Journal of Food and Drug Analysis, Vol. 13, No. 1, 2005, Pages 17-21 藥物食品分析 第十三卷 第一期

Evaluation of the ID Number Method for Structure Generation and Complete Assignment of NMR Spectra of Erythromycin A

SHENG-CHIH CHEN* AND JIN-CHERNG LIEN

Graduate Institute of Pharmaceutical Chemistry, China Medical University, 91 Hsueh-Shih Rd., Taichung City 404, Taiwan, R.O.C.

(Received: April 8, 2004; Accepted: September 13, 2004)

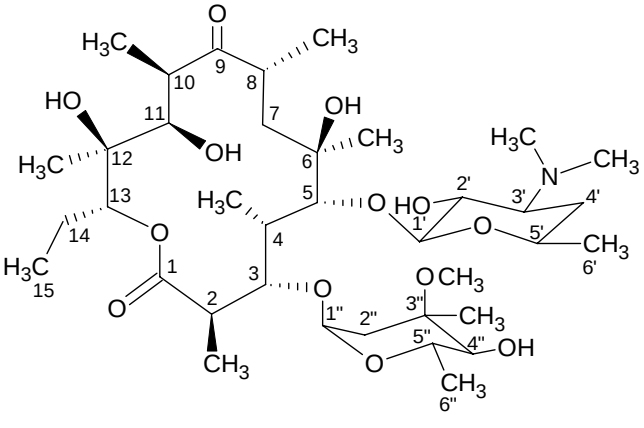
ABSTRACT

The present paper describes the evaluation of the ID number method for the structure generation and complete assignments of proton and carbon NMR spectra of erythromycin A. This study revised the errors found in previous assignments of two carbon signals in the spectra of this compound.

Key words: erythromycin A, structure generation, ID number method, NMR spectra

INTRODUCTION

Erythromycin A is a clinically important macrolide antibiotic which has been used worldwide for at least fifty years[1]. Due to its complex chemical structure and NMR spectra, the structure of erythromycin A has been elucidated by chemical[2] and X-ray crystallographic analysis[3], thereby making it a suitable sample for the evaluation of the identity (ID) number method. The ID number method was established for structure generation and complete assignments of proton and carbon NMR spectra[4].



Erythromycin A

The present paper describes the structure generation and complete assignments of the proton and carbon-13 NMR spectra of erythromycin A to evaluate the ID number method as another approach for using NMR spectra. The result of this study indicated that previous assignments of two carbon signals were in error and thus revised.

* Author for correspondence. Tel: +886-4-22053366 ext. 1008; Fax: +886-4-22055105; E-mail: scchen@mail.cmu.edu.tw

MATERIALS AND METHODS

Forty milligrams of erythromycin A was dissolved in 0.7 mL of $CDCl_3$, and its proton, carbon, DEPT, COSY, HMQC and HMBC spectra were determined on a 400 MHz FT-NMR spectrometer (Bruker DPX-400) with TMS as internal standard. All spectra were determined by use of the built-in programs and parameters. Moreover, an ACD/NMR Processor (ver. 6.0, from Advanced Chemistry Development Inc., Toronto, Canada) was used for peak-picking and identification of the close peaks.

The connections to generate the structure were performed on an IBM-compatible PC with Microsoft Word 2000 purchased from Microsoft Corporation.

RESULTS AND DISCUSSION

A previous study had established an ID number method for structure generation and signal assignment of 1D spectra[4]. In the method, ID numbers were separately given, from low to high field as shown in Table 1, for the signals in the carbon-13 NMR spectrum of erythromycin A. By HMQC, the signals of proton and carbon-13 NMR spectra were correlated and the ID number of each carbon was labeled together with the attached proton to complete the third and fourth columns. Thus, the ^{1}H-^{1}H correlations in COSY and ^{1}H-^{13}C correlations in HMBC could be changed to ^{13}C-^{13}C connectivities of α or β-position. The procedure for the structure generation of erythromycin A then began with a ^{13}C-^{13}C connectivity or fragment derived from a cross peak of COSY. Afterwards, the structure was generated by successive connections as in crossword or jigsaw puzzle. A trial and error process was performed, and the connectivity derived from HMBC would exhibit its allocation effect during the structure generation process.

When the structure was established, the signals of protons and carbons were concomitantly assigned *via* the ID numbers.

The structure of erythromycin A has been established successfully in this study. The resulting structure shown in Figure 1 is consistent with that reported previously[2,3]. The signals in proton and carbon-13 NMR spectra, as indicated in Table 1, were also assigned concomitantly when the structure was established. Therefore, this study shows that the method is useful. The generation process is as follows.

