The Journal of
Organic
Chemistry

VOLUME 59, NUMBER 7 APRIL 8, 1994

Communications

Antineoplastic Agents. 277. Isolation and Structure of Phakellistatin 3 and Isophakellistatin 3 from a Republic of Comoros Marine Sponge[1a]

George R. Pettit,* Rui Tan, Delbert L. Herald, Ronald L. Cerny[1b] and Michael D. Williams

Cancer Research Institute and Department of Chemistry, Arizona State University, Tempe, Arizona 85287-1604

Received October 7, 1993 (Revised Manuscript Received January 14, 1994®)

Summary: The Western Indian ocean sponge *Phakellia carteri* was found to contain two isomeric (at the photo-Trp indole ring juncture) *cyclo*-heptapeptides named phakellistatin 3 (**3**) and isophakellistatin 3 (**4**).

Exploratory investigations of marine Porifera heterocyclic and sterol constituents over 40 years ago by Bergmann[2a] resulted in the discovery of the arabinose-type nucleosides that led to the now well-established anticancer and antiviral drugs ARA-C and ARA-A.[2] Illustrative of new marine sponge heterocyclic constituents are the recent isolations of pyridine,[3ab] isoquinoline,[3c] bis-indole,[3d] carboline,[3e] imidazole,[3f] pyrimidine,[3g] piperazines,[3h] pyrroloquinoline,[3i] oxazole,[3j,k] and thiazole[3k] ring system representatives (*cf.* also heterocyclic compounds from microorganisms associated with sponges).[4]

Our earlier evaluations of marine sponges in the genus *Phakellia* for cell growth inhibitory substances led to isolation of phakellistatin 1 (**1**)[5] from a Western Pacific (Micronesia) *Phakellia* sp. and phakellistatin 2 (**2**)[6] from the Eastern Indian Ocean (Republic of the Comoros) *Phakellia carteri*. We now report that further bioassay guided (murine P388 lymphocytic leukemia cell line) separation of *P. carteri* cytotoxic constituents has resulted in the discovery of phakellistatin 3 (**3**), a cyclic peptide with significant P388 (ED_{50} 0.33 μg/mL) activity. Phakellistatin 3 represents a new type of *cyclo*-heptapeptide containing an amino acid unit apparently derived from a photooxidation product of tryptophan.[7] The *cis*-ring juncture isomer of this modified tryptophan unit was found in an otherwise identical new *cyclo*-heptapeptide designated isophakellistatin 3 (**4**).

The dichloromethane soluble fraction obtained from a 250-kg (wet wt) 1987 collection of *P. carteri* (order Axinellida, class Demospongiae) which was used to obtain phakellistatin 2[6] by a series of solvent partitioning, gel permeation, and partition chromatographic (Sephadex LH-20) procedures was further separated (P388 bioassay guided). Continued column chromatographic separations on Sephadex LH-20 and purification by reversed-phase HPLC afforded phakellistatin 3 (**3**, 4.6 mg, 2.0×10^{-6}%)

• Abstract published in *Advance ACS Abstracts*, February 15, 1994.

(1) (a) For part 276 see: Stewart, J. P.; McGown, A. T.; Prendiville, J.; Pettit, G. R.; Fox, B. W.; Arrand, J. R. Manuscript in preparation. (b) Midwest Center for Mass Spectrometry, Chemistry Dept., University of Nebraska, Lincoln, NE 68588.

(2) (a) Consult for leading references: Pettit, G. R., *Biosynthetic Products for Cancer Chemotherapy*; Plenum Publishing Co.: New York, 1978; Vol. 1. (b) Djerassi, C.; Silva, C. J. *Acc. Chem. Res.* **1991**, *24*, 371.

