Pergamon

Tetrahedron Vol. 50, No. 42, pp. 12267-12280, 1994

0040-4020/94 $7.00+0.00

0040-4020(94)00772-1

"Connectivist" Approach to Organic Structure Determination Lsd-Program Assisted Nmr Analysis of the Insect Antifeedant *Azadirachtin*

Steven V. Ley†, Kevin Doherty†, Georges Massiot¶ and Jean-Marc Nuzillard*¶[1]

†*University Chemical Laboratory, Cambridge University, Lensfield Road, Cambridge, CB2 1EW, U.K.*

¶*Laboratoire de Pharmacognosie, Unité Associée au CNRS no 492, Faculté de Pharmacie, 51, Rue Cognacq-Jay, F-51096 Reims Cédex. France.*

Abstract: This paper sets out a method for 'intelligent' structural determination of complex organic molecules that has as its basis the Logic for structure determination program (Lsd). Here it is used for the automatic analysis of ^{1}H and ^{13}C NMR correlation data derived for the limonoid antifeedant Azadirachtin.

INTRODUCTION

At present there exists an active search for environmentally compatible alternatives to traditional chemical strategies of pest management in modern agriculture.[2] This arises in response to the need for more efficient use of a diminishing farmland resource, to help maximise crop returns and so cope with the demands of an ever increasing world population. At the same time, it is intended to steer a course away from the environmentally damaging consequences of continued intensive use of synthetic insecticides with their attendant non-selective toxicity, and the complications associated with insect acquired resistance to these toxins.[3]

Of the number of secondary metabolites known to have roles in the defence mechanisms that plants have evolved to counter insect predation,[4] the C-seco limonoid Azadirachtin (Figure 1), a tetranortriterpenoid component of the seeds, leaves and bark of the Neem tree (*Azadirachta indica* A. Juss (Meliaceae)), has attracted the most attention. Since the earliest disclosures of its insect antifeedant properties,[5] coupled with a growth modifying behaviour against a range of insect species,[6] and demonstrated non-toxicity toward higher forms of life,[7] this compound has been at the centre of considerable commercial and research activity. Presently azadirachtin, available commercially as a Neem extract, is being used as a novel insect control agent, in integrated pest

management control programmes, and is also the subject of intense chemical investigation as a target of modern synthetic methodology.[8]

Current structure-activity studies are set to determine at the molecular level the minimum structural requirements needed to elicit the antifeedant response,[9,10] and should eventually enable the synthesis of structurally concise analogues of equivalent activity. Significantly this work has highlighted the functional synergy of the main structural components in azadirachtin,[11] notably between the dihydrofuranacetal (fura[2,3b]pyran) and decalin fragments,[12,13] paralleling the known structural demarcation in this now well characterised molecule[14] (Figure 1). Although earlier studies had identified most of the functional groups,[15,16] the structural proposals of the time were fundamentally flawed, omitting important structural elements and/or failing to assign their correct position in the molecule.[17] The presently accepted structure for azadirachtin shown in Figure 1, was the product of later, detailed NMR investigations[18,19] and a final definitive X-ray analysis of the analogue, 22,23-dihydrodetigolylazadirachtin.[20] Comprehensive accounts of the studies leading to the solution of azadirachtin structure can be found in the recent literature.[21]

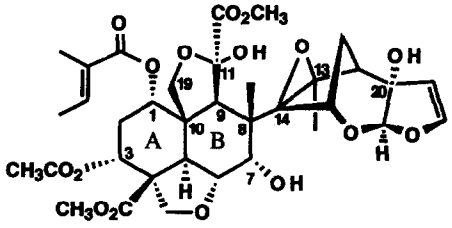


Figure 1. Azadirachtin

Inspection of Figure 1 reveals the structural complexity underlying the biological behaviour and shows that an obvious structural disconnection in the molecule, and one that represents the final step of a current synthetic approach,[8] centres on the C8-C14 bond linking the dihydrofuranacetal (fura[2,3b]pyran) and decalin fragments. Rotational freedom about this bond, manifest as a temperature-dependent behaviour in the NMR spectra,[12] is likely to have important consequences for the antifeedant properties and may explain the functional synergy in the molecule.[11] Collectively the two fragments, which describe the right and left hand sides of the molecule,[12,13,20] present a complex array of oxygen based functionality, whose structural relationship, in particular the pattern of intramolecular hydrogen bonds they establish, may be potentially important to the recognition of azadirachtin by its natural receptor. Foremost among these is a conformationally stabilising hydrogen bond between the C9-C10 annulated tetrahydrofuranacetal, specifically its C11 hemiacetal hydroxyl, and the C13-C14 oxirane ring oxygen,[18] with a weaker such contact shared by the C7 and C20 hydroxyl groups.[20] Clearly azadirachtin is of sufficient structural complexity to be a rigorous test of the structure solving capability of the Logic for structure determination program

(Lsd) and in view of its important biological properties, warrants consideration as a target of the present study of computer aided methodology for structure elucidation.

