JOURNAL OF NATURAL PRODUCTS

Volume 60, Number 3 March 1997

Rapid Communications

Crystal Structure of a 1:1 Complex of Natural Diterpenoids: Absolute Configurations and Unambiguous NMR Spectral Assignments of Neoangustifolin and Epinodosinol

Han-Dong Sun,[†] Sheng-Xiang Qiu,[‡] Long-Ze Lin,[‡] Rong-Ping Zhang,[†] Yasheen Zhou,[‡] Qi-Tai Zheng,[§] Michael E. Johnson,[‡] Harry H. S. Fong,[‡] Norman R. Farnsworth,[‡] and Geoffrey A. Cordell*[,‡]

Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, People's Republic of China, Institute of Materia Medica, Chinese Academy of Medical Science, Beijing 100050, People's Republic of China, and the Program for Collaborative Research in the Pharmaceutical Sciences, Department of Medicinal Chemistry and Pharmacognosy, Center for Pharmaceutical Biotechnology, College of Pharmacy, University of Illinois at Chicago, 833 Wood Street, Chicago, Illinois 60612

Received December 6, 1996[⊗]

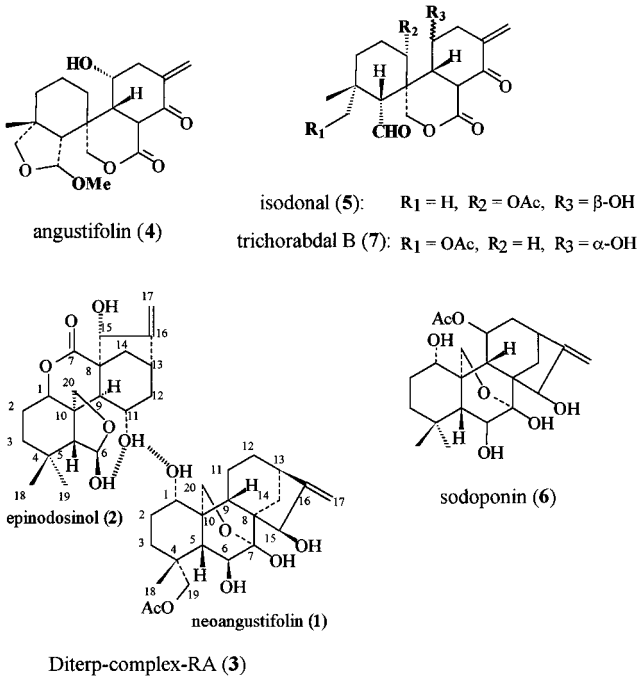
Abstract: Diterp-Complex-RA (DCRA), a unique, inseparable, natural hydrogen-bonded 1:1 complex consisting of a novel diterpenoid neoangustifolin and a known diterpenoid epinodosinol, was isolated from *Rabdosia angustifolia*, along with daucosterol and the known diterpenes sodoponin, trichorabdal A, and cirsiliol. The binding mode of the complex and the absolute configuration of the two substrates were elucidated through X-ray crystallography. Unambiguous NMR data assignments were achieved by a combination of various 2D-NMR techniques. It was clearly demonstrated that DCRA consists of two different diterpene substrates that are bound together by an intermolecular hydrogen bond and the very good hydrophobic close approach of mutual matching surfaces. In addition, intramolecular hydrogen bonds also contribute to the stabilization of the complex.

The diterpene constituents, often referred as bitter principles, of *Rabdosia* (Labiateae) plants have attracted chemists and pharmacologists for a long time because of their structural diversity and significant biological activities, with particular interest centered on their antitumor activity.[1] *Rabdosia angustifolia* (Dunn) Hara, distributed mainly in the Yunnan Province of China, has been used for medical purposes against pyrexia, abdominal distension, and inflammation since ancient times, and the isolation of β-sitosterol and two B-*seco*-kaurene-type diterpenes, angustifolin (**4**) and isodonal (**5**), has been reported.[2,3] In the course of a search for additional, novel diterpenes for biological evaluation from the title plant, an alcoholic extract of the leaves was partitioned successively with petroleum ether, EtOAc, and *n*-BuOH. The EtOAc part was subjected to silica gel, Sephadex LH-20, reversed-phase Rp-8 columns and finally preparative HPLC, resulting in the isolation of a 1:1 complex of natural diterpenoids, tentatively named Diterp-Complex-RA (DCRA, **3**), along with the known compounds daucosterol, sodoponin (**6**), trichorabdal B (**7**), and cirsiliol (**8**).[4] The binding mode and absolute configuration of both subunits of the complex were determined by X-ray crystallography. In addition, unambiguous assignments of the NMR spectra were made by a combination of various 1D- and 2D-NMR techniques.

