JOC Article

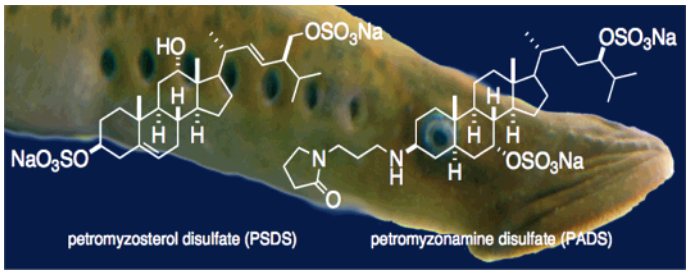
Details of the Structure Determination of the Sulfated Steroids PSDS and PADS: New Components of the Sea Lamprey (*Petromyzon marinus*) Migratory Pheromone

Thomas R. Hoye,*[,†] Vadims Dvornikovs,[†] Jared M. Fine,[‡] Kari R. Anderson,[†] Christopher S. Jeffrey,[†] David C. Muddiman,[§] Feng Shao,[†] Peter W. Sorensen,[‡] and Jizhou Wang[†]

Department of Chemistry, 207 Pleasant Street, SE, University of Minnesota, Minneapolis, Minnesota 55455, Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota, St. Paul, Minnesota 55108, and W.M. Keck FT-ICR MS Laboratory, Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695.

hoye@chem.umn.edu

Received May 7, 2007



The discovery of two new components of the migratory pheromone used by sea lamprey to guide adults to spawning grounds was recently reported. These hold promise for use in a pheromone-based control program for this species, an invasive pest in the Great Lakes. Details of the structure determination of these steroidal bis-sulfates [petromyzosterol disulfate (PSDS, **2**) and petromyzonamine disulfate (PADS, **3**)] are described here. Pattern matching of ^{1}H NMR data was particularly valuable. This involved comparison of spectra of the natural samples of **2** and **3** with those of appropriate steroidal analogues [e.g., petromyzonol sulfate (PS, **1**, a previously known sea lamprey bile acid derivative that is a third component of the migratory pheromone), cholesterol sulfate (**6**), and squalamine (**8**)] and model compounds containing the unprecedented aminolactam substructure present in **3**. The logic underlying the iterative analyses used is presented.

Introduction

Lamprey are among the most ancient (>400 million years) extant vertebrates. The sea lamprey, *Petromyzon marinus* ("stone-sucker"), is native to the Atlantic Ocean. Over the course of the first half of the last century it colonized the Great Lakes of North America. Because it is parasitic and has no significant natural predators in the Great Lakes, the sea lamprey has had a devastating impact on fisheries there. As a result, it is the target of one of the largest control programs in the world for an aquatic invasive species. We recently communicated the isolation and structures of three natural products that comprise the principal components of a pheromone (i.e., a chemical signal, comprised of one or more compounds, that transmits a message between members of the same species) that are released by larvae of this species and used by adults to locate suitable streams for reproduction.[1] These compounds hold considerable promise for use as agents in an environment-friendly sea lamprey control program.[2] We report here the strategies by which the structures of these three sulfated steroid derivatives (Figure 1) were determined. The first, petromyzonol sulfate (PS, **1**), was known;[3]

[†] Department of Chemistry, University of Minnesota.
[‡] Department of Fisheries, Wildlife, and Conservation Biology; University of Minnesota.
[§] W.M. Keck FT-ICR MS Laboratory, Department of Chemistry, North Carolina State University.

(1) Sorensen, P. W.; Fine, J. M.; Dvornikovs, V.; Jeffrey, C. S.; Shao, F.; Wang, J.; Vrieze, L. A.; Anderson, K. R.; Hoye, T. R. *Nat. Chem. Biol.* **2005**, *1*, 324−328.

10.1021/jo070957l CCC: $37.00
Published on Web 08/24/2007

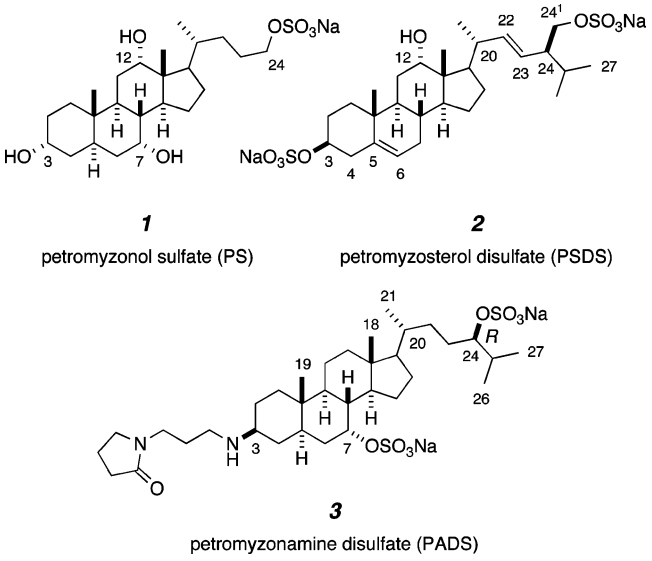


FIGURE 1. Structures of PS (**1**), PSDS (**2**), and PADS (**3**), the three principal components of the sea lamprey migratory pheromone.

the other two, petromyzosterol disulfate (PSDS, **2**) and petromyzonamine disulfate (PADS, **3**), were previously unknown.

Adult sea lamprey detect and are attracted to **1**−**3** at concentrations as low as 10^{-13} M. Sea lamprey larvae (≤20 cm l), which are filter feeders that live in the bottoms of a limited set of streams entering the Great Lakes, release these (and other) steroidal components. Adults presumably have evolved to recognize the pheromone as an indicator of good larval habitat and are guided, thereby, to their upstream spawning grounds. While compound **1** is quite prevalent in larval livers and gall bladders [simple trituration of several gall bladders with methanol provided several milligrams of PS, (**1**)], components **2** and **3** were isolated after the larvae had released them into the water of their laboratory holding tanks. Organic components present in that larval holding water were adsorbed onto Amberlite XAD7HP and eluted with methanol. The eluate was concentrated and fractionated on a bed of Sephadex G15 resin. Fractions found to contain the pheromone components by olfactory activity and ESI-MS were subjected to reversed phase C18 HPLC (1−4 min, 40% methanol, 4−56 min, 40−60% methanol, 56−60 min, 60−100% methanol) to provide, in order of elution, components **2**, **3**, and **1**. Ultimately, 8000 liters of holding water containing ∼35 000 larvae were processed to provide ∼600 and ∼700 μg of the new components **2** and **3**, respectively [along with ∼200 μg of PS (**1**)].