The automatic analysis of spectral data aimed at accurate structural prediction is a topic of much theoretical and practical importance, that has been the focus of considerable attention in the early stages of the development of artificial intelligence related software. In fact many of the concepts integrated within the present generation of programs related to the Dendral project (DENDRitic ALgorithm), have their origins in this earlier work.[22] These 'state of the art' programs serve as an excellent example of how powerful present computer technology can be when used as an analytical tool for the solution of chemical problems.[23] They are intended to provide the practising chemist with constitutional molecular formulae, that are based on a search of the best way of assembling elements of structure that is consistent with available ^{13}C chemical shift data in addition to other sources of spectral and chemical information. The predicted structures arising from this analysis are usually presented in the form of simple two-dimensional drawings with no attempt made to accurately portray bond lengths and bond angles. This offers the chemist an impression of the molecular architecture where perhaps no other structural information had previously been available and serves as a useful starting point for further refinement and discussion. Finally, three-dimensional modelling programs which to date have found their principle role in studies of macromolecular structure[24] using coupling constant and nuclear Overhauser (nOe) data, can also serve here to decipher the stereochemistry in the molecule. Recent attempts to rationalise ^{13}C chemical shifts in terms of stereochemical relationships[25] should also find application in this regard.

The logic for structure determination program (Lsd)[26] discussed here uses mainly the direct and remote carbon-proton chemical shift correlation data of two-dimensional NMR experiments to generate molecular structures. Lsd is a stand-alone program and does not require supplementary spectroscopic archives. Here we provide evidence of its ability to arrive at chemically sound structural predictions for azadirachtin.

RESULTS AND DISCUSSION

The general steps of the structural elucidation involve first the collection of the required spectral information, followed by data reduction and finally by analysis leading to the predicted structure(s). As mentioned, the Lsd program exploits the connectivity data contained in two-dimensional NMR correlation maps. The use mainly of ^{13}C-^{1}H correlation data underpins this whole approach to automated structure solving, complemented by substructural information derived from elementary analysis of ^{13}C spectra. This method provides directly structural information that avoids the uncertainties inherent to structural conclusions founded solely on chemical shift knowledge. In the manner of the information supplied by the correlation experiments, certainties replace assumptions as input to the logic of automated structure solving. The principal source of input is taken from Heteronuclear Multiple Bond Correlation (HMBC) spectra,[27] set up to detect the chemical shifts of ^{13}C and ^{1}H atoms separated by two or three bonds in the molecule ($^{n}J_{C\text{-}H}$, n = 2