DCRA (**3**) was obtained as colorless crystals from acetone: mp 213−5 °C, $[\alpha]_D$ −46.5° (*c* 0.16, MeOH); UV λ_{max}^{EtOH} nm (log ϵ) 217 (4.14), 222 (4.09), 227 (4.26); IR ν_{max} (cm^{-1}, KBr) 3474, 3315, 1738, 1724, 1660, 1235; FABMS (negative) *m/z* 771 [M(**3**) − 1]$^+$, 407 [M(**1**) − 1]$^+$, 363 [M(**2**) − 1]$^+$. It was shown to be homogeneous on TLC using several different solvent systems and by HPLC analysis. However, the ^{1}H- and ^{13}C-NMR spectra taken in DMSO-d_6, methanol-d_4, and C_5D_5N initially appeared to be unduly complex. Closer examination of

* To whom correspondence should be addressed. Tel. (312) 996-7245. Fax. (312) 996-7107. E-mail: cordell@pcog8.pmmp.UIC.edu.
[†] Kunming Institute of Botany.
[‡] The University of Illinois at Chicago.
[§] Institute of Materia Medica.
[⊗] Abstract published in *Advance ACS Abstracts*, March 1, 1997.

S0163-3864(97)00001-3 CCC: $14.00



the NMR spectra, however, disclosed that they were derived from the presence of two diterpene substrates in the same abundance. A single-crystal X-ray analysis (Figures 1−4) of DCRA[5] revealed that it consisted of epinodosinol (**2**) and a new diterpene designated as neoangustifolin (**1**), 19-acetoxy-1α,6*β*,7*β*,15*β*-tetrahydroxy-7α,20-epoxy-*ent*-kaur-16-ene.

Although there is some overlap of resonances in the NMR spectra of DCRA, all of the ^{1}H and ^{13}C-NMR data could be unambiguously assigned (Table 1) by a combination of DEPT, DQF-COSY, HOHAHA, ROESY, HETCOR, and FLOCK NMR experiments.

The complete ^{1}H NMR data assignment of epinodosinol (**2**) was initiated by selecting the characteristic hemiacetal proton H-6, resonating at δ 5.66, which appeared as a singlet because of the near 90° dihedral angle with H-5. The signals for H_3-19 and H_3-18 could be differentiated by ROESY, since cross peaks were observed between H_3-19/H-20a, H_3-18/H-5, and H_3-18/H-6. Moreover, H-15 showed long-range coupling with each of the terminal coupled methylene protons H_2-17.

The methylene protons (by DEPT, HETCOR) at 1.84 ppm, assigned to H_2-2, were scalar coupled with 1-H, which was correlated with H-5 by roe's, and the methylene protons at 1.18 ppm, resulting in the assignment

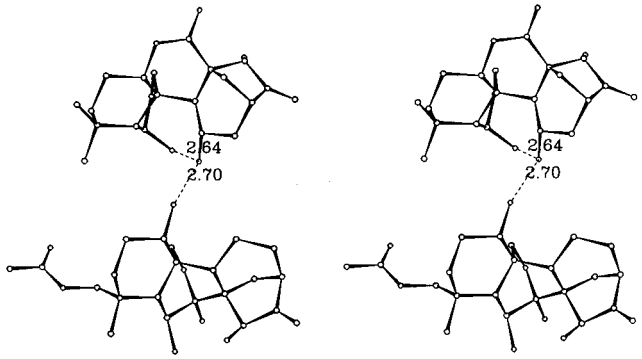


Figure 1. Stereoview of the X-ray structure of DCRA (**3**) showing the binding site and inter- and intramolecular hydrogen bonds (Å) with dotted lines.

