STRUCTURE OF CLEOMISCOSIN B, A COUMARINO-LIGNOID OF CLEOME VISCOSA SEEDS

Anil B. Ray and Sunil K. Chattopadhyay

Department of Medicinal Chemistry, I. M. S., Banaras Hindu University, Varanasi, India

and

Chohachi Konno and Hiroshi Hikino

Pharmaceutical Institute, Tohoku University, Aoba-yama, Sendai, Japan

Abstract —— A new coumarino-lignoid cleomiscosin B has been isolated from the seeds of Cleome viscosa and its structure has been elucidated as I based on chemical and spectroscopic evidence which, in turn, has brought on the revision of the structure of cleomiscosin A as IV.

We recently made a phytochemical investigation on the seeds of Cleome viscosa Linné (syn. C. icosandra Linné, Capparidaceae) used as a folklore medicine in India and reported the isolation of a coumarino-lignoid, cleomiscosin A for which the structure I was advanced.[1]

A thorough search for chemical constituents from the seed extract has now led to the isolation of a new congener which is named as cleomiscosin B. The present paper describes the structure determination of cleomiscosin B and the concurrent structure revision of cleomiscosin A.

Cleomiscosin B, m.p. 274°, $C_{20}H_{18}O_8$ (MS m/e 386.1014, M^+), and cleomiscosin A showed striking resemblance in spectral properties, indicating a close structural similarity between the two molecules. Thus the ^{13}C NMR spectrum of cleomiscosin B disclosed the presence of seven aliphatic carbons (CH_3-O x 2, $-CH_2$-O x 1, >CH-O x 2, -CH=CH- x 1), twelve aromatic carbons (CH x 4, C x 2, C-O x 6) and one carbonyl carbon like its congener, cleomiscosin A (Table I).

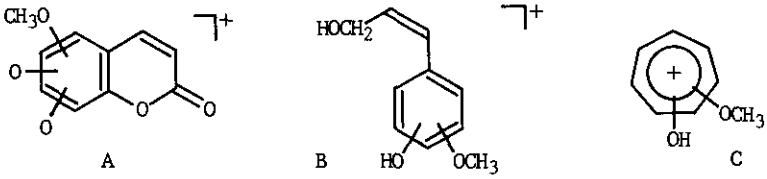
The oxygen functions were first examined. Cleomiscosin B was found to be phenolic in nature (IR band at 3510 cm^{-1}, positive phosphomolybdic acid test) and gave the non-phenolic monomethyl ether (II), m.p. 212-213°, $C_{21}H_{20}H_8$ (MS m/e 400, M^+). Since the ether (II) still exhibited an IR band at 3530 cm^{-1} (hydroxyl), cleomiscosin B was acetylated to furnish the diacetate (III), m.p. 174°, $C_{24}H_{22}O_{10}$ (MS m/e 470, M^+), which showed no IR band due to hydroxyl. Cleomiscosin B was thus proved to bear one phenolic and one alcoholic hydroxyl grouping.

Further, cleomiscosin B displayed a UV maximum at 328 nm with humps at 285 and 230 nm, showed IR bands at 1720 and 1610 cm^{-1} and disclosed two ^1H NMR 1H doublets at δ 6.21 and 7.51 (J 10 Hz), indicating that the same coumarin moiety is present in cleomiscosin B as in cleomiscosin A. The ^1H NMR spectrum also indicated the presence of two methoxyl groups (δ 3.71 and 3.83) in cleomiscosin B. Thus six out of eight oxygen atoms were accounted for and the remaining two oxygens were to be involved in cyclic ether formation in analogy with cleomiscosin A. The structure similarity between two cleomiscosins became further manifest from mass spectral analysis because both the substances showed peaks due to the identical fragment ions (m/e 208.0283 for $C_{10}H_8O_5^+$ (A), m/e 180.0769 for $C_{10}H_{12}O_3^+$ (B) and m/e 137 for $C_8H_9O_3^+$ (C) in cleomiscosin B). The fragment ion A indicated that cleomiscosin B contains a trioxygenated coumarin moiety and that one of these