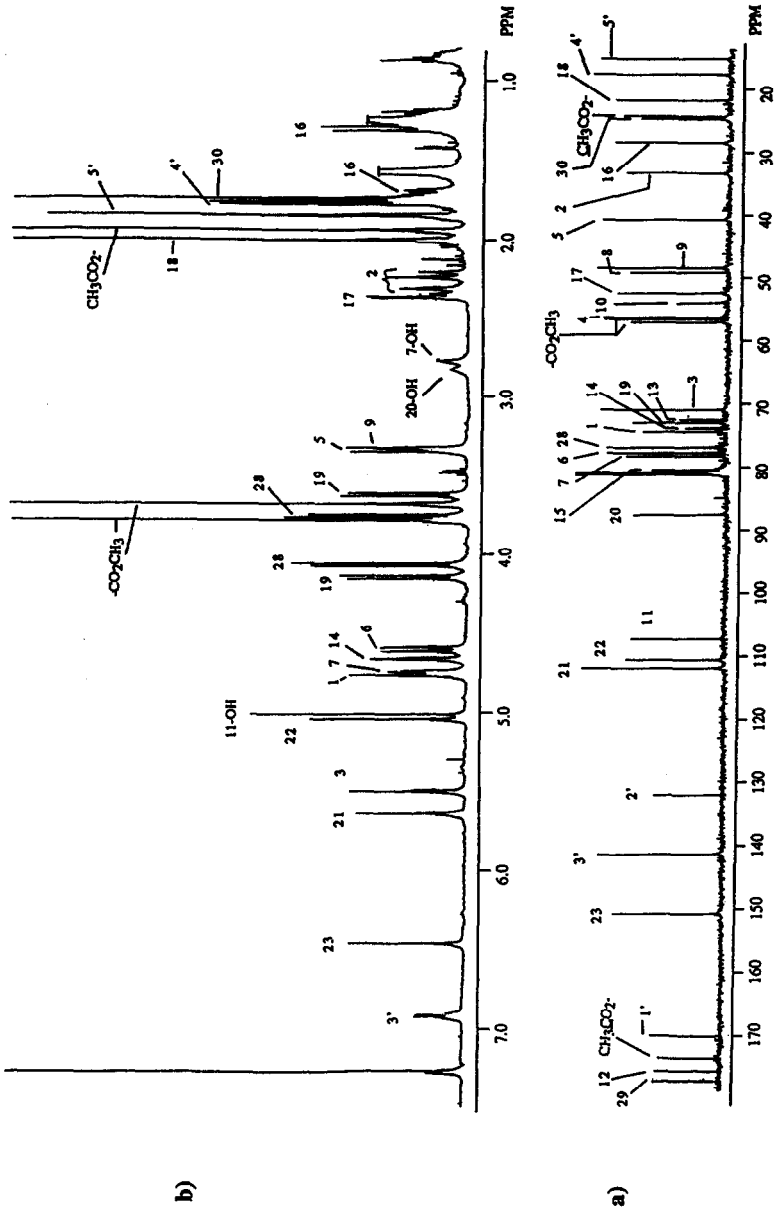


Figure 2. Nmr spectra of azadirachtin, showing peaks numbered according to the convention adopted in ref.19., a) ^{13}C spectrum (75MHz), SI=16K and NS=12000 accumulations, b) ^{1}H spectrum (500MHz), SI=32K and NS=64 accumulations; contents of sample as described in the experimental section.

or 3). When this information is combined with the results of HMQC (Heteronuclear Multiple Quantum Coherence) experiments,[28] establishing via ^{1}H detection (inverse mode),[29] direct ^{13}C-^{1}H correlation ($^{1}J_{C-H}$), it is possible to relate the shifts of carbons separated by up to two bonds ($^{n-1}J_{C-C}$, n = 2 or 3) and in this manner trace the carbon backbone of the molecule. Formally this approach is equivalent to performing a two dimensional INADEQUATE experiment (Incredible Natural Abundance DoublE QUAntum Transfer Experiment)[30] with the difference in the latter case, that there is no requirement for the carbons to be protonated to establish their connectivity and furthermore the INADEQUATE spectra provide only unambiguous single bond ^{13}C-^{13}C correlation. An important advantage of the combined HMBC-HMQC approach described here, apart from being technically more practical, is that the spectra are intrinsically more sensitive ($\gamma_H / \gamma_C = 4$) than those of the INEPT based methods,[31] and moreover both techniques should benefit from the now available pulsed field-gradient hardware. To complement the HMQC experimental data, there are the correlations contained in homonuclear ^{1}H-^{1}H COSY spectra (COrrelation SpectroscopY),[32] which together, allow one-bond (from $^{3}J_{H-H}$ scalar coupling) and/or two bond (from $^{4}J_{H-H}$ scalar coupling) carbon-carbon connectivity in the molecule to be determined.

In the next step of the process the data are reduced, namely the relevant spectral information, *i.e.* the chemical shifts, J-coupling data and correlations are translated into encoded data sets. The latter consist of groups of resonances, specifically ^{13}C peaks, numbered in the order of decreasing chemical shifts in the resolved one-dimensional spectrum (see Figure 2a,b; Figure 3 and Table I), that have been related, in the main, through spectrally established chemical connectivity.

Encoding data for use by the Lsd-program is a manual process requiring careful inspection of plotted spectra; this situation should improve in the near future with the introduction of an 'intelligent' peak picking algorithm for both one and two dimensional nmr spectral analysis. In building data sets, the first task is to correctly number the signals appearing in the one dimensional carbon and proton spectra. Numbering, as indicated above, follows a simple user specified convention, such as the order in which signals appear according to chemical shift in the one dimensional spectrum (Figure 2a,b; Figure 3). The normally good dispersion of ^{13}C shifts facilitates numbering in the carbon spectrum, and also of the proton spectrum, where partially or completely hidden multiplets may, as a consequence of this property of ^{13}C nuclei, be well resolved in the HMQC correlation map. In this manner it is possible to identify most if not all the proton signals for inclusion in the numbering process although severely overlapping peaks can be neglected without serious effect. The same consideration does not apply to the carbon signals however, all of which have to be identified at this time and numbered accordingly. Even superimposed signals are numbered collectively, with HMBC established correlations between any proton signal and a group of such carbons handled by the program, working on the assumption that at least one member in the group should be responsible for the observed connectivity *i.e.*, treating all possible combinations.

Once the numbering is complete, the HMQC, HMBC, and COSY spectra, specifically the correlations they identify, are translated into sets of peak co-ordinates *i.e.*, the correlations observed are each assigned a pair of numbers designating a particular ^{13}C-^{1}H or ^{1}H-^{1}H combination sharing a connection in the molecule. Some care has to be exercised however when interpreting the HMBC spectra since these are especially prone to artifacts and, as a general rule, it is better to discard

suspect data rather than run the risk of introducing uncertainties into the final analysis. Analysis of the resulting sets of numbered correlation data allows the underlying carbon connectivity in the molecule to be determined. Crucial to the structural solution is a proper evaluation of the bond path linking sets of correlated carbon nuclei. The result of this combined treatment of the HMBC / HMQC data is accurately described by an imaginary carbon-carbon correlation spectrum where the observed correlations, representing a fictious coupling $^{n-1}J_{C-C}$ (n = 2 or 3), establish, by definition, the existence of an n-1 bond path between the connected nuclei.