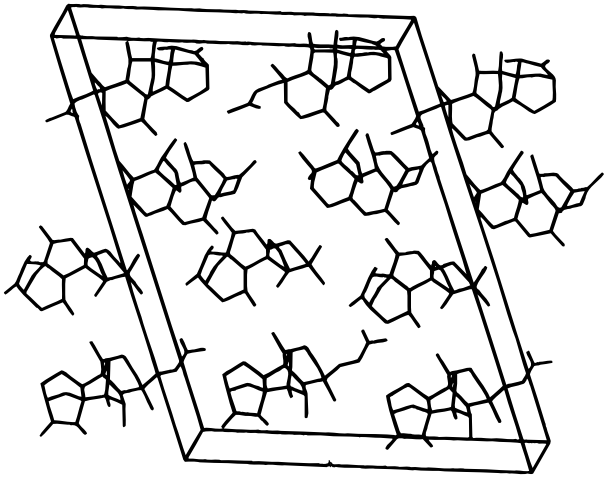


Figure 2. Drawing of the packing mode of the DCRA molecules along *c* axis.

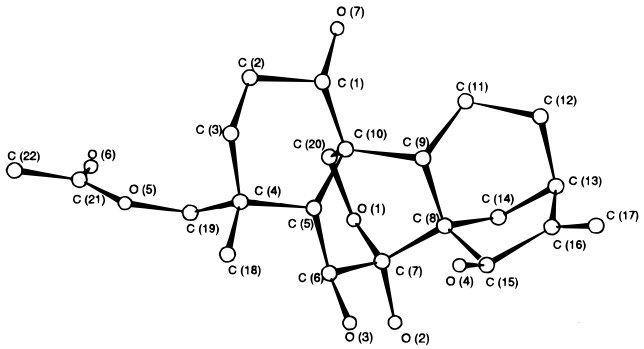


Figure 3. ORTEP drawing of the crystal structure of neoangustifolin (**1**).

of 1.18 ppm to H_2-3, which was further confirmed by the noticeable long-range coupling with 1-H in the HOHAHA spectrum. Analysis of the coupling pattern of the HOHAHA coupling fragment $C_9-C_{11}-C_{12}-C_{13}-C_{14}$ of **2** resulted in the assignment of H-9, H-11, H_2-12, H-13, and H_2-14.

With regard to the ^{13}C-NMR data, the assignment of hydrogen-bearing carbons was made by taking advantage of HETCOR. Thus, the remaining quaternary carbon signals of C-4, C-8, and C-10, except for C-7 and C-16, which could be assigned readily by taking account of the structure, were achieved by FLOCK, a long-range ^{1}H−^{13}C correlation spectrum, in which the deductive cross peaks were detected between C-4/H-5 (two bond coupling), C-4/H-6, C-8/H-9 (two bond coupling), and C-8/H-11, as well as C-4/H_3-18 and C-4/H_3-19. The unambiguous NMR data assignment of epinodosinol is presented here for the first time, although it was isolated previously from *Isodon japonica*.[6]

The complete assignment of neoangustifolin (**1**) was initiated by selecting H-15 at δ 5.13 as the starting point, which showed long-range coupling with the terminal methylene protons at 5.42 and 5.17 ppm, respectively. Two coupling fragments $C_1-C_2-C_3$ and $C_{11}-C_{12}-C_{13}$ were deduced from the HOHAHA spectrum, resulting in the assignment of these protons by analyzing the coupling pattern with the aid of COSY and ROESY correlations. Moreover, ROE correlations could also be applied to clarify the assignments of H_2-19 and H_2-20, since roe's were observed between the proton pairs of H_2-20/H-2α and H_2-20/H-11α. The ^{13}C signal assignments of **1** were achieved as described for

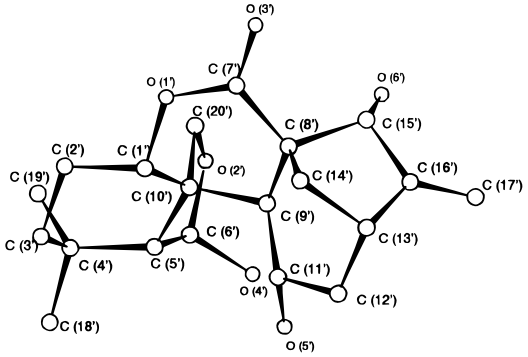


Figure 4. ORTEP drawing of the crystal structure of epinodosinol (**2**).

