ORGANIC LETTERS

2000
Vol. 2, No. 11
1605−1607

Cyclodidemniserinol Trisulfate, a Sulfated Serinolipid from the Palauan Ascidian *Didemnum guttatum* That Inhibits HIV-1 Integrase

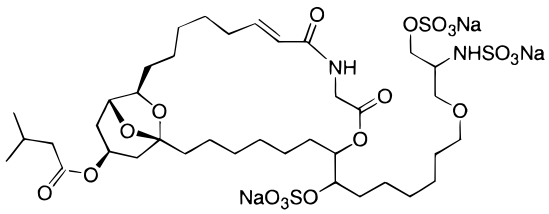
Scott S. Mitchell,[†] Denise Rhodes,[‡] Frederic D. Bushman,[‡] and D. John Faulkner*,[†]

Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093-0212 and Infectious Disease Laboratory, The Salk Institute, 10010 North Torrey Pines Road, La Jolla, California 92037

jfaulkner@ucsd.edu

Received March 27, 2000

ABSTRACT



Bioassay-guided fractionation of extracts of the Palauan ascidian *Didemnum guttatum* led to the isolation of cyclodidemniserinol trisulfate (1) as an inhibitor of HIV-1 integrase, which is an attractive target for anti-retroviral chemotherapy. The structure of cyclodidemniserinol trisulfate (1), the stereochemistry of which was only partially determined, was elucidated by interpretation of NMR and mass spectral data.

HIV encodes three enzymes: reverse transcriptase, protease, and integrase. While inhibitors of reverse transcriptase and protease are used, usually in combination, in the treatment of HIV-infected people, no clinically useful inhibitors of integrase have been described.[1] As part of a larger program to discover and design new inhibitors of HIV-1 integrase,[2] we screened several hundred extracts of marine invertebrates from Palau. Although a number of the active extracts contained sterol sulfates, of which haplosamates A and B were the most interesting examples,[3] we have now concentrated our efforts on finding nonsteroidal integrase inhibitors.

In this paper we report the structural elucidation of cyclodidemniserinol trisulfate (**1**) from the Palauan ascidian *Didemnum guttatum*, Monniot and Monniot 1996.

Although cyclodidemniserinol trisulfate (**1**) is most closely related to didemniserinolipid A (**2**) from an Indonesian *Didemnum* sp.,[4] there are some significant differences between the two structures, most notably the presence of an additional ring containing a glycine unit and the presence of sulfate groups.

A specimen of *Didemnum guttatum* was collected by hand using SCUBA (−5 m) at Ngerchaol Island, Palau, and was maintained frozen until it was extracted. The crude methanolic extract showed selective inhibition of HIV-1 integrase versus the MCV topoisomerase counterscreen.[5] Bioassay-

[†] Scripps Institution of Oceanography.
[‡] The Salk Institute.

(1) Coffin, J. M.; Hughes, S. H.; Varmus, H. E. *Retroviruses*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor. 1997; p 834.

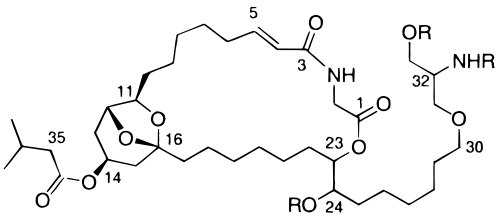
(2) See, for example: (a) Reddy, M. V. R.; Rao, M. R.; Rhodes, D.; Hansen, M.; Rubins, K.; Bushman, F.; Venkateswarlu, Y.; Faulkner, D. J. *J. Med. Chem.* **1999**, *42*, 1901−1907. (b) Carlson, H. A.; Masukawa, K. M.; Rubins, K.; Bushman, F. D.; Jorgensen, W. L.; Lins, R. D.; Briggs, J. M.; McCammon, J. A. *J. Med. Chem.*, in press.

(3) Qureshi, A.; Faulkner, D. J. *Tetrahedron* **1999**, *55*, 8323−8330.

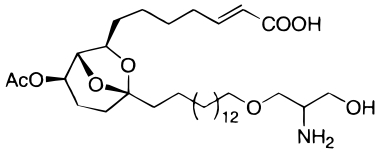
(4) González, N.; Rodriguez, J.; Jiménez, C. *J. Org. Chem.* **1999**, *64*, 5705−5707.