Results and Discussion

PS (1, Petromyzonol Sulfate). The molecular weight of isolated PS (**1**) was determined by using negative ion electrospray ionization coupled to an ion trap mass analyzer (ESI-ITMS) giving rise to a $(M - H^+)^-$ of 473.3, its retention time was determined with HPLC, and its ^{1}H NMR spectrum was obtained. These data were all consistent with an authentic sample. Since neither the ^{1}H NMR data for PS nor the full assignment for petromyzonol (P, **4**), which is the nonsulfated tetraol analogue of PS (**1**), are described elsewhere in the literature, we report them here. The knowledge we gained in the course of that interpretation proved informative in subsequent analyses and assignment of PSDS (**2**) and PADS (**3**) spectra and structures (see below).

The complete set of proton and carbon NMR data for both **1** and **4** are listed in Table 1. Analysis of the multiplicities of the three carbinol methine protons (Figure 2) allowed us to assign

TABLE 1. ^{1}H and ^{13}C NMR (CD_3OD, 800 MHz) Assignments [Chemical Shifts (δ) in ppm and, for ^{1}H, Coupling Constants (*J*, in parentheses) in Hertz (Hz)] for Petromyzonol (P, 4) and Petromyzonol Sulfate (PS, 1)[a]

	petromyzonol (**4**)		petromyzonol sulfate (**1**)	
no.	^{1}H	^{13}C	^{1}H	^{13}C
1β	1.40, m	32.9	1.39, m	32.9
1α	1.44, m		1.43, m	
2β	1.62, m[b]	29.3	1.62, m[b]	29.2
2α	1.66, m[b]		1.66, m[b]	
3β	3.97, dddd (3, 3, 3, 3)	66.9	3.97, m	66.9
4β	1.49, ddd (14.0, 14.0, 3.0)	36.4	1.31, m 1.49, ddd (14.0, 14.0, 3.0)	36.3
4α	1.31, m			
5α	2.15, dddd (13, 13, 3, 3)	32.5	2.14, dddd (13.0, 13.0, 3.0, 3.0)	32.5
6β	1.42, m	37.5	1.42, m	37.4
6α	1.33, ddd (14, 3, 3)		1.32, ddd (14.0, 3.0, 3.0)	
7β	3.78, ddd (3.0, 3.0, 3.0)	68.6	3.78, ddd (3.0, 3.0, 3.0)	68.5
8β	1.47, ddd (11.5, 11.5, 2.5)	40.9	1.46, ddd (11.5, 11.5, 3.0)	40.9
9α	1.66, m	40.2	1.66, m	40.2
10		36.6		36.6
11β	1.55, ddd (14.5, 14.5, 3.0)	29.0	1.55, m	29.2
11α	1.66, m		1.66, m	
12β	3.94, dd (3.0, 3.0)	73.7	3.93, dd (3.0, 3.0)	73.7
13		47.2		47.1
14α	1.93, ddd (12.5, 11.5, 7.5)	43.0	1.93, ddd (12.0, 12.0, 7.5)	42.9
15β	1.10, dddd (12, 12, 12, 6)	23.9	1.11, dddd (12.0, 12.0, 12.0, 6.0)	23.9
15α	1.76, ddd (12, 9, 7.5, 3.5)		1.75, m	
16β	1.89, dddd (13.0, 9.5, 6.5, 3.5)	28.4	1.89, m 1.28, m	28.4
16α	1.28, nfom			
17α	1.85, ddd (9, 9, 9)	48.0	1.84, ddd (9.5, 9.5, 9.5)	48.0
18	0.71, s	12.8	0.71, s	12.7
19	0.81, s	10.3	0.80, s	10.2
20	1.39, m	36.8	1.41, m	36.6
21	1.02, d (6.5)	17.8	1.02, d (6.5)	17.6
22a	1.09, m	33.0	1.15, m	32.8
22b	1.48, m		1.54, m	
23a	1.43, m	30.2	1.56, m	27.0
22b	1.63, m		1.75, m	
24a	3.50, ddd (10.5, 6.5, 6.5)	63.4	3.97, m	69.4
24b	3.51, ddd (10.5, 6.5, 6.5)			

[a] ^{1}H and ^{13}C chemical shift assignments were made on the basis of COSY, HMQC, and HMBC analyses. ^{1}H chemical shift values (δ) for resonances designated as "m" were taken from COSY or HMQC cross-peaks; those having assigned multiplicity (e.g., ddd) were taken directly from the 1D ^{1}H spectrum. ^{13}C chemical shift values were taken from HMQC and/or HMBC cross-peaks. [b] Resonances assigned to Hα and Hβ may be interchanged.

(2) For an overview that discusses various scientific and policy issues related to the development and implementation of sea lamprey pheromones as part of an integrated pest management strategy, see: Sorensen, P. W.; Hoye, T. R. *J. Fish Biol.* In press.

(3) Haslewood, G. A. D.; Tokes, L. *Biochem. J.* **1969**, *114*, 179−184.

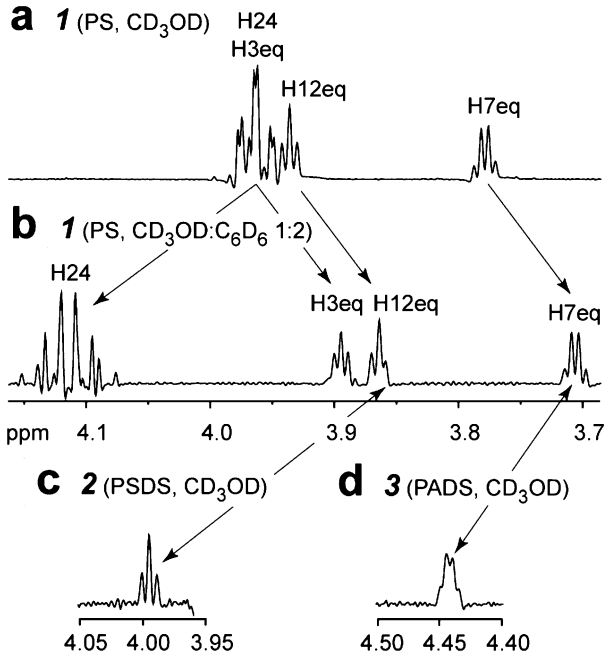


FIGURE 2. A portion of the resolution enhanced ^{1}H NMR spectrum (a) of PS (**1**) in CD_3OD, (b) of PS (**1**) in 2:1 benzene-d_6:CD_3OD, (c) of PSDS (**2**), and (d) of PADS (**3**). The pattern of these methine protons (dd for **2** and ddd for **3** with all J values = 2.5−3 Hz) vis-à-vis those in **1** led to assignment of their location and, therefore, the site and configuration of oxygenated carbons in **2** and **3**.