I. *General Considerations of Analytical Data for Structure Generation*

Assuming that erythromycin A is a new compound, the following were considered before starting the structure generation.

Erythromycin A ($C_{37}H_{67}NO_{13}$; M.W. 733.9; hydrogen deficiency index[5] is 5) should have three rings in structure, since the hydrogen deficiency also arose from carbonyl groups with chemical shifts at 221.9 (ketone) and 175.9 (-COO- or –CO-N-) ppm. In addition, there was no unsaturated proton or carbon signal in 1D spectra.

In DEPT, we could find 12 methyl signals from the 13 methyl groups since the signal at 40.3 ppm was correlated to 6 protons in HMQC. Besides, 4 methylene signals at 38.5, 35.0, 28.8, 21.1 ppm and 5 quaternary carbon signals at 221.9, 175.9, 75.0, 74.7, 72.6 were present in the spectrum. Hence, it could be presumed that there were 15 methine groups in erythromycin A. Thus, there was a difference of 2 methine carbons between the presumption and the methine signals observed in DEPT. With the aid of HMQC, the 2 methine signals absent in carbon spectrum and DEPT were found one at 68.9, which was a signal of 2 carbons, and one at 76.9 ppm, which was coincident with that of solvent.

Total of 62 hydrogen atoms were present in NMR data, which was less than that presented in the molecular formula by 5 (Table 1). Therefore, there might be 5 hydroxy groups in the structure. The signals of 14 carbons that were attached to oxygen atoms appeared between 65.5-103.2 ppm, except that of the carbonyl groups. So, 13 oxygen atoms in the molecular formula of erythromycin A were present as ethers, alcohols, ester and ketone in the structure. The signals at 49.5 ppm in carbon-13 spectrum and at 3.31 ppm in proton spectrum were attributed to a methoxy group. Signal at 40.3 ppm with proton chemical shift at 2.30 ppm should be assigned to 2 methyl groups of an *N, N*-dimethylamino group. Because the signals of these 2 methyl groups were correlated to the signal at 65.5 ppm in HMBC, the dimethylamino group should be connected to this carbon. The signals at 3.83 (*s*) and 3.47 (*m*) ppm in proton spectrum were correlated to the 2 carbons at 68.9 ppm in HMQC. Thus, for convenience it could arbitrarily assign one of the carbons to be 3.83 and the other 3.47 as indicated in Table 1 to establish their identities. The signals

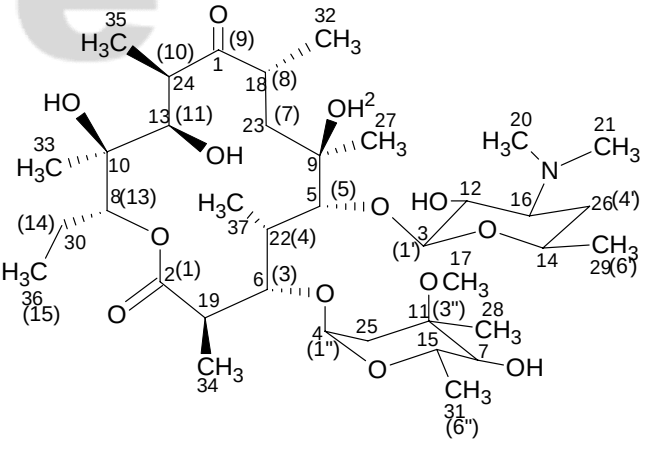


Figure 1. Structure of erythromycin A generated with ID numbers. IUPAC numbering is given in parentheses.

of methyl groups with their multiplicities shown in proton and 2D spectra could also offer useful information for exact connection.

II. *Fragments Generated from COSY*

COSY can provide "hard" connectivity through spin-spin coupling. The connectivities were established by the conversion of the ^1H-^1H correlations in COSY with the aid of HMQC (Table 1). Before starting the structure generation, it is necessary to combine the related connectivities. Thus, fragments generated from COSY were rearranged and combined to form larger fragments as follows:

C^3H-$C^{12}H$-$C^{16}H$-$C^{26}H_2$-$C^{14}H$; C^4H-$C^{25}H_2$; C^5H-$C^{22}H$-$C^{37}H_3$; C^6H-$C^{19}H$-$C^{34}H_3$; C^7H-$C^{15}H$-$C^{31}H_3$; C^8H-$C^{30}H_2$-$C^{36}H_3$; $C^{13}H$-$C^{24}H$-$C^{35}H_3$; $C^{23}H_2$-$C^{18}H$-$C^{32}H_3$.