(5) Counterscreens were carried out by assaying inhibition of the topoisomerase enzyme of Molluscum contagiosum virus (MCV). Assays monitored the breaking and rejoining of a DNA strand characteristic of type 1B topoisomerases and were carried out in a microtiter plate format.

10.1021/ol005866o CCC: $19.00
Published on Web 05/05/2000



1 R = SO_3Na
3 R = H



2

guided fractionation revealed that the active material, cyclodidemniserinol trisulfate (**1**, 5×10^{-3} % dry wt), was a water-soluble compound that could be purified by repeated reversed phase chromatography.

Cyclodidemniserinol trisulfate (**1**), $[\alpha]_D$ −26.6°, was isolated as a colorless oil. The molecular formula, $C_{38}H_{63}N_2O_{19}S_3Na_3$, was determined from the high-resolution mass measurement of the $[M - Na]^-$ ion at 993.2961 (Δ 2.7 ppm) and the low resolution mass of the $[M + Na]^+$ ion at *m*/*z* 1039. The IR spectrum included bands assigned to the ester (1725 cm^{-1}), α,*β*-unsaturated amide (3320, 1665, 1625 cm^{-1}), and sulfate (1220 cm^{-1}) groups. The UV spectrum contained an absorption at 210 nm (ϵ 4500) that was assigned to the α,*β*-unsaturated amide group. An initial analysis of the ^{13}C NMR spectrum (Table 1) provided evidence for three carbonyl groups (*δ* 171.7, 169.5, 165.7), an olefin (*δ* 143.2, 123.8), a ketal (*δ* 107.8), and eight carbons bearing oxygen, of which five were methines (*δ* 79.4, 77.2, 76.3, 74.9, 66.2) and three were methylenes (*δ* 70.5, 69.8, 65.9). The data required that cyclodidemniserinol (**1**) be tricyclic, with two rings being involved in a bicyclic ketal ring system. A search of the marine natural products literature indicated a possible similarity to didemniserinolipid A (**2**).[4]

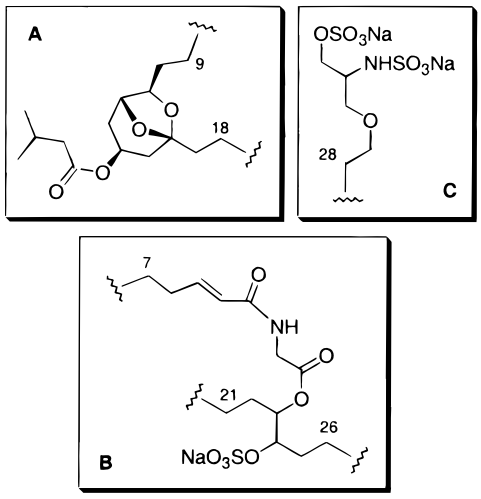


Table 1. ^{13}C (DMSO-d_6, 100 MHz) and ^1H NMR (DMSO-d_6, 500 MHz) Data for Cyclodidemniserinol Trisulfate (**1**)