Table I. Carbon-13 shieldings in cleomiscosin B and ralated substances (δ)

	cleomiscosin B (I) (C_5D_5N)	cleomiscosin B diacetate (III) ($CDCl_3$)	cleomiscosin A (C_5D_5N)	cleomiscosin A diacetate ($CDCl_3$)
C-2	160.7 s	160.4 s	160.8 s	160.4 s
C-3	113.8 d	114.3 d	113.6 d	114.4 d
C-4	144.4 d	143.5 d	144.5 d	143.5 d
C-5	101.2 d	100.8 d	101.1 d	100.5 d
C-6	146.2 s	145.7 s	146.3 s	145.8 s
C-7	138.1 s	136.3 s	138.4 s	136.9 s
C-8	133.2 s	132.1 s	133.0 s	131.7 s*
C-9	139.4 s	140.7 s	139.3 s	140.8 s
C-10	111.8 s	111.8 s	111.9 s	111.9 s
C-1'	127.5 s	133.5 s	127.5 s	133.5 s*
C-2'	112.3 d	111.3 d	112.3 d	111.5 d
C-3'	150.1 s	151.6 s	150.0 s	151.7 s
C-4'	149.1 s	138.8 s	149.0 s	138.8 s
C-5'	116.5 d	123.2 d	116.6 d	123.3 d
C-6'	121.7 d	119.8 d	121.7 d	119.9 d
C-7'	77.1 d	76.0 d	77.5 d*	76.7 d
C-8'	80.2 d	75.6 d	79.9 d*	75.1 d
C-9'	61.1 t	62.4 t	60.7 t	62.4 t
OCH_3	55.9 q	56.0 q	55.8 q	56.0 q
OCH_3	56.1 q	56.4 q	56.2 q	56.3 q
CH_3CO		20.6 q		20.6 q
CH_3CO		20.6 q		20.6 q
CH_3CO		168.6 s		168.5 s
CH_3CO		170.2 s		170.2 s

*The assignments were reversed in the previous paper.[1]



oxygen functions is a methoxyl. This was further corroborated by the occurrence of a 1H singlet at δ 6.53 for an insulated aromatic hydrogen which, however, showed a long range coupling with the signal at δ 7.57 for the C-4 hydrogen in the coumarin nucleus and an intramolecular nuclear Overhauser effect (19%) with the methoxy hydrogens (δ 3.92). The signal at δ 6.53 must, therefore, be due to the C-5 hydrogen and a methoxyl group must be at C-6 of the coumarin moiety. Hence the two oxide linkages in cleomiscosin B are at C-7 and C-8 or in other words the substitution pattern of the coumarin moiety in cleomiscosin A and B are identical. In conformity with this observation, *the* ^{13}C *NMR parameters for the coumarin carbons* in cleomiscosin B and its diacetate (III) fitted with those in cleomiscosin A and its diacetate (Table I).

Again in the mass spectrum of cleomiscosin B, the fragment ions B and C suggested that cleomiscosin B has a phenylpropane unit as cleomiscosin A. The ^1H NMR spectrum of the diacetate (III) exhibited aromatic hydrogen signals at δ 6.98 (1H doublet, J 8 Hz), 7.00 (1H singlet) and 7.05 (1H doublet, J 8 Hz) which were comparable with those in cleomiscosin A diacetate,[1] a fact which showed the presence of a 3,4-dioxygenated phenyl side chain in cleomiscosin B. The mass spectral peak due to the ion C pointed to two oxygen functions to be allocated to a hydroxyl and a methoxyl. Comparison of the predicted values, calculated by adding the acetylation parameters[2] to the observed values for the phenyl carbons in cleomiscosin B, with the observed values in

cleomiscosin B diacetate (III) showed that the calculated values were discrepant from the observed values if cleomiscosin B had the 3-hydroxy-4-methoxy phenyl side chain but consistent with the observed values except for the two carbons bearing oxygen functions if cleomiscosin B had the 4-hydroxy-3-methoxy phenyl side chain (Table II). Discrepancy between calculated and observed