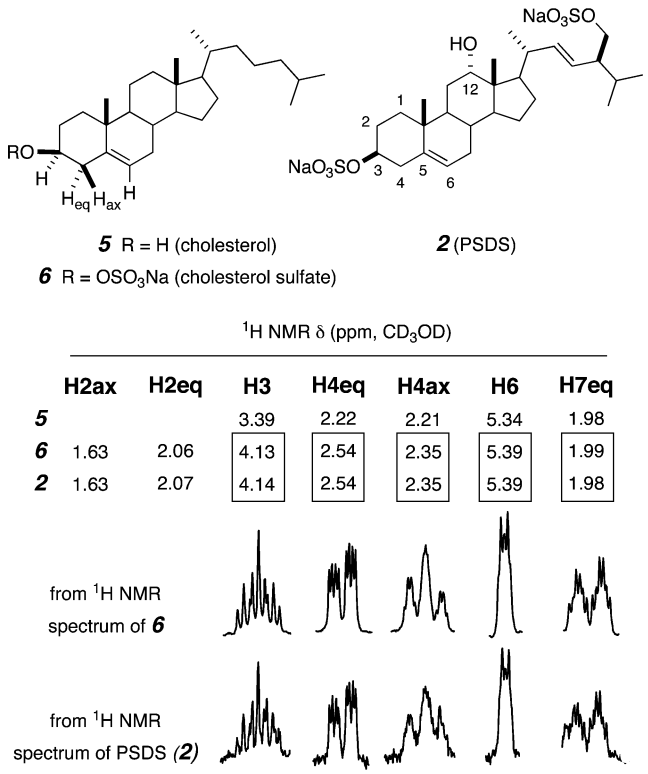


FIGURE 3. Chemical shift comparisons and pattern matching of resonances for selected protons in the ^{1}H NMR spectra of cholesterol sulfate (**6**) vs PSDS (**2**).

the chemical shift of each of these equatorially oriented (as reinforced by the vicinal J values of 2.5−3 Hz, throughout) protons. Namely, the dddd (app pentet) at δ 3.90 belongs to H3eq; the ddd (app quartet) at 3.71 to H7eq; and the dd (app triplet) at 3.86 to H12eq. Recognition of these patterns played a critical role in locating the position of carbocyclic oxygenation in the unknown components **2** and **3**. That is, a downfield dd (at δ 3.99 ppm) in the spectrum of PSDS allowed us to place an axial hydroxyl group at C12 in **2** and a downfield ddd (at δ 4.45 ppm) in the spectrum of PADS implicated the (sulfated) axial hydroxyl group at C7 in **3**.

PSDS (2, Petromyzosterol Disulfate). The first of the two unknown components, PSDS (**2**), gave a negative pseudomolecular ion of *m*/*z* 589.4 [ESI-ITMS, $(M - H^+)^-$] and an ion at *m*/*z* 491.3; the latter was attributed to the loss of H_2SO_4 and therefore suggested the presence of at least one sulfate in PSDS. A doubly charged species at *m*/*z* 294.2, indicative of a second ionizable group (e.g., sulfate, sulfonate, or carboxylate), was also observed. The high-resolution negative ion mass spectral data [(ESI-TOFMS) 589.2498 $(M - H^+)^-$, 491.2813 $(M - H_2SO_4 - H^+)^-$] pointed to a molecular formula of $C_{28}H_{46}O_9S_2$, which implied six degrees of unsaturation for PSDS (**2**).

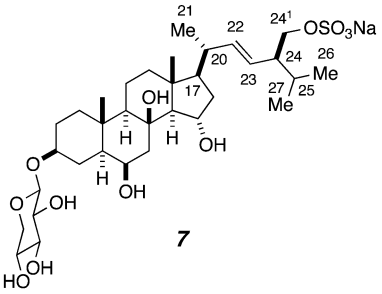
The ^{1}H NMR data suggested that **2** was a steroid [two methyl singlets (δ 0.76 and 1.03) and three methyl doublets (δ 0.87, 0.95, 1.09)]. The three most downfield resonances at δ 5.39 (br s, 1H), 5.34 (dd, J = 15.1, 9.0 Hz, 1H), and 5.21 (dd, J = 15.2, 9.2 Hz, 1H) were indicative of three olefinic protons, the first attached to an endocyclic trisubstituted double bond, and the other two to the same, vicinally disubstituted *trans*-alkene. Each allylic carbon of the latter was branched, bearing only one proton. The other two downfield resonances at δ 4.14 and 3.99 were consistent with the presence of two oxygenated methine protons and the resonance at δ 3.95 (d, J = 6.5 Hz, 2H) with an oxygenated pair of methylene protons.

Upon noticing similarities between the ^{1}H NMR spectrum of **2** and that of cholesterol (**5**, in $CDCl_3$), we acquired the spectrum of **5** in CD_3OD. The *multiplicities* of the resonance of the H6 alkene proton at δ 5.34 and of the H3 methine at δ 3.39 matched well with those of the PSDS (**2**) alkene proton at δ 5.39 and methine at δ 4.14, respectively. We hypothesized that the differences in the chemical shifts were due to the presence of a 3β-*O*-sulfate in **2**. Therefore, cholesterol sulfate (**6**)[4] was prepared and its ^{1}H NMR spectrum (in CD_3OD) exhibited an excellent match in the *multiplicity and chemical shift* of protons H2−H7 (Figure 3). From this comparison we deduced the identity between the A/B ring substructure of cholesterol sulfate (**6**) and PSDS (**2**).

Comparison of the multiplicity of the H12β (ddd or app t) in PS (**1**) (cf. panel a, Figure 2) with the resonance at δ 3.99 in PSDS (**2**) (cf. panel c, Figure 2) allowed us to locate another site of oxygenation as 12α-OH. This assignment was confirmed by HMBC (see below), in which a strong correlation between (CH_3)18 (δ_H 0.76) and C12 (δ_C 73.7) was observed.

There was no evidence in the NMR spectrum for additional oxidation in the tetracyclic steroid framework. Therefore, we assumed the composition of the C17 side chain to be a $C_9H_{17}O_4S$ subunit. Alkene proton resonances at 5.21 and 5.34 (each with J = 15 and 9 Hz) suggested the presence of a disubstituted *trans*-alkene having a single proton at each allylic carbon (R^1R^2-CH−HC=CH−CHR^3R^4). The two-proton doublet (J = 6.5 Hz) at 3.98 ppm suggested the presence of a sulfated primary alcohol

(4) Dusza, J. P.; Joseph, J. P.; Bernstein, S. *Steroids* **1985**, *45*, 317−323.