C no.	δ_C	δ_H	mult, no. of H, *J*(Hz)	HMBC
1	171.7			
2	42.2	3.70	d, 2 H, 6	C1, C3
NH-2		8.42	t, 1 H, 6	C3, C2
3	165.7			
4	123.8	5.95	d, 1 H, 15.5	C3, C6
5	143.2	6.65	dt,1 H, 15.5, 7	C3, C6, C7
6	31.0	2.07	m, 2 H	C8
7	27.7	1.35	m, 2 H	
8	28.3	1.40	m, 2 H	
9	24.8	1.08	m, 2 H	
10	34.5	1.25	m, 2 H	
11	79.4	3.83	m, 1 H	C9, C13, C16
12	77.2	4.20	br m, 1 H	C14, C16
13	34.0	2.03	m, 1 H	C14, C15
		1.58	m, 1 H	C12
14	66.2	5.05	tt, 1 H, 10, 7	C31
15	41.0	1.98	m, 1 H	C13, C14, C16
		1.34	m, 1 H	C14, C16
16	107.8			
17	36.2	1.54	m, 2 H	C16, C19
18	22.3	1.37	m, 2 H	
19	*a*			
20	*a*			
21	24.9	1.34	m, 2 H	
22	30.2	1.42	m, 2 H	
23	74.9	4.81	dt, 1 H, 10, 6	C1, C21, C22 C24
24	76.3	4.11	dt, 1 H, 10, 5	C22, C23, C26
25	28.9	1.40	m, 2 H	
26	24.6	1.22	m, 2 H	
27	*a*			
28	25.8	1.20	m, 2 H	
29	29.3	1.40	m, 2 H	
30	70.5	3.30	m, 2 H	C28, C29, C31
31	69.8	3.45	dd, 1 H, 10, 5	C30, C32, C33
		3.30	dd, 1 H, 10, 5	C30, C32, C33
32	52.3	3.30	1 H, m	C31, C33
33	65.9	3.82	dd, 1 H, 10, 5	C31, C32
		3.72	dd, 1 H, 10, 5	C31, C32
34	171.7			
35	42.8	2.14	d, 2 H, 7	C34, C36, C37 C38
36	25.4	1.90	hept, 1 H, 6.5	
37, 38	22.1	0.89	d, 6 H, 6.5	C35, C36, C37 C38
NH-32		7.17	br s, 1 H	

[a] Unassigned signals at δ_C 29.9 (δ_H = 1.09), 29.7 (1.14), and 29.3 (1.40).

Analysis of the NMR data (Table 1) led to the structural elucidation of substructures **A**, **B**, and **C**. In substructure **A**, the isovalerate ester residue (C-34 to C-38) was easily identified from the ^1H, ^{13}C, COSY, HSQC, and HMBC data summarized in Table 1. There was a three-bond correlation from the C-34 carbonyl to the H-14 methine signal at *δ* 5.05 (tt, 1 H, *J* = 10, 7 Hz), which is attached to a methine carbon with an unusual upfield chemical shift (δ_C 66.2) that signified further oxygen atoms on the *β*-carbons, which was not

compatible with the bicyclic ketal ring system of **2**. Analysis of the COSY data revealed that H-14 was coupled to methylene signals at δ 2.03 and 1.34 (H_2-15; δ_C 41.0), which showed no further coupling. The H-14 signal was also coupled to methylene signals at 1.98 and 1.58 (H_2-13; δ_C 34.0), that were coupled to a signal at 4.20 (br s, 1 H, H-12), which was not coupled further. HMBC correlations from the H_2-15 signals placed the ketal at C-16, which showed additional correlations to the H-12 signal and to the H-11 signal at δ 3.83. The lack of coupling between H-11 and H-12 was explained by molecular modeling that revealed that the H-11/H-12 dihedral angle was approximately 90°. A ROESY experiment showed strong NOEs from both H-12 and H-14 to H-11, confirming the structure and allowing assignment of the stereochemistry shown for the bicyclic ketal substructure **A**. By using the HMBC data, we were able to extend substructure **A** to account for all signals in the C-9 to C-18 region.

Substructure **B** contains an α,β-unsaturated amide linked through the amide bond to a glycine unit which is in turn joined through an ester bond to an aliphatic chain containing a vicinal diol. The methylene protons of the glycine unit show HMBC correlations to the carbonyl signals at δ_C 169.5 (C-1) and 165.7 (C-3), which in turn is correlated to the olefinic proton signals at 5.95 (d, 1 H, J = 15.5 Hz, H-4) and 6.65 (dt, 1 H, J = 15.5, 7 Hz, H-5). By judicious use of the COSY and HMBC data, the methylene chain could be extended to C-8, but no direct connection with substructure **A** was observed. The HMBC correlation from H-23 signal at δ 4.81 (dt, 1 H, J = 10, 6 Hz) to C-1 and the COSY correlation to the H-24 signal at 4.11 (dt, 1 H, J = 10, 5 Hz) revealed that the glycine unit was joined through an ester bond to a vicinal diol, of which the second oxygen must be sulfated to comply with the molecular formula. The HMBC data allowed the alkyl chain to be extended from C-21 to C-26 but there was no overlap of carbon signals with those of substructure **A**.