Other sources of chemical and spectral (^{1}H NMR,IR,UV,MS) input at this stage of the data reduction can offer an additional degree of refinement in building the data sets *i.e.*, grouping of atoms, that impose constraints on the number of structural solutions sought by the Lsd-program in the resolution process. Information needed includes a knowledge of the molecular composition and of the 'atom status' for each non-hydrogen atom in the structure. The atom status is defined by an atom-assigned number (user specified, discussed above), state of hybridisation and valency *i.e.* number of bound hydrogens; for carbons in the molecule this information can readily be obtained from ^{13}C and DEPT subspectra.[31] Similar information has also to be provided for all heteroatoms in the structure (oxygen, nitrogen *etc*), which are numbered arbitrarily. In situations where the state of hybridisation is not easily determined *e.g.*, carbons resonating around 100 ppm which can be either sp^2 or acetalic sp^3 in nature, separate data sets must be built to take account of the various possibilities in atom status. Special properties deduced from an elementary analysis of spectral data *i.e.*, chemical shifts, couplings *etc.*, are also used as input to specify the status of neighbouring atoms and thereby restrict the sets of possible atom-atom pairings in the bond generating process. The Lsd-program has the means to build data sets according to atom status and the special properties data. In arriving at structural solutions the program first treats correlated nuclei assigned with either one or two bond connectivity in the bond forming process, followed then by bond generation between pairs of incompletely connected atoms not appearing in the correlation data. Similarly, if a bond or an element of structure is known to be present, *e.g.*, the decalin fragment *i.e.* A-B ring system of azadirachtin,[15] it can be introduced into the data base as a collection of sub-bonds between sub-atoms having their own numbering system and used in the structure generation.

The current edition of the Lsd-program is written in C-language with the advantage that it should find much wider application, in addition to offering resolution times on the order of ~10^2 faster than those achieved with the earlier Prolog based version. The processing of data is based on a recursive generation and test algorithm, with a simple topological criterion used by the program to search out anti-Bredt structures[33] and discard them, allowing for the sorting of unrealistic computer generated structures according to this type of steric constraint. In partial mode the program will generate partial structures that take account of all of the correlation data, with the possibility of completing these structures *i.e.*, adding in bonds at free positions to yield valid solutions, constantly verified. The output comprises lists of atom status and of the bonds generated in all structural solutions arrived at by the program, satisfying the given constraints. These listings can then be directed to the input of other programs for either two dimensional representation of structures or three dimensional molecular modelling. Options also available include single step driven operation and report generation, intended to provide a log-book of the resolution process.

Table I. (a) Carbon atoms numbered in order of decreasing chemical shift (Figure 2a). (b) List of carbon chemical shifts measured at 75 MHz with TMS as internal standard. (c) Carbon position according to convention adopted in ref. (19) and given in Figure 2a. (d) Proton reference number of directly bound hydrogen atom(s) from HMQC experimental data, see Table II. (e) Proton reference number of remotely connected proton spins from inspection of HMBC data.

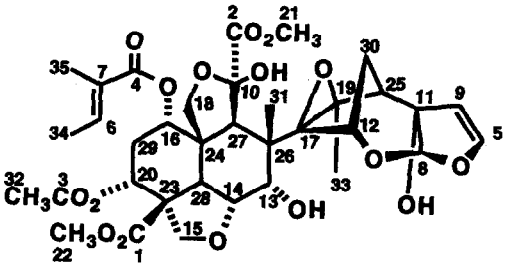
Carbon Atom No.[a]	δ (^{13}C)ppm[b]	Position [c]	HMQC [d]	HMBC [e]
1	173.4	29	-	12, 13, 17
2	171.8	12	-	15, 18
3	169.6	C(=O)OCH$_3$	-	4, 24
4	166.2	1'	-	1, 7, 25
5	147.0	23	2	3, 6
6	137.6	3'	1	25, 26
7	128.6	2'	-	25, 26
8	108.7	21	3	2, 6, 9, 21
9	107.4	22	6	2, 21
10	104.2	11	-	11
11	103.6	20	-	2, 6, 21
12	76.4	15	9	3, 21, 28
13	74.3	7	8	10, 27
14	73.8	6	10	17
15	72.9	28	12	-
16	70.5	1	7	11, 18
17	70.0	14	-	21, 23, 27, 28
18	69.0	19	11	18
19	68.5	13	-	9, 21, 23, 28
20	66.9	3	4	7, 12
21	53.2	C(=O)OCH$_3$	15	-
22	52.7	C(=O)OCH$_3$	13	-
23	52.4	4	-	4, 17
24	50.1	10	-	17
25	48.6	17	21	3, 9, 23
26	45.4	8	-	18, 27
27	44.6	9	18	8
28	37.0	5	17	12
29	29.7	2	22	-
30	25.0	16	28	-
31	21.3	30	27	18
32	20.8	CH$_3$CO$_2$	24	-
33	18.3	18	23	-
34	14.3	4'	26	1
35	11.9	5'	25	1

Table II. (a) Protons numbered in order of decreasing chemical shift (Figure 2b). (b) List of proton chemical shifts measured at 500 MHz, with TMS as internal standard. (c) Proton position according to convention adopted in ref. (19), see Table I and Figure 2b (d) Proton reference numbers of directly correlated hydrogen atom(s) inferred from COSY experimental data.