7

^{1}H NMR δ (ppm, CD_3OD)

	H20	H21	H22	H23	H24	H24^1	H25	H26/H27
7	2.06	1.03	5.33	5.24	2.15	3.99	1.86	0.89/0.95
2	2.05	1.09	5.34	5.21	2.12	3.95	1.86	0.87/0.93

FIGURE 4. Comparison of chemical shifts of protons residing in the side chains of asteriidoside L (**7**) and PSDS (**2**), from which the constitution of that subunit in **2** was deduced.

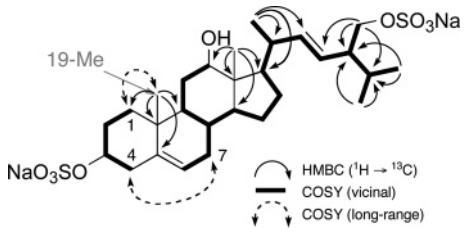


FIGURE 5. Key skeletal connectivities for PSDS (**2**), as deduced from COSY and HMBC spectra.

adjacent to a single methine proton ($R^5R^6CHCH_2OSO_3Na$). Three, three-proton doublets (J = 6.5−7 Hz) indicated the presence of three methyl groups, each having a single vicinal proton. COSY and HMBC correlations supported a $\Delta^{22(23)}$ location for the double bond, a sulfated hydroxymethyl substituent attached to C24, and methyl branching at C20 and C25. To accommodate these facts, we proposed the side chain shown in Figure 4. This substructural unit is present in the sulfated ergostane asteriidoside L (**7**).[5] The ^{1}H NMR shift data for all of the side chain protons in PSDS (**2**) vs **7** (Figure 4) strongly support this constitutional assignment for **2**. The question of configuration at C24 is not easily addressed. While it is tempting to propose that **2** and **7** also share the same *configuration* at C24, we have been unable to locate any reported examples of C24^1 sulfated ergostanes that are epimeric at C24 in order to assess whether chemical shift differences would be distinctive. Moreover, since it is known that there is a high degree of similarity of the proton and carbon NMR spectra for each of the C24-epimers of several pairs of model ergostan-24^1-ols that had free hydroxyl groups at C24^1, an unambiguous answer to this question must await resolution by synthesis, which we are pursuing.

Finally, COSY and (selected) HMBC correlations observed for PSDS (**2**) are summarized in Figure 5. They are in full accord with our structure of **2**. The long-range ^{1}H−^{1}H correlations (dashed arrows) are consistent with (i) the known long-range axial methyl (C19) to axial vicinal proton (H1ax) coupling within rigid cyclohexane derivatives[6] and (ii) allylic−allylic coupling between protons on C4 and C7.

PADS (3, Petromyzonamine Disulfate). The second of the two unknown components, PADS (**3**), had a nominal molecular weight of 704 Da, as indicated by negative ion ESI-ITMS with m/z values of 703.3 $(M - H^+)^-$ and 725.3 $(M - 2H^+ + Na^+)^-$. A doubly charged species with an m/z of 351.3 $[(M - 2H^+)^{2-}]$, indicative of a second ionizable group (e.g., sulfate, sulfonate, or carboxylate), was also observed. ESI-ITMS/MS analysis of the deprotonated molecular ion gave rise to product ions with m/z values of 605.5 and 623.3, suggestive of loss of H_2SO_4 (−98) and SO_3 (−80), respectively. Positive ion ESI-ITMS data resulted in m/z values of 705.2 $(M + H^+)^+$ and 749.3 $(M - H^+ + 2Na^+)^+$, which reinforced the neutral nominal molecular weight of the intact species as 704 Da. Positive ion MALDI-TOFMS data exhibited ions at m/z ratios (tentative assignment) of 607 $(M - H_2SO_4 + H^+)^+$, 509 $(M - 2H_2SO_4 + H^+)^+$, and 527 $(M - H_2SO_4 - SO_3 + H^+)^+$, which suggested the presence of at least two sulfate groups in PADS.

A comparison of ESI-ITMS data for two samples of PADS in CH_3OH vs CD_3OD proved informative. The negative ion observed at m/z 725.3 in CH_3OH corresponding to $(M - 2H^+ + Na^+)^+$ appeared, instead, at m/z 726.3 in CD_3OD; an analogous change was observed for the dianion (m/z 351.2 vs 351.7). This implied the presence of one (and only one) remaining, exchangeable proton (OH or NH) in these anions. The positive ion at m/z 749.3 in CH_3OH corresponding to $(M - H^+ + 2Na^+)^+$ appeared, instead, *two m/z* units higher at 751.3 in CD_3OD. This implied the presence of two remaining, exchangeable protons in these cations, an observation most consistent with the presence of a secondary amine group (RR′NH) in PADS (**3**).

Negative ion MS analysis (m/z 703.3634, 605.3968, 623.4081, and 351.1738) implicated $C_{34}H_{60}N_2O_9S_2$ as the molecular formula of PADS (**3**). More precise mass characterization of PADS was carried out with high resolving power, high mass measurement accuracy dual ESI-Fourier transform ion cyclotron resonance mass spectrometry (dualESI-FTICRMS).[7] Defining elemental composition of unknowns of increasing molecular weight is quite challenging. However, as previously demonstrated by Marshall and co-workers, using high-resolution mass spectrometry, one can directly count the number of sulfur atoms in a molecule.[8] The combination of high mass measurement accuracy and the ability to determine accurately the number of sulfurs is a powerful tool for making confident assignments of elemental compositions of unknowns. Negative ion dualESI-FTICRMS (9.4 T) measurements (Figure 6) with a resolving power of 180 000_{FWHM} *directly* revealed that there were 2 sulfurs present [((M + 2)/M) × 100 = 8.6%] in PADS. Since there are 2 sulfurs and the measurement has a mass accuracy of ±1 ppm (m/z measured = 703.3663_1, m/z theoretical = 703.3667), $C_{34}H_{60}N_2O_9S_2$ was established as the definitive molecular formula of PADS, a composition that indicates six degrees of unsaturation.

^{1}H NMR spectroscopic analysis indicated that PADS (**3**), like PSDS (**2**), was a steroid; the spectrum had two methyl singlets

(5) De Marino, S.; Iorizzi, M.; Palagiano, E.; Zollo, F.; Roussakis, C. *J. Nat. Prod.* **1998**, *61*, 1319−1327.

