown color.

After the NMR tube was cut open, preliminary interrogation of the sample by LC/MS showed that none of the starting alkaloid remained in the sample. Analytical HPLC showed the sample to contain 26 components (Figure 1) ranging from the major components DP-1 (~35%) and DP-2 (~16%) (based on peak area uncorrected for relative response factors) to numerous small components in the 2-3% range. The deuterodimethylsulfoxide solution was diluted with an aqueous acetonitrile-water-trifluoroacetic acid mobile phase and injected onto a 21.5 x 250 mm semi-preparative Kromasil C18 column. Detection was accomplished by monitoring at 270 nm.

Mass spectrometric data for the isolated samples were acquired either on PE Sciex Q-Star time of flight mass spectrometer (low resolution and MS/MS measurements or using a Finnigan MAT-900ST mass spectrometer operating in the micro-ESI (micro electrospray ionization) mode. High resolution mass spectral data for the DP-1 and DP-2 isolates were obtained on a few micrograms of each of the isolates dissolved in mobile phase consisting of 50:50 methanol:water with 2% formic acid added. Accurate mass measurement was carried out by linear E-scan peak matching at a resolution near 8,500 (m/Δm, 10 % valley definition) using selected reference ions from PEG 400 to bracket the various sample pseudomolecular ions. The accurate masses of isolate DP-1 and DP-2 were measured as their protonated ions. These data established the molecular formulae of cryptolepinone (**2**) and the unknown, DP-1

(**7**). The latter, with an empirical formula of $C_{32}H_{22}N_4O$ based on high resolution measurements, represents a loss of 26 Da relative to the structure of cryptospirolepine (**1**). Fragmentation observed in the MS/MS experiments is discussed in the text.

NMR data for an ~500 µg sample of cryptolepinone (**2**) were acquired by dissolving the sample in ~150 µL of deuterodimethylsulfoxide (99.996% D, Cambridge Isotope Laboratories) and then transferring the sample to a 3 mm NMR tube (Wilmad). The data were acquired on a two channel Varian Inova 400 MHz NMR spectrometer equipped with a Nalorac Z•SPEC™ MIDG-400-3 gradient inverse 3 mm NMR probe. The sample was identified as cryptolepinone (**2**) from COSY and GHSQC data.

NMR data for DP-2 (**7**) were obtained by halving the sample and dissolving ~100 µg of the isolate in ~150 µL of deuterodimethylsulfoxide (99.996% D, Cambridge Isotope Laboratories) and then transferring the sample to a 3 mm NMR tube (Wilmad). All data were acquired on a Varian Inova 500 MHz three channel NMR instrument equipped with either a Nalorac Z•SPEC™ MIDTG-500-3 gradient inverse triple resonance probe or a Varian 5 mm gradient inverse triple resonance cryogenic Chili-probe™ operating at 25 K. Homonuclear NMR data were acquired using the conventional 3 mm NMR probe as were the GHSQC data, which were acquired overnight. On initially recording a proton reference spectrum, the aromatic proton resonances were broad and suggestive of traces of acid contamination from the isolation procedure. The sample was initially orange in color. After being treated with several bubbles of ammonia gas from the headspace of a bottle of conc. ammonium hydroxide, the color shifted to a deep purple color consistent with the extended conjugation of cryptolepine with commensurate sharpening of the proton resonances. Long-range 6 Hz optimized ^{1}H-^{13}C HMBC data, both magnitude calculated and phase-sensitive, were acquired overnight using the Chili-probe running the 3 mm tube coaxially in the 5 mm probe. Processing details are described in the figure captions. Long-range ^{1}H-^{15}N CIGAR-HMBC data [13,29-31] optimized for 3-6 Hz were acquired in approximately 72 h over a weekend, and gave correlations to 3 of the 4 nitrogen resonances contained in the structure of DP-2. Chemical shifts for the 3 nitrogens observed were consistent with ^{15}N chemical shifts for similar molecules in this series of alkaloids [29].

REFERENCES AND NOTES

* To whom inquiries should be addressed: e-mail: gary.e.martin@pharmacia.com; +269-833-6283; fax +269-833-2030.

[1] The sample used for this study was the original sample of cryptospirolepine prepared when the original structure elucidation was undertaken in late 1991. See: A. N. Tackie, G. L. Boye, M. H. M. Sharaf, P. L. Schiff, P., Jr., R. C. Crouch, T. D. Spitzer, R. L. Johnson, J. Dunn, D. Minick, and G. E. Martin, *J. Nat. Prod.*, **56**, 653 (1991).

[2] M. E. Munk, *J. Chem. Inf. Comput. Sci*., **38**, 997 (1998).

[3] K. Funatsu, S. Sasaki, *J. Chem. Inf. Comput. Sci*., **36**, 190 (1996).

[4] C. Peng, S. Yuan, C. Zheng, and Y. Hui., *J. Chem. Inf. Comput. Sci*., **34**, 805 (1994).

[5] C. Peng, S. Yuan, C. Zheng, Y. Hui, H. Wu, K. Ma, and X. Han, *J. Chem. Inf. Comput. Sci*., **34**, 814 (1994).