Table II. Carbon-13 shieldings of the phenyl side chain in cleomiscosin B and its acetate

	C-1'	C-2'	C-3'	C-4'	C-5'	C-6'
in the case of 3-hydroxy-4-methoxy derivative						
cleomiscosin B (obs)	127.5	116.5	149.1	150.1	112.3	121.7
the diacetate (calc)	125.5	126.4	142.8	160.4	110.3	129.6
the diacetate (obs)	133.5	119.8	138.8	151.6	111.3	123.2
in the case of 4-hydroxy-3-methoxy derivative						
cleomiscosin B (obs)	127.5	112.3	150.1	149.1	116.5	121.7
the diacetate (calc)	135.4	110.3	160.4	142.8	126.8	119.7
the diacetate (obs)	133.5	111.3	151.6	138.8	123.2	119.8

values of adjacent aromatic carbons carrying oxygen functions is, however, known.[2] The side chain of cleomiscosin B was thus deduced to be 4-hydroxy-3-methoxyphenyl as in cleomiscosin A.

The apliphatic portion was next subjected to examination. In the ^{1}H NMR spectra of the derivatives (II and III), methine doublets (J 8 Hz) appeared at δ 5.12 and 5.04 which were in accord with those (δ 5.05 and 5.03) in the ethyl ether and the diacetate of cleomiscosin A. The counterparts of the 1H doublets occurred at δ 4.0 and 4.4 in the derivatives (II and III), revealing that the second methine is juxtaposed with the first one. There were also two 1H signals (δ 3.53 and 4.0 in the ether (II) and δ 4.12 and 4.4 in the diacetate (III)) as the AB part of ABC system. These signals were attributed to a methylene which should be attached to the second methine because no other position was left for this group. It was demonstrated that the signals for the methylene underwent downfield shifts both in the ^{1}H NMR (Δδ 0.4–0.7 ppm) and ^{13}C NMR (Δδ 1.3 ppm) spectra on passing from the ether (II) to the diacetate (III), indicating that the methylene is attached to a hydroxyl and consequently the remaining two methines bear ethereal oxygens. The phenylpropane moiety of cleomiscosin B was thus concluded to have the same arrangement as cleomiscoisn A. This was further supported by the fact that the chemical shifts of the ^{13}C NMR signals of the phenylpropane carbons in cleomiscosin B and its diacetate (III) were comparable with those in cleomiscosin A and its diacetate (Table I).

Hence the difference between cleomiscosin A and B should be in the stereochemistry at C-7' and C-8' and/or in the mode of attachment of the coumarin part with the phenylpropane part.

The possibility of cleomiscosin B being a stereoisomer of cleomiscosin A was rule out from the measurement of the coupling constant between the ^{1}H NMR signals for the C-7' and C-8' methine hydrogens which was found to be 8 Hz in cleomiscosin B and its derivatives (II and III). This value agrees with that in cleomiscosin A and indicated that the phenyl and the hydroxymethyl are trans-oriented. Again, cleomiscosin B showed no optical activity and so it is racemic like cleomiscosin A. Obviously, therefore, cleomiscosin B is a position isomer of cleomiscosin A.

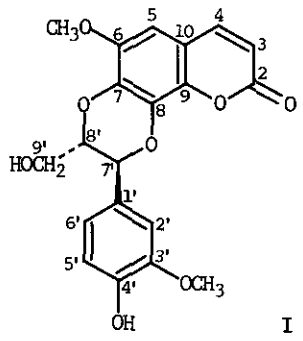
In the structure study of cleomiscosin A,[1] the mode of linkage of the coumarin moiety with the phenylpropane residue was settled by measurements of the ^{13}C–^{1}H spin couplings using cleomiscosin A diacetate. Thus irradiation at the C-7' hydrogen signal at δ 5.03 converted the carbon signal at δ 133.5, formerly assigned to C-8, from a double doublet (J 8 and 3 Hz) into a doublet (J 8 Hz) and irradiation at the C-8' hydrogen signal at δ 4.4 sharpened the doublet (J 8 Hz) at δ 136.9 attributed to C-7. Cleomiscosin A was thus concluded to have the structure I.