Hydrogen Atom No.[a]	δ (^{1}H)ppm [b]	Position [c]	COSY [d]
1	6.92	3'	26
2	6.43	23	6
3	5.63	21	-
4	5.48	3	22
5	5.06	11-OH	OH
6	5.02	22	2
7	4.74	1	22
8	4.71	7	10
9	4.64	14	-
10	4.58	6	8, 17
11	4.13	19	-
12	4.04	28	-
13	3.78	OMe	-
12	3.75	28	-
15	3.66	OMe	-
11	3.61	19	-
17	3.35	5	10
18	3.22	9	-
19	3.09	7-OH	-
20	2.99	20-OH	-
21	2.36	17	28
22	2.32	2	-
22	2.22	2	4, 7
23	1.98	18	-
24	1.93	$\underline{C}H_3C(=O)O$	-
25	1.83	5'	-
26	1.77	4'	1
27	1.72	30	-
28	1.68	16	10
28	1.30	16	10

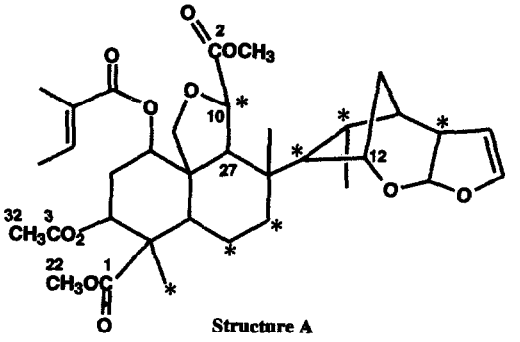
A complete description of the results of NMR experiments used in the structural elucidation of azadirachtin is presented in Tables I and II. Inspection of these data, in addition to providing insight concerning the carbon connectivity in the molecule, also supplies supplementary input *i.e.*, substructural information: an important contributor to the data base processed by the program and outlined in the following section.

The molecular formula of azadirachtin has been determined as $C_{35}H_{44}O_{16}$, with the presence of a decalin fragment *i.e.*, intact A-B ring system, inferred by the results of earlier chemical modification studies.[15] DEPT experimental data established that in azadirachtin there are seven methyl, four methylene and twelve methine carbon sites, from which we conclude the twelve remaining carbons to be quarternary in character. In addition, a proton inventory reveals the presence of three hydroxyl groups. From the carbon data listed in Table I it is clear that atoms numbered 1 through 4 are carbonyls, signalling the presence of eight sp^2 paired atoms in the structure *i.e.*, four carbon and four oxygen atoms. The chemical shifts reported for carbons numbered 5 to 7 identify these as sp^2 in character, from which it can safely be said that among the carbons resonating in the 100 ppm region, namely atoms 8, 9 and 10 (for the special case of atom 11 see below), one or all of these must be similarly hybridised, in keeping with the requirement for sp^2 sites to be paired *i.e.*, the number of such sites in the structure to be even. As mentioned earlier, signals appearing in this region of the carbon spectrum could equally arise from sp^2 or acetalic sp^3 hydridised sites and so account must be taken of these possibilities in assigning atom status for azadirachtin. Among carbons 8, 9, and 10, those which are sp^2 are bound only to carbon, and the others are by necessity sp^3 hybridised and of the acetal type. Furthermore it can be seen from Table I that all atoms numbered above 10 are sp^3 hybrid sites in the structure *i.e.*, carbons numbered 11 through 22, and that on the basis of their shifts, all are bound to exactly one sp^3 oxygen atom. The true identity of carbon 11 as an sp^3 hybrid site in the molecule, in fact as a tertiary alcohol, is established upon inspection of Figure 2b and Table II, the unusually high and somewhat misleading chemical shift value (Table I) is attributed to the proximity and deshielding influence of the π-orbital electrons in the adjacent enol ether moiety of the dihydrofuranacetal fragment. Inspection of the data in Figures 2a,b and Table I shows that the methylene protons bound to carbons 15 and 18 appear as pure AB-subspectra and that, according to their chemical shifts and the absence of any additional couplings elsewhere in the molecule, we conclude that these sites are adjacent to a single oxygen and a quarternary centre in the molecule. Carbons 31, 32, 33 and 35 are also each directly bound to a quarternary centre, based on the observation that proton signals arising from this group of atoms appear only as methyl singlets. The splitting of the proton signal assigned to carbon 34 signifies its presence in the molecule next to a methine carbon. The data also show that atoms 7, 23, 24 and 26 are only bound to adjacent carbon sites. Finally, examination of the COSY results listed in Table II, establish the existence of one bond connectivity between seven pairs of sites in the molecule, namely carbon atoms numbered 6-34, 5-9, 20-29, 13-14, 16-29, 14-28 and sites 25-30. The supplementary information detailed above together with the observed correlations arising from the HMBC / HMQC experiments, contributed to the data base processed by Lsd in solving for the structure of azadirachtin.