Substructure **C** comprises an alkyl chain with a terminal serinol ether disulfate. The serinol unit was clearly defined by the NMR data, particularly the HSQC-TOCSY data and the chemical shift of C-32 at δ_C 52.3. The ether linkage was established by the HMBC correlations from the H-31 signals to a methylene carbon signal at δ_C 70.5. The corresponding methylene proton signal at δ_H 3.30 (m, 2 H, H-30) showed HMBC correlations to C-28, C-29, and C-31, thereby completing the assignments for substructure **C**, which again showed no carbon signals in common with either substructures **A** or **B**. Three carbon signals at δ_C 29.9, 29.7, and 29.3 remained to be assigned.

The linkage between substructures **A** and **B** was observed in a 70 ms TOCSY experiment, which showed a weak correlation between the allylic methylene protons (H_2-6) and H-11, but the correlation did not define the length of the carbon chain joining the substructures. The length of the chain between C-5 and C-11 was determined by oxidative degradation of **1** using acidified potassium permanganate, followed by methylation of the resulting acids using diazomethane to obtain dimethyl pimelate (MeOOC−$(CH_2)_5$−COOMe), which was detected by GC-MS using dimethyl adipate, dimethyl pimelate, and dimethyl suberate as standards. Analysis of the mass spectrum of **1** was complicated by the presence of the sulfate groups, which were removed by acid-catalyzed hydrolysis[6] to obtain cyclodidemniserinol (**3**), which was not present in the crude extracts. The mass spectra of **3**, including MS/MS spectra, contained key peaks at *m*/*z* 376, 334 and 203 (Figure 1) that confirmed length of

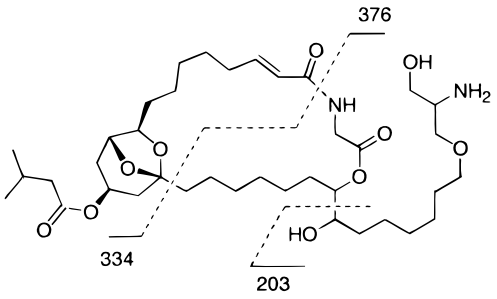


Figure 1. Key fragmentation peaks in the mass spectra of cyclodidemniserinol (**3**).

the A−B linkage and supported the proposed linkages from C-18 to C-21 and from C-26 to C-28.

The inhibition of purified integrase by cyclodidemniserinol trisulfate (**1**) was assayed. Reactions monitoring the ability of HIV integrase to form covalent bonds between a DNA that models the end of the unintegrated HIV DNA and a target DNA[1,7] were carried out in microtiter plates as described earlier.[2b,8] Cyclodidemniserinol trisulfate (**1**) inhibited purified integrase with an IC_{50} of 60 μg/mL. In contrast with the results obtained for the crude extracts, cyclodidemniserinol trisulfate (**1**) inhibited MCV topoisomerase with an IC_{50} of 72 μg/mL. Purified cyclodidemniserinol trisulfate (**1**) therefore exhibited almost no selectivity for integrase inhibition. Further research will concentrate on other water-soluble components of the crude extract.

Acknowledgment. We thank the Government of the Republic of Palau for research permits and Patrick L. Colin (Coral Reef Research Foundation, Palau) for identifying the ascidian by comparing our specimen with a voucher previously identified by Dr. F. Monniot (Paris). This research was funded by grants from the National Institutes of Health (CA 49084 to D.J.F. and GM 56553 to F.D.B.).

Supporting Information Available: Experimental procedures and ^{1}H NMR, ^{13}C NMR, COSY, HMBC, and HSQC-TOCSY spectra for cyclodidemniserinol trisulfate (**1**). This material is available free-of-charge via the Internet at http://pubs.acs.org.

OL005866O

(6) Albertini, E.; Barco, A., Benetti, S.; De Risi, C.; Polini, G. P.; Zanirato, V. *Tetrahedron Lett.* **1997**, *38*, 681−684.
(7) Hansen, M. T. S.; Carteau, S.; Hoffmann, C.; Li, L.; Bushman, F. In *Genetic Engineering. Principles and Methods*; Setlow, J. K., Ed.; Plenum Press: New York, 1998; pp 41−62.
(8) Craigie, R.; Mizuuchi, K.; Bushman, F. D.; Engelman, A. *Nuc. Acids Res.* **1991**, *19*, 2729−2734.