(6) Jackman, L. M.; Sternhell, S. In *Applications of Nuclear Magnetic Resonance Spectroscopy in Organic Chemistry*; International Series of Monographs in Organic Chemistry, Vol. 5; Pergamon Press: Oxford, UK, 1969; pp 335−338.

(7) Nepomuceno, A. I.; Muddiman, D. C.; Bergen, H. R.; Craighead, J. R.; Burke, M. J.; Caskey, P. E.; Allan, J. A. *Anal. Chem.* **2003**, *75*, 3411−3418.

(8) Shi, S. D.; Hendrickson, C. L.; Marshall, A. G. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11532−11537.

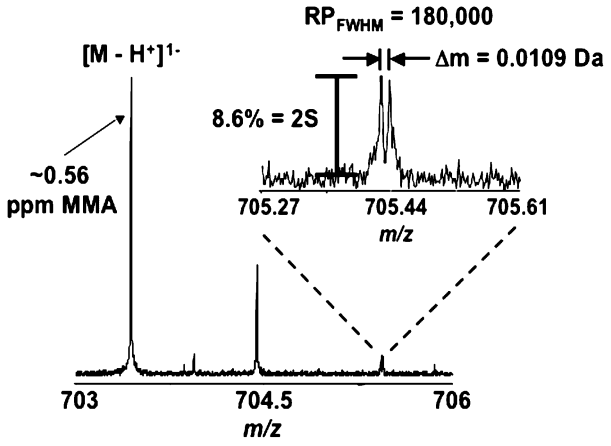
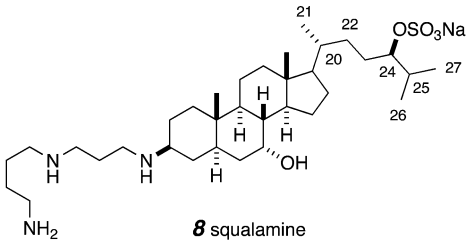


FIGURE 6. Negative-dual electrospray ionization FT-ICR mass measurement of PADS (**3**).



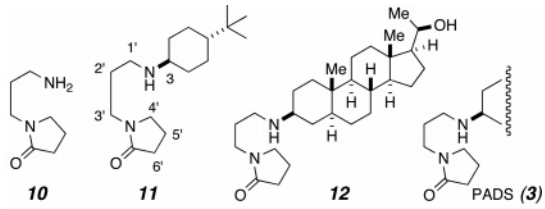
^{1}H NMR δ (ppm, CD_3OD)

	H20	H21	H22a/b	H23a/b	H24	H25	H26/H27
8	1.44	0.97	1.31/1.54	1.53/1.71	4.13	2.06	0.94/0.96
3	1.43	0.95	1.15/1.53	1.55/1.71	4.12	2.02	0.93/0.95
9	1.41	0.95	n.d.	1.57/1.67	4.12	2.02	0.91/0.96

FIGURE 7. Comparison of chemical shifts of protons residing in the side chains of squalamine (**8**), natural PADS (**3**), and synthetic 24-*epi*-PADS (**9**).

(δ 0.69 and 0.84) and three methyl doublets (δ 0.93, 0.95, 0.95). Literature searches of nitrogenous steroids pointed us to squalamine (**8**)[9] (and its congeners[10]), an antibiotic produced by the dogfish shark, *Squalus acanthias*, another ancient fish. The previously reported structural studies that were carried out with squalamine were instructive. COSY and HMBC analyses and chemical shift comparison established that the side chain in **3** is the same as that of squalamine (**8**) (Figure 7) and includes a sulfate at C24. The resonance for a proton on an oxygenated carbon atom (δ 4.45 ppm) within the carbocyclic framework was particularly revealing. Namely, the multiplicity of this methine proton (ddd, J = 2.3, 2.3, 2.3 Hz) was similar to that of H7eq in PS (**1**) (δ 3.78 Figure 2). The analogous H7eq proton in squalamine (**8**) appears at δ 3.79 ppm, also as a narrow multiplet ($\sum J \approx 8$ Hz). The significant downfield shift of H7eq in **3** led us to place the second site of sulfation as the C7ax oxygen atom.

The partial structure comprising the steroidal skeleton, the squalamine-like, 24-sulfated C20−C27 side chain, and the C7 sulfate had an atomic composition of $C_{27}H_{47}O_8S_2$; what remained was a fragment containing the residual atoms $C_7H_{13}N_2O$. There were no obvious resonances in the ^{1}H NMR spectrum from protons associated with any additional oxygen substituents within the tetracyclic steroid skeleton. Moreover, the presence of two independent sets of trimethylene ($-CH_2CH_2CH_2-$) subunits was evident in the COSY and ^{1}H NMR spectra. We deemed it most likely that the entire set of $C_7H_{13}N_2O$ atoms was present as a single subunit attached at C3, having an unsaturation value of two and no asymmetry. A methine resonance at δ 2.70 ppm (and δ 2.50 ppm, dddd, J = 11.5, 11.5, 4.0, 4.0 Hz, free base form) suggested that the linkage to C3 was through a basic nitrogen, equatorially oriented.



^{1}H NMR δ (ppm, CD_3OD)

ammonium ions	H1'	H2'	H3'	H4'	H5'	H6'	H3
***10*·H⁺**	2.94	1.94	3.42	3.55	2.10	2.47	na
***11*·H⁺**	3.00	1.94	3.41	3.51	2.09	2.43	2.99 ax
***12*·H⁺**	3.00	1.94	3.41	3.51	2.09	2.42	3.08 ax
***PADS*·H⁺**	2.99	1.90	3.39	3.50	2.09	2.42	3.08 ax
***cis11*·H⁺**	3.02	2.02	3.43	3.52	2.10	2.43	3.38 eq
free base							
11	2.64	1.76	3.34	3.47	2.06	2.39	2.45 ax
12	2.63	1.76	3.34	3.47	2.06	2.39	2.54 ax
PADS	2.54/2.58	1.73	3.31	3.46	2.05	2.38	2.50 ax
cis-11	2.53	1.76	3.35	3.47	2.05	2.38	2.77 eq

FIGURE 8. Comparison of chemical shifts of protons residing in the aminopropyl pyrrolidinone side chains of model amines **10**−**12** with those of PADS (**3**).