In the diacetate (III), when the C-7' hydrogen signal at δ 5.04 was irradiated, the double

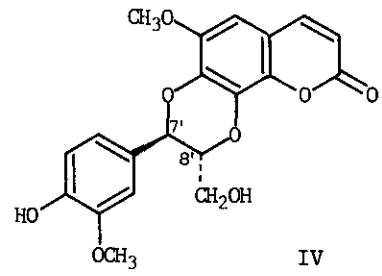
doublet (J 8 and 3 Hz) at δ 133.5 changed to a doublet (J 8 Hz), and when the C-8' hydrogen signal at δ 4.4 was irradiated, the doublet (J 8 Hz) sharpened significantly. These results led to the contrary conclusion that cleomiscosin A and B possess the same structure in spite of the fact that they are different substnaces. This anomaly was considered to be due to the inappropriate assignement of the double doublet (J 8 and 3 Hz) at δ 133.5 to C-8. The carbon whose ^{13}C NMR signal was equally altered on irradiation of the C-7' hydrogen signal in both the diacetates might as well be C-1' provided that the signal at δ 133.5 is assigned to this carbon. This carbon may couple with the C-5' hydrogen signal in addition to the C-7' hydrogen signal and its splitting pattern, a doublet of doublets, may thus be rationalized. Consequently, the ^{13}C NMR signal at δ 133.5 was assigned to C-1' and not to C-8 and the signal originating from C-8 was to occur at δ 131.7 in cleomiscosin A diacetate and at δ 132.1 in cleomiscosin B diacetate (III). The reason for the improper assignment which was made earlier is due to the assumption that the C-8 signal should not be significantly displaced on passing from cleomiscosin A to its diacetate.

Based on the above assignments, the mode of fusion of the coumarin part and the phenylpropane part in cleomiscosin A and B was reexamined. In cleomiscosin A diacetate, when the C-7' and C-8' hydrogen signals (δ 5.03 and 4.4, respectively) were irradiated, the carbon signals for C-7 (δ 136.9) and C-8 (δ 131.7), respectively, showed significant sharpening. It was thus demonstrated that the structure of cleomiscosin A should be revised to IV. In cleomiscosin B diacetate (III), irradition at the C-7' and C-8' hydrogen signals (δ 5.04 and 4.4, respectively) induced sharpening of the C-8 and C-7 carbon signals (δ 132.1 and 136.3), respectively, cleomiscosin B was thus proved to have structure I.

It is of interest to note that the process of the coupling of the coumarin part with the phenylpropane unit is not selective in the biosynthesis of cleomiscosin A and B as judged from the fact that these position isomers, cleomiscosin A and B, co-occur in the same plant and that both are racemic.



I



IV

After the publication of the paper concerning the structure of cleomiscosin A,[1] it became to our knowledge that a similar substance was isolated from the mineral-stained wood of Acer saccharum Marsh. (Aceraceae) and its structure was alleged to be I based on some chemical evidence and biogenetic speculation.[3] After the isolation of the two position isomers, cleomiscosin A and B, the biogenetic speculation does not seem to be convincing to determine the mode of fusion of the coumarin moiety with the phenylpropane unit. It is not certain whether the substance from A. saccharum is to be identical with cleomiscosin A or B.

Recently, we obtained a sample of the methyl ether of cleosandrin isolated from the same plant source (the seeds of C. icosandra),[4] which was identified as cleomiscosin A methyl ether.

REFERENCES

1) A. B. Ray, S. K. Chattopadhyay, C. Konno and H. Hikino, Tetrahedron Letters, 1980, **21**, 4477.
2) M. Tamada, K. Endo and H. Hikino, Heterocycles, 1979, **12**, 783.
3) J. F. Manville and N. Lavintin, Environ. Can. For. Serv., 1974, **30**, 3.
4) A. G. R. Nair, Indian J. Chem., 1979, **17B**, 438.

Received, 12th August, 1981