2; hydrogen-bearing carbon signals could be assigned by HETCOR, and the quaternary carbon signals of C-4, C-8, and C-10 were assigned by means of FLOCK, in which informative correlations were observed between H_3-18/C-4, H-1/C-10, H-13/C-8, and H-9/C-8.

The structure characterizations of the known compounds sodoponin (**6**), trichorabdal (**7**), and cirsiliol (**8**) were accomplished by direct comparison of ^{1}H, ^{13}C NMR, mp, and $[\alpha]_D$ with published data.[7,8]

X-ray crystallography of DCRA: colorless crystals, $C_{42}H_{60}O_{13}$, MW = 772.93, monoclinic, space group C_2, with a = 21.893(5) Å, b = 6.496(1) Å, c = 28.561(6) Å, β = 73.65(2)°, V = 3898.21 Å^3, Z = 4. Collection and processing were carried out on a R3m/E four-circle diffractometer, graphite-monochromated Cu Kα X-radiation in the range $0° < \theta < 57°$, W/2Q scan. The structure model was elucidated by a direct method (SHELXTL), and all of the 55 non-hydrogen atoms were positioned by difference Fourier synthesis. The positional parameters were refined by full-matrix least-squares methods to the final R = 0.0553 for 2735 refractions used in all calculations [$I > 3\sigma(I)$] out of 4280 unique refractions measured.

It was clearly demonstrated by X-ray crystallography that DCRA consists of the two substrates **1** and **2**. The major contact surface between **1** and **2** in the crystal

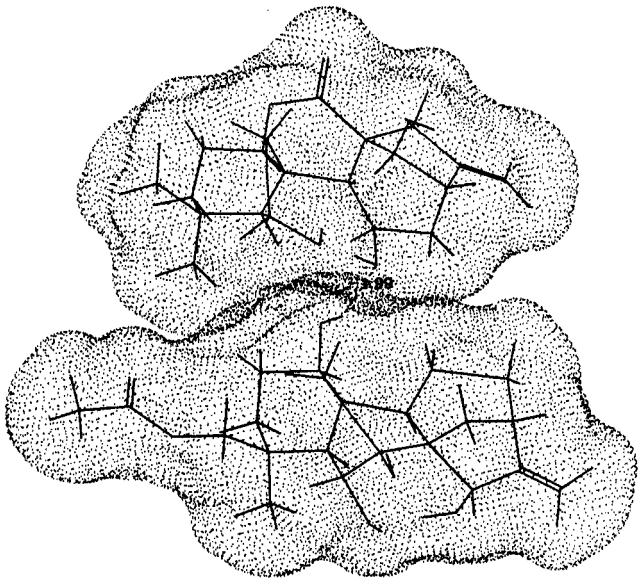


Figure 5. Possible interaction mode of subunits **1** and **2** in DCRA (**3**). The intermolecular hydrogen bond is shown in a broken line and the Connolly surfaces for each subunits are displayed in dots.

shows that **1** and **2** are bound together by an intermolecular hydrogen bond between the 11-OH of **2** and the 1-OH of **1** and the very good hydrophobic close approach of mutual matching surfaces (as shown in Figure 5); intramolecular hydrogen bonds also contribute to the stabilization of the complex.

It is interesting how DCRA can exist in a stable form in polar solvents such as $CHCl_3$, MeOH, and H_2O, since complexation competes with solvolysis in solution; i.e., the stabilizing "force" must be strong enough to compensate for the desolvation energy of both subunits and the loss of entropy on going from two free molecules to form a single molecule of the complex. It is also notable that **1** binds **2** with very high specificity because neither nonbound **1** or **2** were isolated nor were complexes found among other analogs of **1** and **2** (**4**−**7**) isolated simultaneously from the same source. Decomposition of the

Table 1. ^{1}H and ^{13}C NMR Assignments of Neoangustifolin (**1**) and Epinodosinol (**2**) in DCRA[a]