An *N*-(3-aminopropyl)pyrrolidin-2-one substructure was proposed to account for all of the above features. The pattern of the methylene resonances in the ^{1}H NMR spectrum of the commercially available, parent primary amine, *N*-(3-aminopropyl)pyrrolidin-2-one (**10**), is similar to that ascribed to the remaining subunit in **3**. To test this assignment through experimentation, we prepared model compounds containing a 3-(2-oxo-1-pyrrolidinyl)propylamino substituent. Namely, reductive amination of 4-*tert*-butylcyclohexanone and of 5α-pregnane-3,20-dione gave, predominantly, the equatorial amines **11** and **12**, respectively, through preferential axial attack by sodium borohydride on the imines formed by preincubation of each ketone with **10** in methanol or ethanol. The spectra from these amines gave considerable support to our hypothesis that **3** was the structure of PADS. The chemical shifts of the protons associated with these amine side chains are tabulated in Figure 8.

During the isolation phase we obtained different samples of PADS (**3**), varying in their purification history, that gave different chemical shifts for, principally, protons H3 and H1′. This nagging issue was clarified once a synthetic sample of PADS (**3**) was obtained (see below). Spectra of fully protonated, free base, and even partially protonated samples of the natural material gave rise to, now, understandable and reproducible

(9) (a) Moore, K. S.; Wehrli, S.; Roder, H.; Rogers, M.; Forrest, J. N., Jr.; McCrimmon, D.; Zasloff, M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1354−1358. (b) Wehrli, S. L.; Moore, K. S.; Roder, H.; Durell, S.; Zasloff, M. *Steroids* **1993**, *58*, 370−378.

(10) Rao, M. N.; Shinnar, A. E.; Noecker, L. A.; Chao, T. L.; Feibush, B.; Snyder, B.; Sharkansky, I.; Sarkahian, A.; Zhang, X.; Jones, S. R.; Kinney, W. A.; Zasloff, M. *J. Nat. Prod.* **2000**, *63*, 631−635.

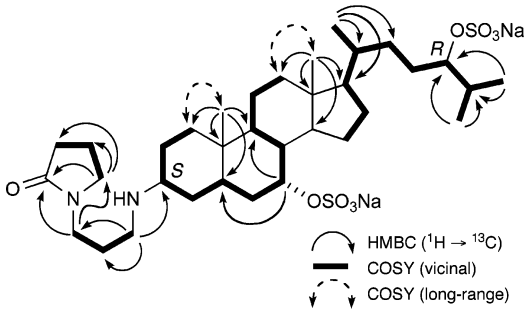


FIGURE 9. Key skeletal connectivities for PADS (**3**), as deduced from COSY and HMBC spectra.

chemical shift patterns for protons close to the amino group. These differences can be judged from the data for the fully protonated forms (of models **10**−**12** vs **3**; top entries) vs the free base forms (of **12** vs **3**; bottom entries) included in Figure 8.

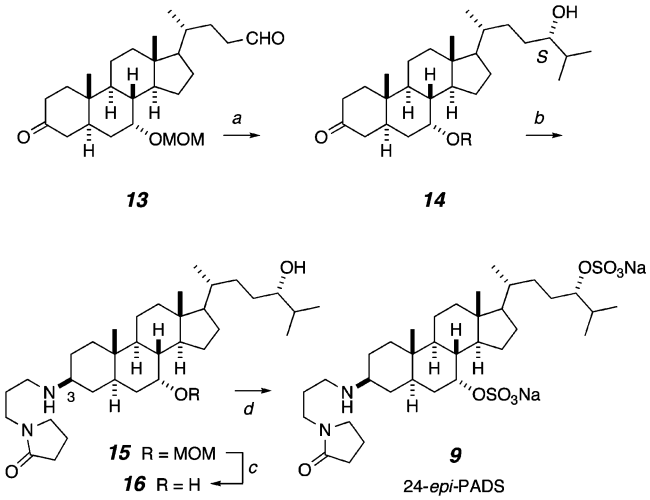
Finally, the configuration of C3 in PADS (**3**) was revealed by both the chemical shift of the H3 resonance (cf. data in gray box for **11**·H^+, having an axial proton vs *cis*-**11**·H^+ having an equatorial proton vs **3**·H^+) and its coupling constants (two large and two small for **11**, **12**, and **3** vs four small for *cis*-**11**, regardless of protonation state).

Parallel HMQC and HMBC analyses (see Figure 9) of PADS (**3**) allowed identification and assignment of all 34 carbon atom chemical shifts. In addition, the connectivity of the entire skeleton deduced from these analyses was fully consistent with structure **3**. To prove this structural assignment for PADS (**3**), we prepared a sample by chemical synthesis, starting from chenodeoxycholic acid, and showed its identity with the natural sample.[1] Synthetic PADS (**3**) was found to (i) elicit electrophysiological activity from the sea lamprey olfactory system (EOG response)[11] and (ii) attract[12] adult sea lamprey in behavioral test mazes with activity comparable to that of the isolated PADS.

We also prepared 24-*epi*-PADS (**9**), whose ^{1}H and ^{13}C NMR spectral data were slightly but diagnostically different from those of **3**, primarily in the side chain methyl chemical shifts (cf., Figure 7 and the Supporting Information). The synthesis of **9** is summarized in Scheme 1. Addition of isopropylmagnesium chloride to the known aldehyde **13**[13] gave an ca. 1:1 mixture of epimeric C24-carbinols **14** and the corresponding, known 24*R*-epimer[13] (not shown). Remarkably, these were relatively easily separable by silica gel chromatography. Mosher ester analysis (Supporting Information) was used to independently establish the configuration at C24 for each.[14] Reductive amination of the ketone in **14** with amine **10** and MOM ether cleavage gave diol **16** via **15**. Sulfation of both hydroxyl groups completed the synthesis of 24-*epi*-PADS (**9**).

Finally and interestingly, one can envision that the *N*-(3-aminopropyl)pyrrolidin-2-one subunit, unique to **3** among natural products, arises from oxidation of the primary amino group in squalamine (**8**). Such an oxidation event is precedented for the parent triamine, spermidine [$NH_2(CH_2)_4NH(CH_2)_3NH_2$].[15] This leads to our suggestion[1] that squalamine (**8**), isolated from the dogfish shark, another ancient organism, is also likely present in sea lamprey and that it is the biosynthetic precursor to PADS (**3**).

SCHEME 1[a]



[a] Reagents and conditions: (a) *i*PrMgCl, THF, −78 °C; (b) (i) *N*-(3-aminopropyl)pyrrolidin-2-one, 3 Å MS, MeOH, (ii) $NaBH_4$, −78 °C; (c) HCl (anhyd), MeOH; (d) py·SO_3, pyridine, 50 °C.