(3) (a) Kobayashi, J.; Zeng, C.; Ishibashi, M.; Shigemori, H.; Sasaki, T.; Mikami, Y. *J. Chem. Soc., Perkin Trans.* **1992**, 1291. (b) Talpir, R.; Rudi, A.; Ilan, M.; Kashman, Y. *Tetrahedron Lett.* **1992**, *33*, 3033. (c) Davidson, B. S. *Tetrahedron Lett.* **1992**, *33*, 3721. (d) Wright, A. E.; Pomponi, S. A.; Cross, S. S.; McCarthy, P. *J. Org. Chem.* **1992**, *57*, 4772. (e) Adesanya, S. A.; Chbani, M.; Pais, M.; Debitus, C. *J. Nat. Prod.* **1992**, *55*, 525. (f) He, H.; Faulkner, D. J.; Lee, A. Y.; Clardy, J. *J. Org. Chem.* **1992**, *57*, 2176. (g) Berlinck, R. G. S.; Braekman, J. C.; Daloze, D.; Bruno, I.; Riccio, R.; Rogeau, D.; Amade, P. *J. Nat. Prod.* **1992**, *55*, 528. (h) Dillman, R. L.; Cardellina, J. H., II. *J. Nat. Prod.* **1991**, *54*, 1159. (i) Stierle, D. B.; Faulkner, D. J. *J. Nat. Prod.* **1991**, *54*, 1131. (j) Walker, M. A.; Heathcock, C. H. *J. Org. Chem.* **1992**, *57*, 5566. (k) Gunawardana, G. P.; Koehn, F. E.; Lee, A. Y.; Clardy, J.; He, H.; Faulkner, D. J. *J. Org. Chem.* **1992**, *57*, 1523.

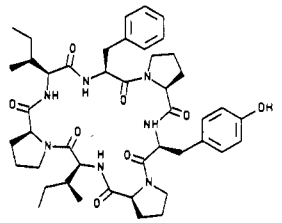
(4) (a) Shigemori, H.; Bae, M.-A.; Yazawa, K.; Sasaki, T.; Kobayashi, J. *J. Org. Chem.* **1992**, *57*, 4317. (b) Stierle, D. B.; Faulkner, D. J. *J. Nat. Prod.* **1991**, 1134.

(5) Pettit, G. R.; Cichacz, Z.; Barkoczy, J.; Dorsaz, A.-C.; Herald, D. L.; Williams, M. D.; Doubek, D. L.; Schmidt, J. M.; Tackett, L. P.; Brune, D. C.; Cerny, R. L.; Hooper, J. N. A.; Bakus, G. J. *J. Nat. Prod.* **1993**, *56*, 260.

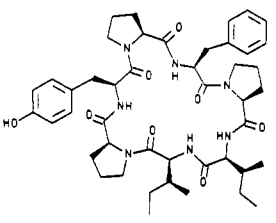
(6) Pettit, G. R.; Tan, R.; Williams, M. D.; Tackett, L.; Schmidt, J. M.; Cerny, R. L.; Hooper, J. N. A. *BioMed. Chem. Lett.*, submitted.

(7) Nakagawa, M.; Kato, S.; Kataoka, S.; Kodato, S.; Watanabe, H.; Okajima, H.; Hino, T. *Chem. Pharm. Bull.* **1981**, *29*, 1013. The *cis/trans* nomenclature here refers to the relative orientations of the C-52 hydroxyl group and C-17 carbonyl substituent.