	^{1}H		^{13}C	
carbon	**1**	**2**	**1**	**2**
1	3.77 (dd, 5.0, 11.0)	4.79 (t, 6.5)	73.45	76.45
2	1.82 (m)	1.84 (m)	24.33	30.18
3	1.83 (m, α); 1.17 (m, β)	1.18 (m)	33.67	37.15
4			37.58	31.67
5	1.84 (br.s)	3.11 (br.s)	58.89	54.32
6	4.36 (d, 7.0)	5.66 (br.s)	73.75	102.21
7			97.33	175.22
8			52.65	53.34
9	2.63 (m)	3.53 (d, 9.5)	43.73	46.30
10			41.47	51.00
11	1.93 (m), 2.29 (m)	4.38 (ddd, 11.5, 9.5, 8.5)	19.13	63.03
12	1.58 (m), 2.23 (m)	1.87 (m), 2.90 (m)	33.08	45.56
13	2.60 (m)	2.71 (m)	37.10	37.11
14	2.04 (m)	1.67 (m)	26.89	34.45
15	5.13 (br s)	5.21 (br.s)	75.25	77.74
16			162.34	158.12
17	5.42 (br.s), 5.17 (br.s)	5.44 (br.s), 5.17 (br.s)	106.86	108.54
18	1.31 (s)	0.93 (s)	26.97	33.15
19	4.85, 4.46 (ABd, 11.0)	0.94 (s)	66.44	23.18
20	4.82, 4.31 (ABd, 11.0)	4.18, 4.13 (ABd, 11.0)	64.23	73.20
−OAc	1.91 (s)		170.81	
			20.71	

[a] Recorded in pyridine-d_5, chemical shift values were reported as δ values (ppm) from internal TMS at 500 MHz. Signal multiplicity and coupling constants (Hz) are shown in parentheses.

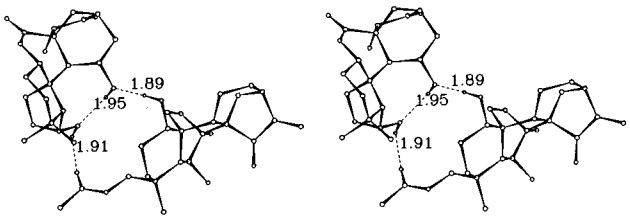


Figure 6. Stereoview of a possible binding mode of **1** and **2** derived from minimization calculation of the original crystal structure. The H-bond network on the contact surface of **1** and **2** is displayed by the distances from the polar hydrogen atoms to the donor aceptor atoms.

complex also did not occur on melting (215 °C), the TLC chromatographic pattern remained the same, and the FAB-MS (positive mode) was identical to the isolated starting material.

The inseparable nature of DCRA suggests that compounds **1** and **2** bind each other with high affinity in a thermodynamic equilibrium. The major interaction contact surface between **1** and **2** in crystal form, as shown in Figure 1, shows no conceivable barriers for their separation, in which both monomers are in their lowest energy conformation with maximum intramolecular H-bonds. However, the single intermolecular H-bond between **1** and **2** in the crystal structure could barely explain their high-affinity binding, whereas other hydrophobic interactions are not expected to be a major contributor to the binding in organic solvents. In order to investigate the possible complex forms of DCRA, the Monte Carlo searching method was applied using Macromodel software (version 4.5) and MM2 forcefield. One of the lower energy complex forms derived is displayed in Figure 6. In this complex structure, while keeping the original intermolecular H-bond, an additional H-bond is established between O(6) of **1** and O(5) of **2**. This H-bond could be a substantial contributor to the binding of **1** and **2**. The H-bond pattern on the contact surface of this model also explains the specificity between **1** and **2**, whereas the crystal structure could not. Compound **6**, which does not have the acetoxyl group of **1** that is involved in an intermolecular bond with **2**, is not capable of complexation with **2**. From the primary modeling analysis performed, we speculate that the binding modes between **1** and **2** in solvents could be more complicated than they initially appear and could be very different from what was shown by X-ray crystallography. Further modeling analysis will be carried out to explore the possible recognition mechanisms between **1** and **2**.