Experimental Section

(5α,7α,24*S*)-24-Hydroxy-7-(methoxymethoxy)cholestan-3-one (14). Isopropyl magnesium chloride (0.27 mL, 1.77 M in Et_2O, 0.47 mmol) was added dropwise to a solution of aldehyde **13**[13] (100.0 mg, 0.239 mmol) in THF (3.0 mL) at −78 °C. The solution was stirred at this temperature for 2 h and then quenched by addition of saturated aqueous NH_4Cl at −78 °C. The aqueous layer was extracted with ethyl acetate (3 × 10 mL). The combined organic phase was washed with saturated $NaHCO_3$ and brine, dried over $MgSO_4$, filtered, and concentrated at reduced pressure. Purification by flash chromatography (SiO_2, 2:1, hexanes:ethyl acetate) gave a mixture of the C24-epimeric alcohols (66.4 mg, 56%). The 24*S*-epimer (**14**) was separated from the 24*R*-epimer by medium-pressure liquid chromatography (SiO_2, 3:1, hexanes:ethyl acetate). ^{1}H NMR (500 MHz, $CDCl_3$) δ 4.69 (d, J = 6.9 Hz, 1H), 4.60 (d, J = 6.9 Hz, 1H), 3.62 (br q, J = 3.3 Hz, 1H), 3.35 (s, 3H), 3.32 (br ddd, J = 4, 4, 8 Hz, 1H), 2.38 (ddd, J = 15.6, 13.6, 6.6 Hz, 1 H), 2.31 (dddd, J = 15.5, 4.9, 2.1 Hz, 1H), 2.26 (app t, J = 14.2, 1H), 2.10−1.92 (m, ca. 5H), 1.86 (dddd, J = 13.3, 9.6, 9.6, 6.4 Hz, 1H), 1.73−1.02 (m, ca. 25H), 1.01 (s, 3H), 0.94 (d, J = 6.6 Hz, 3H), 0.93 (d, J= 6.9 Hz, 3H), 0.90 (d, J = 6.8 Hz, 3H), and 0.67 (s, 3H). ^{13}C NMR (125 Hz, $CDCl_3$) δ 212.1, 96.4, 77.1, 77.0, 75.1, 56.1, 55.9, 45.7, 44.4, 42.7, 39.8 (2×), 39.4, 38.4, 38.3, 35.9, 35.8, 34.0, 33.7, 32.2, 30.7, 28.4, 23.9, 19.1, 18.8, 17.4, 12.0, and 10.7. HR ESI-MS calcd for $C_{29}H_{50}NaO_4$ $[M + Na]^+$ 485.3601, found 485.3590. IR (film) 2935, 2869, 1713, 1464, 1445, 1372, 1260, 1144, 1090, and 1038 cm^{-1}. TLC R_f 0.50 (2:1, hexanes:ethyl acetate).

(3β,5α,7α,24*S*)-1-[3-[[7-(Methoxymethoxy)-24-hydroxycholestan-3-yl]amino]propyl]-2-pyrrolidinone (15). 1-(3-Aminopropyl)pyrrolidin-2-one (**10**) (5.0 μL, 0.035 mmol) and 3 Å MS (400 mg) were added to a solution of ketone **14** (5.1 mg, 0.011 mmol) in anhydrous methanol (1 mL). This mixture was stirred

(11) Christopher S. Jeffrey and Peter W. Sorensen, unpublished observations.

(12) Fine, J. M. Ph.D. Thesis, University of Minnesota, St. Paul, MN, 2006.

(13) Zhang, D.-H.; Cai, F.; Zhou, X.-D.; Zhou, W.-S. *Org. Lett.* **2003**, *5*, 3257−3259.

(14) Specifically, each of **14** and its 24*R*-diastereomer was derivatized with *R*- and *S*-MTPA-Cl to give respectively the *S*- and *R*-MTPA ester derivatives (see the Supporting Information).

(15) Seiler, N.; Knödgen, B.; Haegele, K. *Biochem. J.* **1982**, *208*, 189−197.

overnight at rt and then cooled to −78 °C. $NaBH_4$ (10.2 mg, 0.27 mmol) was added. After being stirred at −78 °C for 2 h the reaction mixture was quenched with phosphate buffer (pH 7) and allowed to warm to room temperature. The resulting suspension was filtered through celite with exhaustive washing with MeOH and CH_2Cl_2. The filtrate was concentrated under reduced pressure. The aqueous phase was extracted with 10% MeOH in $CHCl_3$ (3 × 15 mL) and the combined organic extracts were dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Silica gel chromatography (90:9:1;CH_2Cl_2:MeOH:NH_4OH) provided amine **15** (3.0 mg, 0.0051 mmol, 45%) as a colorless oil. ^{1}H NMR (300 MHz, $CDCl_3$; assignments supported by COSY) δ 4.70 (d, J = 6.8 Hz, 1H, $-OCH_aH_bOMe$), 4.61 (d, J = 6.8 Hz, 1H, OCH_aH_bOMe), 3.61 (br ddd, J = 2, 2, 2 Hz, 1H, H7), 3.38 (s, 3H, OMe), 3.43−3.25 (m, ca. 5H), 2.62 (t, J = 7.0 Hz, 2H, H1′), 2.46 (1H, br m, 1H, H3), 2.39 (t, J = 8.2 Hz, 2H, H6′), 2.03 (pent, J = 7.2 Hz, 2H, H5′), 2.0−0.98 (m, ca. 29H), 0.93 (d, J = 6.8 Hz, 6H, 2×Me), 0.89 (d, J = 6.9 Hz, 3H), 0.79 (s, 3H, H19), and 0.64 (s, 3H, H18). ^{13}C NMR ($CDCl_3$, 75 MHz) δ 175.4, 96.2, 77.4, 77.0, 75.4, 57.7, 56.1, 55.9, 50.3, 47.4, 46.4, 44.1, 42.7, 40.5, 39.9, 39.6, 38.2, 37.4, 36.3, 33.8, 33.3, 32.4, 31.2, 31.0, 29.9, 29.2, 28.4, 28.1, 23.9, 21.1, 19.3, 19.0, 18.1, 16.9, 12.0, and 11.6. ESI-MS calcd for $C_{36}H_{65}N_2O_4$ [M + H]$^+$ 589.4939, found 589.4946. TLC R_f 0.75 (80:18:2, $CHCl_3$:MeOH:NH_4OH).