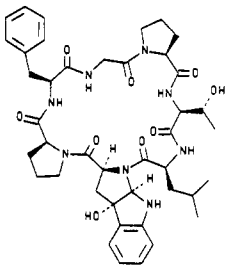
0022-3263/94/1959-1593$04.50/0



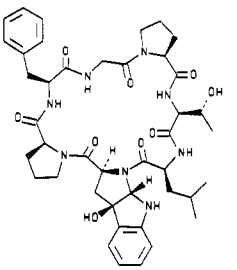
1, phakellistatin 1



2, phakellistatin 2



3, phakellistatin 3



4, isophakellistatin 3

as an amorphous powder (P388 ED_{50} 0.33 μg/mL) [mp 178–180 °C; TLC R_f 0.17 (hexane–ethyl acetate–methanol (4:2:1)); $[\alpha]^{24}_D$ −147° (*c* 0.22, CH_3OH); UV λ_{max} nm (log ε) in CH_3OH, 214 (4.17), 234 (4.04), 291 (3.52); IR ν_{max} cm^{-1} (film) 3335, 2926, 1663, 1522, 1456]; and isophakellistatin 3 (4, 30 mg, 1.3 × 10^{-5}% yield): crystals from acetone melting at 218–220 °C; TLC R_f 0.20 (hexane–ethyl acetate–methanol (4:2:1)); $[\alpha]^{23}_D$ −138° (*c* 0.21, CH_3OH); UV λ_{max} nm (log ε) in CH_3OH, 211 (4.41), 238 (3.91), 295 (3.37); IR ν_{max} cm^{-1} (KBr) 3453, 3370, 2955, 2872, 1670, 1634, 1520; FABHRMS *m/z* calcd. for $C_{42}H_{55}N_8O_9$ 815.4092, found 815.4090 $[M + H]^+$.

Comprehensive high-field (400- and 500-MHz) 2D-NMR analyses of *cyclo*-heptapeptides 3 and 4 employing APT, $^1H,^1H$-COSY, HMQC, and HMBC[8] techniques revealed spin systems corresponding to Gly, Leu, Pro (two units), Phe, Thr, and a Trp-like amino acid unit (3a-hydroxypyrroloindole). Amino acid analyses added further support for the common amino acid components. Comparison of the ^1H- and ^{13}C-NMR spectra revealed significant differences in only two ^{13}C-signals: a carbon at δ 171.49 (Trp CO) and a methine at δ 62.80 (Trp αCH) for peptide 3 *vs* δ 173.29 and δ 61.63 for the same carbons of peptide 4. Again in the ^1H-NMR spectra only two signals differed, namely a peptide 3 methylene proton at δ 2.00 (Trp βCH_2) and a methine proton at δ 5.54 (Trp δCH) *vs* δ 2.24 and δ 5.75 for peptide 4. By elimination, all of these signal shift differences were found to arise from the tryptophan-derived unit. Furthermore, sequence analyses by MS/MS techniques indicated peptides 3 and 4 differed only by stereochemistry.

Fortunately, a specimen of isophakellistatin 3 (4) suitable for X-ray crystal structure determination crystallized from acetone: crystal data; $C_{42}H_{54}N_8O_9 \cdot C_3H_6O \cdot H_2O$; crystal dimensions 0.26 × 0.28 × 0.28 mm; Cu Kα radiation (λ = 1.541 84 Å) at 27 ± 1 °C; monoclinic cell parameters and calculated volume of a = 9.994(3) Å, b = 18.094(3) Å, c = 12.961(3) Å, β = 98.60(2)°, and V = 2317.40 Å^3; space group $P2_1$; 8531 reflections were collected, of which 7789 were unique and not systematically absent (includes Friedel reflections).

The structure was solved by direct methods using SIR88.[9] Of the 64 non-hydrogen atoms in the asymmetric unit, 55 non-hydrogen atoms of the main peptide molecule were located in the first run of SIR88 using the default settings. Recycling of the 55-atom fragment through SHELXS86[10] using the TEXP 250 option provided the remainder of the structure, excluding one aromatic carbon of the phenylalanine ring. Subsequent difference Fourier runs using CRYSTALS[11] revealed the presence of the remaining aromatic carbon atom as well as the two solvate molecules, i.e., one molecule of water and one molecule of acetone.