Each of above connectivities has been confirmed by HMBC, and they were used in structure generation.

III. *Structure Generation*

The structure of erythromycin A was generated step by step as indicated in Figure 2. After normalization, the structure and ID numbers are given in Figure 1.

IV. *Revised Assignments of the Signals Reported Previously*

In the previous report by Ager and Sood[6], complete assignment of the carbon-13 spectrum of erythromycin A had been carried out by using 2D INADEQUATE and DEPT. However, difficulty occurred with the assignments of signals in methyl and methine regions of the spectra due to coincidence or overlapping. For example, the signals at 68.9 (2 × CH), 65.6, 65.5, 40.3 (2 × CH_3), 21.5, 21.4, 21.1, 18.6 and 18.3 ppm were too close to be exactly assigned. In this paper, we used ACD/NMR Processor to pick the peaks and clarify the ambiguities, and also used the data of HMQC and COSY to confirm the assignments. So, as can

be found in Figure 1 and Table 1, the assignments of the 2 signals at 45.1 and 44.9 ppm previously reported were found to be in error and should be revised. As shown in HMQC, the signals at 2.69 and 2.87 ppm were unambiguously correlated to C^{18} and C^{19}, respectively. In COSY, both of the signals of the $C^{23}H_2$ protons at 1.71 and 1.92 ppm coupled separately to that of $C^{18}H$, demonstrating that $C^{18}H$ was vicinal to both $C^{23}H_2$ and C^1, whereas $C^{19}H$ was vicinal to both C^6H and C^2. These connections were exactly confirmed by HMBC.

ACKNOWLEDGEMENTS

This work was supported by the China Medical University, Taichung, Taiwan, R.O.C. (CMC91-PC-01).

Table 1. ^{13}C-^{13}C connectivities derived from NMR spectra data[a]