1-[3-[[(3β,5α,7α,24S)-7,24-Dihydroxycholestan-3-yl]amino]-propyl]-2-pyrrolidinone (16). Amine **15** in 3:1 MeOH:CH_2Cl_2 (1 mL) was added to a solution of HCl in MeOH (prepared by adding 100 μL of AcCl to 5 mL of anhydrous MeOH) and the mixture was heated to 55 °C for 2 h. The solvent was evaporated to provide the aminodiol **16** as its HCl salt (3.3 mg), which was used without further purification. **16·HCl**: ^{1}H NMR (500 MHz, CD_3OD) δ 3.79 (br ddd, J = 2, 2, 2 Hz, 1H), 3.5 (t, J = 7.3 Hz, 2H), 3.40 (t, J = 6.4 Hz, 2H), 3.20 (ddd, J = 8, 4, 4 Hz, 1H), 3.08 (br m, 1H), 3.00 (t, J = 7.6 Hz, 2H), 2.42 (t, J = 8.0 Hz, 2H), 2.08 (pent, J = 7.8 Hz, 2H), 2.10−1.10 (m, ca. 29H), 0.95 (d, J = 6.5 Hz, 3H), 0.91 (d, J = 6.4 Hz, 3H), 0.88 (d, J = 6.2 Hz, 3H), 0.86 (s, 3H), 0.70 (s, 3H). ^{13}C NMR (75 MHz, CD_3OD) δ 179.1, 78.3, 68.4, 58.9, 57.7, 51.8, 46.9, 43.8, 43.2, 41.1(2 × overlapped), 40.6, 38.6, 37.7, 37.6, 36.9, 34.7, 33.7, 32.3, 31.9, 31.8, 29.4, 26.2, 25.5, 24.6, 22.2, 19.7, 19.5, 19.0, 17.7, 12.4, and 11.6. ESI-MS calcd for $C_{34}H_{61}N_2O_3$ [M + H]$^+$ 545.4677, found 545.4667. TLC R_f 0.65 (80:18:2, $CHCl_3$:MeOH:NH_4OH).

(3β,5α,7α,24*S*)-1-[3-[[7,24-Bis(sulfooxy)cholestan-3-yl]amino]-propyl]-2-pyrrolidinone, Disodium Salt [24-*epi*-Petromyzonamine Disulfate (24-*epi*-PADS, 9)]. The aminodiol **16** (3.3 mg, 0.0057 mmol) was dissolved in pyridine (0.5 mL) and the solution was treated with SO_3·py (45%, 20.5 mg, 0.058 mmol) and heated to 50 °C. After 2 h the solution was cooled to room temperature, MeOH was added, and the solvent was removed under reduced pressure. The residue was purified by silica gel chromatography (12:4:1; CH_2Cl_2:MeOH:NH_4OH v/v/v) to give 24-*epi*-PADS (**9**) as a white solid (3.0 mg, 0.0042 mmol, 82%, from the 7-MOM amine **15**). ^{1}H NMR (500 MHz, CD_3OD; assignments (see the Supporting Information) are supported by COSY and HMQC) δ 4.45 (br m, 1H), 4.12, (ddd, 1H, J = 5.9, 5.9, 4.1 Hz), 3.50 (t, 2H, J = 7.0 Hz), 3.41 (ddd, 1H, J = 14.2, 7.7, 7.7 Hz), 3.39 (ddd, 1H, J = 14.2, 7.4, 7.4 Hz), 3.11 (ddd, 1H, J = 11.8, 11.8, 4.3, 4.3 Hz), 3.01 (dt, 1H, J = 12.5, 5.2 Hz), 2.99 (dt, 1H, J = 12.6, 5.3 Hz), 2.16 (ddd, 1H, J = 14.0, 3.0, 3.0 Hz), 2.09 (pent, 2H, J = 7.6 Hz), 1.93 (pent, 2H, J = 7.0 Hz, H2′), 2.1−0.98 (m, 31H), 0.96 (d, 3H, J = 6.7 Hz), 0.95 (d, 3H, J = 6.5 Hz), 0.91 (d, J = 6.9 Hz, 3H), 0.85 (s, 3H), and 0.69 (s, 3H) (same data as listed in Table S3, Supporting Information). ^{13}C NMR (75 MHz, CD_3OD) δ 178.9, 85.9, 78.1, 58.8, 57.4, 51.3, *49.0*, 47.5, 43.9, 43.2, 41.1, 40.8, 40.7, 38.7, 37.7, 37.2, 36.7, 34.0, 32.7, 31.9 (2× overlapped), 31.7, 29.3, 28.3, 26.3, 25.5, 24.5, 22.2, 19.4, 19.02, 18.97, 17.6, 12.4, and 11.9 (same data as listed in Table S3, Supporting Information). ESI-MS calcd for $C_{34}H_{59}N_2Na_2O_9S_2$ (M − H$^+$ + 2Na$^+$) 749.3452, found 749.3443. IR (KBr) 3440 (br), 3097, 2967, 2933, 2870, 1655, 1451, 1406, 1242, 1205, and 1056 cm^{-1}. TLC (6:3:1, CH_2Cl_2:MeOH:NH_4OH) R_f 0.35. $[\alpha]^{RT}_D$ +1.3 (*c* 0.15).

Acknowledgment. Some NMR instrumentation was acquired with funds from the National Science Foundation (BIR-961477), the University of Minnesota Medical School, and the Minnesota Medical Foundation. Access to some MS instrumentation was provided through the courtesy of the University of Minnesota Mass Spectrometry Consortium for the Life Sciences. Helpful advice with various aspects of mass spectrometric studies was provided by Katherine Fuller, LeeAnn Higgins, Thomas P. Krick, Suddha Marimanikkuppam, and Dana R. Reed. A comparison sample of squalamine was kindly provided by Professor Michael Zasloff. This research was supported by the Great Lakes Fishery Commission (P.W.S. and T.R.H.), the Minnesota Agricultural Experiment Station (P.W.S.), the National Institutes of Health (T.R.H., GM65597 and CA76497), the W. M. Keck Foundation (D.C.M.), and the Mayo Clinic College of Medicine (DCM).

Supporting Information Available: Experimental procedures for preparation of **11**, *cis*-**11**, and **12**; characterization data for all additional new compounds; copies of NMR spectra of **1** (^{1}H), **2** (^{1}H), **3** (^{1}H/^{13}C), **4** (^{1}H), **6** (^{1}H), **9** (^{1}H/^{13}C), **11** (^{1}H/^{13}C), *cis*-**11** (^{1}H), and **12** (^{1}H/^{13}C); details of Mosher ester analyses (and copies of their ^{1}H NMR spectra). This material is available free of charge via the Internet at http://pubs.acs.org.

JO070957L