The crystal structure of isophakellistatin 3 (4) is shown in Figure 1. Each molecule of isophakellistatin 3 was found to be associated with a single molecule of acetone and a molecule of water. Based upon the A- - -B bond length observed in an A–H- - -B hydrogen bond, isophakellistatin 3 appears to be involved in both intermolecular and intramolecular H-bonding. Generally, peptide N–H- - -O bonds have an A- - -B length in the range of ~2.79 ± 0.12 Å and O–H- - -O bonds have an A- - -B length in the range of 2.76 ± 0.09 Å. Thus, intramolecular H-bonding seems probable between the N1- - -O43 (phenylalanine amide–leucine carbonyl, 2.827 Å) hetero atoms and possibly between the O54- - -O55 (hydroxyl–carbonyl, 2.820 Å) atoms. From the bond distances, intermolecular H-bonding between O36- - -O60 (threonine hydroxyl–water, 2.789 Å), O36- - -O29 (threonine hydroxyl–phenylalanine carbonyl, 2.670 Å), and O59- - -O60 (proline carbonyl–water, 2.795 Å) was also predicted. A computer-generated perspective drawing depicting the (*S*)-absolute configuration of the constituent amino acids (determined *via* chiral GLC analysis[5] of the peptide hydrolysis products), as well as the conformation and proposed H-bonding of the cyclic peptide (4), appears in Figure 1.

With the crystal structure of isophakellistatin 3 in hand, a series of NMR NOESY experiments relating peptides

(8) Pettit, G. R.; Kamano, Y.; Inoue, M.; Dufresne, C.; Boyd, M. R.; Herald, C. L.; Schmidt, J. M.; Doubek, D. L.; Christie, N. D. *J. Org. Chem.* **1992**, *57*, 429.

(9) Burla, M. C.; Camalli, M.; Cascarano, G.; Giacovazzo, C.; Polidori, G.; Spagna, R.; Viterbo, D. *J. Appl. Crystallogr.* **1989**, *22*, 389.

(10) Sheldrick, G. M., *SHELXS-86. A Program for Solution of Crystal Structure from Diffraction Data*; Institute für Anorganische Chemie der Universität Göttingen: Tammanstr. 4, D-3400 Göttingen, Germany, 1986.

(11) Watkin, D. J.; Carruthers, J. R.; Betteridge, P. W. *Crystals User Guide*; Chemical Crystallography Laboratory, University of Oxford: Oxford, England, 1985. The hydrogen atom coordinates for both the parent molecule and acetone solvate molecule were calculated at optimum positions and added in the final stages of the least-squares refinement and structure-factor calculation process but were not refined. Full-matrix least-squares refinement was performed with CRYSTALS, using the Robust-Resistant (Tukey and Prince) weighting method (scheme 15 in CRYSTALS). Anomalous dispersion effects were included in F_c. A total of 6891 reflections having intensities greater than 3.0 times their standard deviation were used in the refinement process. Near the conclusion of the refinement process, an additional least-squares absorption correction (see ref 12) was made to the data using the CRYSTALS program DIFABS. The final cycle of least-squares refinement included 578 variable parameters (anisotropic refinement on all non-hydrogen atoms of the peptide and solvate molecules) and converged with unweighted and weighted agreement factors of $R = \Sigma||F_o| - |F_c||/\Sigma|F_o| = 0.076$ and $R_w = [\Sigma w(F_o - F_c)^2/\Sigma w(F_o)^2] = 0.099$.

(12) (a) Walker, N.; Stuart, D. *Acta Crystallogr.* **1983**, *A39*, 158. (b) Ugozzli, F. *Comput. Chem.* **1987**, *11*, 109. (c) Katayama, C. *Acta Crystallogr.* **1986**, *A42*, 19.

(13) Preparation of Figure 1 was accomplished with SHELXTL-PLUS; G. Sheldrick, G. Siemens Analytical X-Ray Instruments, Inc., Madison, WI 53719.

(14) The author has deposited atomic coordinates for 4 with the Cambridge Crystallographic Data Centre. The coordinates can be obtained, on request, from the Director, Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, CB2 1EZ, UK.