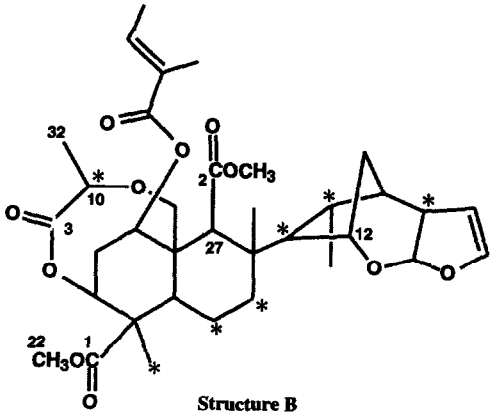


Azadirachtin

Figure 3. Numbering convention employed for azadirachtin, reflects carbons in order of increased shielding in the resolved one dimensional ^{13}C spectrum, see Figure 2a and Table I.



Structure A



Structure B

Figure 4. Lsd predicted solutions for azadirachtin structure

Running in partial mode, the Lsd-program arrived at two solutions, shown as structures A and B in Figure 4, where the stars denote free positions to which oxygen atoms can be added to complete the structure. Both these solutions were obtained by imposing the constraint that atoms numbered 8 and 10 are acetals, with carbon 9 the sp^2 partner to any of the atoms 5, 6, or 7, as explained earlier. In the absence of this restriction, all other possible combinations cause the program to fail in its attempts to arrive at a structural solution. It can hopefully be appreciated that structure A offers a very close fit and can be completed readily by filling in the free positions, to generate either the now well accepted structure of azadirachtin (Figure 1),[20] or those which had been postulated earlier.[12,17] Structure B shows that the acetate group at site 20 has been missed, with carbon atom 10 inserted between sites 3 and 32, accounting for the possibilities presented by the combined HMBC / HMQC correlation data, which suggest that these positions could exist either directly attached as presented in solution A or separated by a common atom as depicted in B. A further result of this structural relationship in B is that carbons 2 and 27 now share a direct bonding situation, instead of being separated by atom 10, as seen for the hemiacetal component correctly identified in structure A. Restricting here the structural search by assuming that atoms 3 and 32 *a priori* share a direct bonded connection would result in the program arriving at a single solution for the structure of azadirachtin.

CONCLUSION

The important outcome of the predictions based on the Lsd processing of the correlation data and supplementary information of this study has been an accurate representation of the carbon backbone, as well as the correct identification and disposition of key structural elements known to be present in azadirachtin (Structure A),[20] which should attest to the general usefulness of this approach as an aid to structural elucidation of complex natural products. It is hoped that this report will draw the attention of synthetic chemists to the existence of the new generation of programs, such as Lsd, for the automated analysis of spectral data and encourage their use on a routine basis, as a first step toward accurate prediction of molecular architecture, that serves as a useful complement to X-ray structural determinations as demonstrated here for the test case of the structurally complex molecule azadirachtin.

EXPERIMENTAL

NMR experiments were performed mainly on a Bruker AC300 spectrometer, modified for inverse detection, equipped with an inverse configuration probehead and BFX-5 heteronuclear decoupler. The 500MHz ^1H NMR data of azadirachtin shown in Figure 2b, were collected on Bruker AM500 spectrometer. Samples typically contained 15mg of compound in 0.5ml deuteriochloroform, with all the data collected at 27°C unless stated otherwise. Pure azadirachtin was isolated from a crude sample (30%) of Neem seed extract, kindly provided by the Rohm &

Haas Co., with rapid column chromatography on silica gel using a 7:3 mixture of ethyl acetate / petroleum-ether (40-60°) as the solvent system.

(^{13}C-^{1}H) Heteronuclear multiple bond correlation spectra (**HMBC**) were obtained with the 'INVDR2LP' pulse sequence: (^{13}C) PW_{90} = 6.5 μs, (^{1}H) PW_{90} = 13 μs, SI1 = 256, SI2 = 2K, SW2 = 2066Hz (6.89ppm), SW1 = 12820 Hz (170ppm), DS = 4, AQ = 0.5s, NS = 112, t_d = time delay to detect long range coupling = 70 ms.

(^{13}C-^{1}H) Heteronuclear multiple quantum coherence spectra (**HMQC**) were collected in phase sensitive mode using TPPI and the 'BIRDD9' pulse sequence: (^{13}C) PW_{90} = 6.5 μs, (^{1}H) PW_{90} = 13 μs, SI1 = 512, SI2 = 2K, SW2 = 2066Hz (6.89ppm), SW1 = 10550 Hz (140ppm), DS = 4, AQ = 0.5s, NS = 40.