ID[b]	C-13 (ppm)[b]	DEPT[c]	Attached-proton correlated by HMQC[c,d] (ppm (*J*, Hz))	^{13}C-^{13}C connected by COSY[e]	^{13}C-^{13}C connected by HMBC[e]	Position[f] A	Position[f] B
1	221.9	C^1=O			C1-C13,C23,C24,C32	9	9
2	175.9	C^2=O			C2-C8,C6,15[g],C19,C34	1	1
3	103.2	C^3H	4.40 (*d*, 7.1)	-O-C^3H-$C^{12}H$	C3-C5,C12	1'	1'
4	96.3	C^4H	4.88 (*d*, 4.7)	-O-C^4H-$C^{25}H_2$	C4-C6,15[g],C25,C31	1"	1"
5	83.6	C^5H	3.55 (*d*, 7.7)	-O-C^5H-$C^{22}H$	C5-C3,C6,15,C22,C27,C37	5	5
6	80.0	C^6H	3.99 (*d*, 9.2)	-O-C^6H-$C^{19}H$	C6-C4,C5,C19,C37,C34	3	3
7	78.0	C^7H	3.00 (*br d*, 7.4)	-O-C^7H-$C^{15}H$	C7-C25,C31,C28,C6,15[g]	4"	4"
8	76.9	C^8H	5.03 (*dd*, 10.9,2.0)	$C^{30}H_2$-C^8H-$C^{27}H_3$	C8-C36,C13,C33	13	13
9	75.0	C^9		-O-C^9	C9-C5,C22,23,30[g],C27	6	6
10	74.7	C^{10}		-O-C^{10}	C10-C8,C33	12	12
11	72.6	C^{11}		-O-C^{11}	C11-C4,C17,C25,C28	3"	3"
12	71.0	$C^{12}H$	3.22 (*dd*, 10,7.5)	-O-$C^{12}H$-$C^{16}H$	C12-C26	2'	2'
13	68.9	$C^{13}H$	3.83 (*s*)	-O-$C^{13}H$-$C^{24}H$	C13,14-C8,C29,C35,C33	11	11
14	68.9	$C^{14}H$	3.47 (*m*)	$C^{14}H$-$C^{26}H_2$,$C^{29}H_3$	C14-	5'	5'
15	65.6	$C^{15}H$	4.00 (*dm*)	-O-$C^{15}H$-$C^{31}H_3$	C15,16[g]-C4,C12,C20,21, C26,C31	5"	5"
16	65.5	$C^{16}H$	2.46 (*m*)	$C^{16}H$-$C^{26}H_2$	C16-	3'	3'
17	49.5	$C^{17}H_3$	3.31 (*s*)	O$C^{17}H_3$	C17-	3"-OMe	3"-OMe
18	45.1	$C^{18}H$	2.69 (*m*)	$C^{23}H_2$-$C^{18}H$-$C^{32}H_3$	C18-	2	8
19	44.9	$C^{19}H$	2.87 (*m*)	$C^{19}H$-$C^{34}H_3$	C18,19[g]-C6,15[g],C23,30[g], C34,C32,C23,26	8	2
20	40.3	$C^{20}H_3$	2.30 (*s*)		C20,21-C16,C20,21	3'-NMe	3'-NMe
21	40.3	$C^{21}H_3$	2.30 (*s*)		C21-	3'-NMe	3'-NMe
22	39.4	$C^{22}H$	1.97 (*m*)	$C^{22}H$-$C^{37}H_3$	C22-C6,15[g],C19,C37	4	4
23	38.5	$C^{23}H_2$	1.92 (*m*), 1.71 (*d*, 14.6)		C23-C5,C27,C32,C18,OH[2]	7	7
24	37.9	$C^{24}H$	3.08 (*q*, 6.8)	$C^{24}H$-$C^{35}H_3$	C24-C35,33[g]	10	10
25	35.0	$C^{25}H_2$	2.36 (*d*, 15.2), 1.56 (15.2,5.0)		C25-C28	2"	2"
26	28.8	$C^{26}H_2$	1.68 (*m*), 1.22	$C^{26}H$-$C^{29}H_3$	C26-C29	4'	4'
27	26.9	$C^{27}H_3$	1.46 (*s*)		C27-C5,C23,30[g]	6-Me	6-Me
28	21.5	$C^{28}H_3$	1.23 (*s*)		C28-	3"-Me	3"-Me
29	21.4	$C^{29}H_3$	1.22 (*d*, 6)		C29-C25	6'	6'
30	21.1	$C^{30}H_2$	1.22 (*m*), 1.90 (*m*)		C30-C36	14	14
31	18.6	$C^{31}H_3$	1.27 (*d*, 6)		C31-	6"	6"
32	18.3	$C^{32}H_3$	1.15 (*d*, 7.9)		C32-C18,C23	8-Me	8-Me
33	16.2	$C^{33}H_3$	1.12 (*s*)		C33-C8	12-Me	12-Me
34	16.0	$C^{34}H_3$	1.17 (*d*, 7.4)		C34-C19	2-Me	2-Me
35	12.0	$C^{35}H3$	1.14 (*d*, 7.6)		C35-C13,C24,C23,C32,35[g]	10-Me	10-Me
36	10.7	$C^{36}H_3$	0.84 (*t*, 7.3)	$C^{30}H_2$-$C^{36}H_3$	C36-C8,C23,30[g]	15	15
37	9.2	$C^{37}H_3$	1.10 (*d*, 7.7)		C37-C6,15[g],C22	4-Me	4-Me
		O^1H	3.13 (*br*)				
		O^2H	1.80 (*br*)				6-OH
Total 37		Total 62					

[a]Spectra were determined at 400 MHz in $CDCl_3$.

[b]ID numbers were given for carbon signals from low to high field.

[c]DEPT and HMQC were used to characterize the carbon signal as primary, secondary, tertiary or quaternary carbon. In DEPT column, each C are labeled with its ID number for identity.

[d]HMQC was used to correlate the attached proton and to find out a coincident carbon signal. The multiplicity of proton signal is represented by *s*, *d*, *t* and *m* for singlet, doublet, triplet and multiplet, respectively.

[e]Cross peaks in COSY and HMBC were used to change ^1H-^1H and ^1H-^{13}C correlations to ^{13}C-^{13}C connections.

[f]IUPAC numbering is used; A: data from reference 6; B: data from the present study.

[g]Ambiguous connectivities are represented as C6,15, C22,23,30, *etc.* in the table and were used by pertinent selection in structure generation.