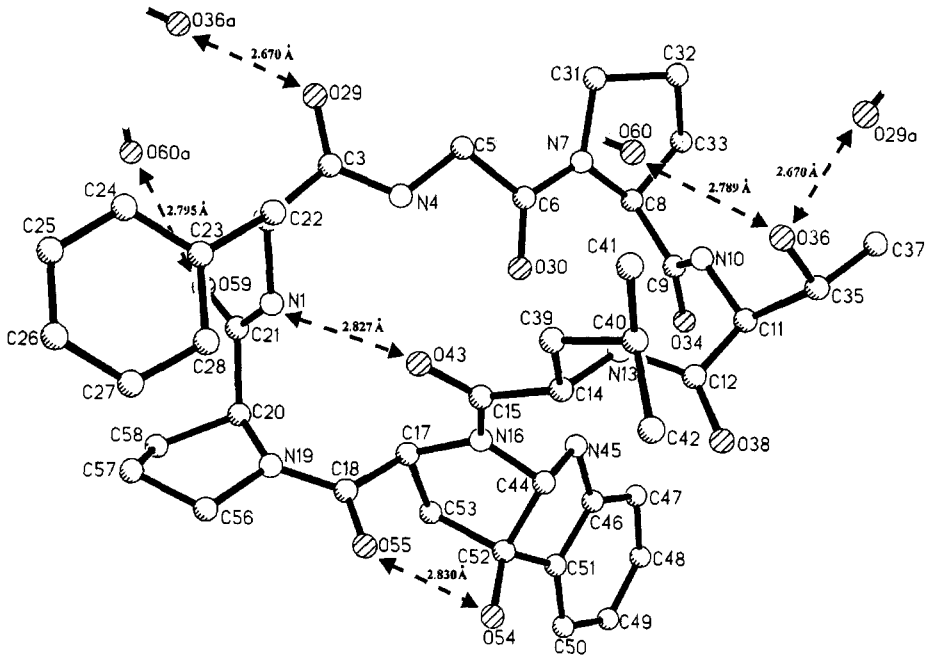


Figure 1. Isophakellistatin 3.

3 and 4 was conducted. These allowed assignment of structure 3 to phakellistatin 3. The result was consistent with the original study[7] of tryptophan photooxidation where two photoproducts isomeric at the ring juncture were identified as *cis*- and *trans*-3a-hydroxyl-1,2,3,3a,8,-8a-hexahydropyrrolo[2,3-b]indole 2-carboxylic acid (herein named *cis*- and *trans*-photo-Trp). To our knowledge *cyclic*-heptapeptides 3 and 4 represent the first examples of photo-Trp serving as a natural peptide unit. The inhibitory effect of the *trans*-ring juncture (peptide 3) on growth of the P388 cell line inhibition *vs* lack of such activity with the *cis*-isomer (peptide 4) is also noteworthy. Extended antineoplastic evaluations of phakellistatin 3 are underway.

Acknowledgment. The critical financial support for this investigatiion was provided by Outstanding Investigator Grant CA 44344-01A1-05 awarded by the Division of Cancer Treatment, National Cancer Institute, DHHS, the Arizona Disease Control Research Commission, Eleanor W. Libby, Virginia Piper, the Eagles Art Ehrmann Cancer Fund, and the Robert B. Dalton Endowment Fund. For other necessary assistance we thank the Republic of the Comoros (Damir Ben Ali, Mohammed Chaher, Karl Danga, W. R. Carlson), Drs. Daniel C. Brune, Cherry L. Herald, Fiona Hogan-Pierson, Ronald A. Nieman and Jean M. Schmidt, Mr. Larry Tackett, Ms. Denise Nielsen-Tackett, Mr. Lee Williams, Mrs. Kim M. Weiss, the U. S. National Science Foundation (Grants CHE-8409644 and BBS-88-04992), and the NSF Regional Instrumentation Facility in Nebraska (Grant CHE-8620177).

Supplementary Material Available: ^{1}H- and ^{13}C-NMR data for phakellistatin 3 and isophakellistatin 3 (6 pages). This material is contained in libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.