Host−guest complexation of artificial receptors drawn by binding forces due to hydrophobic proximity and hydrogen bonding with directionality have been extensively studied.[9,10] However, the observation of such phenomena involving naturally occurring organic molecules are seldom reported, as exemplified by the complexation involving alkaloids (−)-ormosanine and (−)-podopetaline,[11] 6-hydroxydopamine hydrochloride, and a *p*-quinone,[12] as well as the complexation of the C-4 diastereoisomers of manicoline B.[13] These novel complexes appear interesting and worthy of detailed study because of their high degree of selectivity, as shown in the current case, and the high binding tightness that is evident from the observation that they have "survived" other weak-bound complexes under a broad spectrum of conditions such as different solvents, chromatography, high temperature, and repeated recrystallizations. Knowledge of the complexation of natural products may present a clue to a deeper understanding of more complicated biomolecules, offering the opportunity that may permit the design of new ligands capable of binding other small biological molecules. Attempted separation of DCRA, which may make possible a determination of the binding constant and characterization of further natural products capable of binding to other substrates, is in progress.

Acknowledgment. This work was supported, in part, by grants from the Key Laboratory of Phytochemistry, Kunming Institute of Botany, Chinese Academy of Sciences, People's Republic of China, to the Chinese research group. The authors thank Z.-W. Lin of Kunming Institute of Botany, Chinese Academy of Sciences, and B. Wu and Y. Lu of the Institute of Materia Medica, Chinese Academy of Medical Science, for their assistance in isolation and X-ray analysis, respectively; Prof. H.-W. Li for the identification of the plant materials; Mr. R. Dvorak for the MS data; and the Research Resources Center of the University of Illinois at Chicago for the provision of NMR spectroscopic facilities.

References and Notes

(1) Fujita, E.; Node, M. *Prog. Chem. Org. Nat. Prod.* **1983**, *46*, 78−157.
(2) Sun, H.-D.; Lin, Z.-W.; Yoshinori, M.; Teruyoshi, M. *Zhongguo Yaoli Xuebao* **1984**, *19*, 633−636.
(3) Takeda, T.; Fujita, T.; Sun, H.-D. *Chem. Pharm. Bull.* **1990**, *38*, 1877−1880.
(4) Dried and powdered leaves (4.0 kg) were extracted with 95% EtOH. After evaporation of the solvent, the residue was partitioned with petroleum ether, EtOAc, and *n*-BuOH, successively. The EtOAc-soluble part (70 g) was subjected to column chromatography (silica gel, Sephadex LH-20) and finally preparative HPLC (Partisil-ODS, solvent system MeOH:H_2O = 8.5:1.5, flow rate 1 mL/min) and afforded compounds **3** (64 mg, yield 0.0016%), **4** (240 mg, yield 0.006%), **5** (536 mg, yield 0.0134%), **6** (425 mg, yield 0.0106%), **7** (46 mg, yield 0.0012%), and **8** (43 mg, yield 0.0011%).
(5) The atomic coordinates, equivalent isotropic displacement parameters, lengths, bond angles, anisotropic displacement parameters, H-atom coordinates, and isotropic displacement parameters for this work can be obtained, upon request, from Dr. Olga Kennard, University Chemical Laboratory, 12 Union Road, Cambridge CB2 1EW, UK. Any request should be accompanied by the full literature citation for this paper.
(6) Fujita, E.; Fujita, T.; Taoka, M.; Katayama, H.; Shibuya, K. *Tetrahedron Lett.* **1970**, 421−424.
(7) Fujita, E.; Fujita, T.; Taoka, M.; Katayama, H.; Shibuya, K. *Chem. Pharm. Bull.* **1973**, *21*, 1357−1363.
(8) Sun, H.-D.; Lin, Z.-W. *Acta Bot. Yunnan.* **1988**, *10*, 215−218.
(9) Cram, D. J. *Science* **1988**, *240*, 760−767.
(10) Lehn, J.-M. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 90−112.
(11) Misra, R.; Wong-Ng, W.; Cheng, P.-T.; Mclean, S.; Nyburg, S. *J. Chem. Soc., Chem. Commun.* **1980**, 659−660.
(12) Andersen, A. M.; Mostad, A.; Romming, C. *Acta Chem. Scand.* **1975**, *B29*, 45−50.
(13) Polonsky, J.; Prange, T.; Pascard, C. *Tetrahedron Lett.* **1984**, *25*, 2359−2362.

NP970001D