[6] C. Peng, S. Yuan, C. Zheng, Z. Shi, and H. Wu, *J. Chem. Inf. Comput. Sci*., **35**, 539 (1995).

[7] J.-M. Nuzillard and G. Massiot, *Tetrahedron*, **47**, 3655 (1991).

[8] C. Steinbeck, *Angew. Chem. Int. Ed. Engl*., **35**, 1984 (1996).

[9] C. Steinbeck, *J. Chem. Inf. Comp. Sci*., **41**, 1500 (2001).

[10] J. Meiler, M. C. Will, *J. Chem. Inf. Comp. Sci*., **41**, 1535 (2001).

[11] T. Lindel, J. Junker, M. Koeck, *Eur. J. Org. Chem*., **3**, 573 (1999).

[12] G. E. Martin and C. E. Hadden, *J. Nat. Prod.*, **63**, 543 (2000); R. Marek and A. Lycˇka, *Curr. Org. Chem.*, **6**, 35 (2002).

[13] G. E. Martin, "Qualitative and Quantitative Exploitation of Heteronuclear Coupling Constants," Ann. Rep. NMR Spectrosc., Vol **46**, G. A. Webb, Ed., Academic Press, New York, 2002, pp. 37-100.

[14] K. A. Blinov, M. E. Elyashberg, S. G. Molodtsov, A. J. Williams, and E. R. Martirosian, *Fresenius J. Anal. Chem*., **369**, 709 (2001).

[15] R. Araya-Maturana, T. Delgado-Castro, W. Cardona, and B.E. Weiss-Lopez, *Curr. Org. Chem.*, **5**, 253 (2001).

[16] M. E. Elyashberg, K. A. Blinov, and E. R. Martirosian, *Lab. Automat. Inf. Management*, **34,** 15 (1999).

[17] M. H. M. Sharaf, Ph.D. Dissertation, School of Pharmacy, University of Pittsburgh, 1993, pp. 177-192.

[18] A. Paulo, E. T. Gomes, P. J. Houghton, *J. Nat. Prod.*, **58**, 1485 (1995).

[19] M. M. Cooper, J. A. Lovell, and J. A. Joule, J. A. *Tetrahedron Lett*., **37**, 4283 (1996).

[20] M. H. M. Sharaf, P. L. Schiff, Jr., A. N. Tackie, and G. E. Martin, *J. Heterocyclic Chem.* **35**, 1365 (1998).

[21] ACD/StrucEluc, V5.08, Advanced Chemistry Development, 90 Adelaide Street W., Suite 702, Toronto, Ontario, M5H 3V9, Canada. All calculations were performed on a Pentium 1 GHz with 512 Mbytes of RAM

[22] E. Gellert, Raymond-Hamet, and E. Schlittler *Helv. Chim. Acta*, **34**, 642 (1991).

[23] D. Dwuma-Badu, J. S. K. Ayim, N. I. Y. Fiagbe, J. E. Knapp, P. L. Schiff, Jr., and D J. Slatkin, *J. Pharm. Sci*., **67**, 433 (1978).

[24] S. Y. Ablordeppy, C. D. Hufford, R. F. Bourne, and D. Dwuma-Badu, *Planta Med.*, **56**, 416 (1990).

[25] A. N. Tackie, M. H. M. Sharaf, P. L. Schiff, Jr., G. L. Boye, R. C. Crouch, and G. E. Martin, *J. Heterocyclic Chem.*, **28**, 1429 (1991).

[26] M. H. M. Sharaf, P. L. Schiff, Jr., A. N. Tackie, C. H. Phoebe, Jr., L. Howard, C. Meyers, C. E. Hadden, S. K. Wrenn, A. O. Davis, C. W. Andrews, D. Minick, R. L. Johnson, J. P. Shockcor, R. C. Crouch, and G. E. Martin, *Magn. Reson. Chem.*, **33**, 767 (1995).

[27] Work in the author's laboratory and results recently reported (see: W. Reynolds and R. G. Enriquez, *Magn. Reson. Chem.* **39**, 531 (2001)) have shown that with very small samples it is disadvantageous to use gradient-selected long-range experiments. Better results are obtained with conventional non-gradient phase-cycled experiments.

[28] Only strong (^{2}J or ^{3}J) peaks were entered for the first HMBC experiment using a 6Hz optimization. A total of 31 peaks were entered. Only 12 strong ^{3}J peaks were entered for the COSY experiment. Peaks in the second HMBC experiment, optimized at 8 Hz were manually separated into two groups ($^{2}J/^{3}J$ and $^{3}J/^{4}J$). A total of 22 $^{2,3}J$ peaks and 23 $^{3,4}J$ peaks were entered. The molecular formula of $C_{32}H_{22}ON_4$ was included as input.

[29] G. E. Martin and C. E. Hadden, *J. Nat. Prod.*, **63**, 543 (2000).

[30] G. E. Martin and C. E. Hadden, *Magn. Reson. Chem.*, **38**, 251 (2000).

[31] C. E. Hadden, G. E. Martin, and V. V. Krishnamurthy, *J. Magn. Reson.*, **2000** 38, 143 (2000).

[32] A total of three long-range ^{1}H-^{15}N correlations were added to the input dataset. One was entered as a long-range (^{4-5}J) coupling.