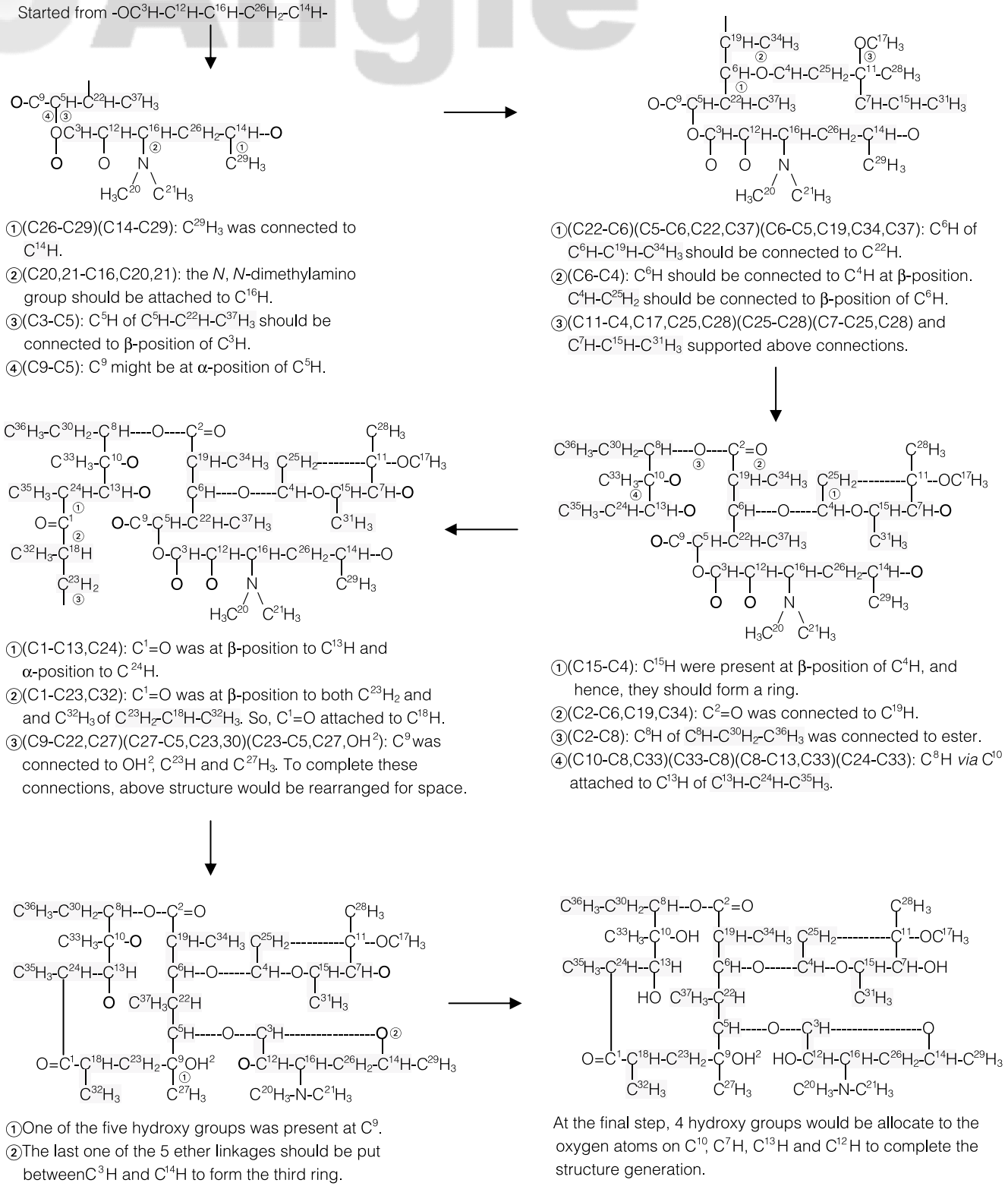


Figure 2. Connection and structure generation of erythromycin A. HMBC connectivities are given in parentheses.

REFERENCES

1. Merck Research Laboratories. 2001. "The Merck Index". 13th ed. pp. 654-655. Merck & Co., Inc. Whitehouse Station, New Jersey, U. S. A.
2. Wiley, P. F., Gerzon, K., Flynn, E. H., Sigal, M. V., Jr., Weaver, O., Quarck, O. C., Chauvette, R. R. and Monahan, R. 1957. Erythromycin. X. Structure of erythromycin. J. Am. Chem. Soc. 79: 6062-6070.

3. Harris, D. R., McGeachin, S. G. and Mills, H. H. 1965. The structure and stereochemistry of erythromycin A. Tetrahedron Lett. 679-685.
4. Chen, S. C. 2002. Another choice for structure elucidation by using NMR spectra: structure generation by ID number method. Chemistry (The Chinese Chem. Soc., Taipei) 60: 153-162.
5. Silverstein, R. M., Bassler, G. C. and Morrill, T. C. 1991. "Spectrometric Identification of Organic Compounds". 5th ed. p. 12. John Wiley & Sons. New York, U. S. A.
6. Ager, D. J. and Sood, C. K. 1987. The complete, unambiguous assignment of the ^{13}C NMR spectrum of erythromycin A. Magn. Reson. Chem. 25: 948-954.