Homonuclear ^{1}H **COSY-45** data were acquired with the 'COSY' pulse sequence and the following parameters: (^{1}H) PW_{90} = 5 μs, SI1 = 256, SI2 = 1K, SW1 = SW2 = 2250 Hz (7.5ppm), DS = 4, AQ = 0.23s, NS = 8.

DEPT data was obtained using the pulse sequence 'DEPTVAR' and the parameters: SW = 15840Hz (210ppm), SI = 16K, NS = 12000.

ACKNOWLEDGEMENT

The authors gratefully acknowledge financial assistance from the N.A.T.O. *(fellowship to JMN)* and CNRS, France *(GM and JMN)*. We would also like to thank the S.E.R.C., U.K. *(SVL and KD)* and the BP research endowment *(to SVL)* at Cambridge, for generous support.

REFERENCES AND NOTES

1. Author to whom correspondence should be addressed. On study leave at the University Chemical Laboratory, Cambridge University, U.K., from the Laboratoire de Pharmacognosie, Unité Associée au CNRS no 492, Faculté de Pharmacie, 51, Rue Cognacq-Jay , F-51096 Reims Cédex. France.

2. Pickett, J.A., *Chem. Brit*, **1988**, *24*, 137; Warthen, D.J.; Morgan, E.D., *'Insect Feeding Deterrents'*, Chapter 2 in Handbook of Natural Pesticides, Vol VI, Morgan, E.D. and Mandava, N.B., Eds.; CRC Press, Boca Raton, **1990**, p.23.

3. van Beek, T.A.; ed Groot, Ae., *Rec. Trav. Chim. Pays-Bays*, **1986**, *105*, 513; Gonzalez, R.H.; Buyckx, E.J.; Brader, L., 'Prospects for New Control Practices of Agricultural Pests in Developing Countries' in 'Natural Products and the Protection of Plants', Marini-Bettolo, G.B., Ed.; *Pontificiae Academiae Scientiarium Scripta Varia*, **1977**, *41*, 687; Rembold, H., *Trans. Bose Res. Inst.*, **1980**, *43*, 53; Lewis, W.J., 'Semiochemicals: Their Role with Changing Approaches to Pest Control', Chapter 1 in *'Semiochemicals: Their Role in Pest Control'*, Norland, D.A.; Jones, R.L. and Lewis W.J., Eds.; John Wiley, New York, **1981**, p.3.

4. Jones, C.G.; Firn, R.D., *Phil. Trans. Roy. Soc. Lond. B*, **1991**, *333*, 273; Williams, D.H.; Stone, M.J.; Hanck, P.R.; Rahman, S.K., *J. Nat. Prod.*, **1989**, *52*, 1189.

5. Munakata, K., *Pure Appl. Chem.*, **1975**, *42*, 57; Ley, S.V.; Toogood, P.L., *Chem. Brit*, **1990**, *26*, 31; Warthen Jr., J.D., *'Azadirachta indica: A Source of Insect Feeding Inhibitors and Growth Regulators'*, U.S. Dept. Agric. Rev. Man. ARM-Ne-4, **1979**.

6. Schmutterer, H.; Rembold, H., *Z. Angew. Entomol*, **1980**, *89*, 179.; 'Natural Pesticides from the Neem Tree (*Azadirachta indica A. Juss*)', and Other Tropical Species', Schmutterer, H. and Ascher, K.R.S., Eds.; GTZ Press, FRG, **1984**; *ibid* **1987** ; Rembold, H.; Sharma, G.K.; Czoppelt, Ch.; Schmutterer, H., *Z. Angew. Entomol.*, **1982**, *93*, 12; Yamasaki, R.B.; Klocke, J.A., *J. Agric. Food. Chem.*, **1987**, *35*, 467 ; Simmonds, M.S.J.; Blaney, W.M.; Ley, S.V.; Anderson, J.C.; Toogood, P.L., *Entomol. exp. appl*, **1990**, *55*, 169.

7. Jacobson, M., 'Pharmacology and Toxicity of Neem' in *'1988 Focus on Phytochemical Pesticides. Vol. 1., The Neem Tree'*, Jacobson, M., Ed.; C.R.C. Press Inc., **1989**, p.133

8. Ley, S.V.; Denholm, A.A.; Wood, A., *Nat. Prod. Reports*, **1993**, 109.

9. Rembold, H., 'Isomeric Azadirachtins and their Mode of Action' in *'1988 focus on Phytochemical Pesticides. Vol 1., The Neem Tree'*; Jacobson, M., Ed.; C.R.C. Press Inc., **1989**, p.47

10. Ley, S.V.; Santafianos, D.; Blaney, W.M.; Simmonds, M.S.J., *Tetrahedron Lett*, **1987**, *28*, 221; Ley, S.V.; Anderson, J.C.; Blaney, W.M.; Lidert, Z.; Morgan, E.D.; Robinson, N.G.; Santafianos, D.; Simmonds, M.S.J.; Toogood, P.L., *Tetrahedron Lett*, **1989**, *30*, 5175; Blaney, W.M.; Simmonds, M.S.J.; Ley, S.V.; Anderson, J.C.; Toogood, P.L., *Entomol. exp. appl*, **1990**, *55*, 149.

11. Blaney, W.M.; Simmonds, M.S.J.; Ley, S.V.; Anderson, J.C.; Smith, S.C.; Wood, A., *Pestic. Sci.*, **1994**, *40*, 169.

12. Kubo, I.; Matsumoto, T.; Matsumoto, A., *Tetrahedron Lett*, **1984**, *25*, 4729; Kubo, I.; Matsumoto, A.; Matsumoto, T.; *Tetrahedron*, **1986**, *42*, 489.

13. Klenk, A.; Bokel, M.; Kraus, W., *J. Chem. Soc., Chem. Commun*, **1986**, 523.

14. Taylor, D.A.H., *Tetrahedron*, **1987**, *43*, 2779.

15. Butterworth, J.H.; Morgan, E.D.; Percy, G.R., *J. Chem. Soc., Perkin Trans. 1*, **1972**, 2445.

16. Butterworth, J.H.; Morgan, E.D., *J. Insect Physiol*, **1971**, *17*, 969.

17. Zanno, P.R.; Muira, I.; Nakanishi, K.; Elder, D., *J. Am. Chem. Soc*, **1975**, *97*, 1975.

18. Kraus, W.; Bokel, M.; Klenk, A.; Pöhnl, H., *Tetrahedron Lett*, **1985**, *26*, 6435.

19. Bilton, J.N.; Broughton, H.B.; Ley, S.V.; Lidert, Z.; Morgan, E.D.; Rzepa, H.S.; Sheppard, R.N.; Slawin, A.M.Z.; Williams, D.J., *J. Chem. Soc., Chem. Commun*, **1985**, 969.

20. Broughton, H.B.; Ley, S.V.; Lidert, Z.; Slawin, A.M.Z.; Williams, D.J.; Morgan, E.D., *J. Chem. Soc., Chem. Commun*, **1985**, 46.

21. Turner, C.J.; Tempesta, M.S.; Taylor, R.B.; Zagorski, M.G.; Termini, J.C.; Schroeder D.R.; Nakanishi, K., *Tetrahedron*, **1987**, *43*, 2789; Bilton, J.N.; Broughton, H.B.; Jones, P.S.; Ley, S.V.; Lidert, Z.; Morgan, E.D.; Rzepa, H.S.; Sheppard, R.N.; Slawin, A.M.Z.; Williams, D.J., *Tetrahedron*, **1987**, *43*, 2805; Kraus, W.; Bokel, M.; Bruhn, A.; Cramer, R.; Klaiber, I.; Klenk, A.; Nagl, G.; Pöhnl, H.; Sadlo, H.; Vogler, B., *Tetrahedron*, **1987**, *43*, 2817.

22. Smith, D.H.; Gray, N.A.B.; Nourse, I.G.; Crandell, C.W., *Anal. Chim. Acta*, **1981**, *133*, 471.

23. Dubois, J.E.; Carabedian, M.; Dagane, I., *Anal. Chim. Acta*, **1984**, *158*, 217; Christie, B.D.; Munk, M.E., *J. Am. Chem. Soc*, **1991**, *113*, 3750; Funatsu, K.; Susuta, Y.; Sasaki, S., *J. Chem. Inf. Comput. Sci.*, **1989**, *29*, 6;

24. Duncan, B.; Buchanan, B.G.; Hayes-Roth, B.; Lichtarge, O.; Altman, R.; Brinkley, J.; Hewett, M.; Cornelius, C.; Jardetzky, O., *Bull. Magn. Reson*, **1986**, *8*, 111.

25. Gastmans, J.P.; Zurita, J.C.; Sahao, J.R., *Anal. Chim. Acta*, **1989**, *217*, 85.

26. Nuzillard, J.M.; Massiot, G., *Tetrahedron*, **1991**, *47*, 3655.

27. Bax, A.; Summers, M.F., *J. Am. Chem. Soc*, **1986**, *108*, 2093.

28. Bax, A.; Subramanian, S., *J. Magn. Reson*, **1986**, *67*, 565.

29. Müller, L., *J. Am. Chem. Soc*, **1979**, *101*, 4481.

30. Bax, A.; Freeman, R.; Freeman, T.A.; Levitt, M.H., *J. Magn. Reson*, **1981**, *43*, 478.

31. Dodrell, D.M.; Pegg, D.T.; Bendall, M.R., *J. Magn. Reson*, **1982**, *48*, 323.

32. Aue, W.P.; Bartholdi, E.; Ernst, R.R., *J. Chem. Phys*, **1976**, *64*, 2229.

33. Bredt, J.; Thouet, H.; Schmitz, J., *Liebigs Ann. Chem*, **1924**, *437*, 1; Nuzillard, J.M., *J. Chem. Inf. Comput. Sci.*, **1994**, accepted for publication.

(*Received in UK* 25 *July* 1994; *revised* 25 *August* 1994; *accepted* 2 *September* 1994)