Pachymoside A — A novel glycolipid isolated from the marine sponge *Pachymatisma johnstonia*[1]

Kaoru Warabi, William T. Zimmerman, Jingkai Shen, Annick Gauthier, Marilyn Robertson, B. Brett Finlay, Rob van Soest, and Raymond J. Andersen

Abstract: Crude extracts of the North Sea marine sponge *Pachymatisma johnstonia* showed promising activity in a new assay for inhibitors of bacterial type III secretion. Bioassay-guided fractionation resulted in the isolation of the pachymosides, a new family of sponge glycolipids. A major part of the structural diversity in this family of glycolipids involves increasing degrees of acetylation and differing positions of acetylation on a common pachymoside glycolipid template. All of the metabolites with these variations in acetylation pattern were converted into the same peracetylpachymoside methyl ester (**2**) for purification and spectroscopic analysis. Pachymoside A (**1**) is the component of the mixture that has natural acetylation at the eight galactose hydroxyls and at the C-6 hydroxyls of glucose-B and glucose-D. Chemical degradation and transformation in conjunction with extensive analysis of 800 MHz NMR data was used to elucidate the structure of pachymoside A (**1**).

Key words: Pachymatisma johnstonia, marine sponge, pachymoside, glycolipid.

Résumé : Les produits bruts extraits de l'éponge marine de la mer du Nord *Pachymatisma Johnstonia* présentent une activité intéressante dans un nouvel essai pour les inhibiteurs de la sécrétion bactérienne de type III. Des fractionnements orientés par des bioessais ont permis d'isoler les pachymosides, une nouvelle famille de glycolipides des éponges. Une partie importante de la diversité structurelle de cette famille de glycolipides implique des degrés croissants d'acétylation et des différences dans les positions de l'acétylation sur un gabarit commun du glycolipide pachymoside. Tous les métabolites comportant ces variations dans le patron d'acétylation ont été transformés dans le même ester méthylique du peracétylpachymoside (**2**) aux fins de purification et d'analyse spectroscopique. Le pachymoside A (**1**) est le composant du mélange qui correspond à l'acétylation naturelle au niveau des huit hydroxyles du galactose et aux hydroxyles en C-6 du glucose-B et du glucose-D. Afin d'élucider la structure du pachymoside A (**1**), on a utilisé une dégradation et une transformation chimique de concert avec une analyse extensive des données RMN à 800 MHz.

Mots clés : Pachymatisma johnstonia, éponge marine, pachymoside, glycolipide.

[Traduit par la Rédaction]

Introduction

Several important gram-negative pathogenic bacteria deliver virulence factors directly into host cells via a "type III secretion system" (TTSS) (1–4). These include *Bordetella pertussis*, the cause of whooping cough, *Chlamydia trachomatis*, the cause of a sexually transmitted disease that leads to blindness, enterohemorrhagic *Escherichia coli* (EHEC), the cause of "hamburger" disease, and *Salmonella typhi*, the cause of typhoid fever. EHEC (5) and the closely related enteropathogenic *E. coli* (EPEC) (6), the major cause of infantile diaherria that kills several thousand children worldwide each year, have served as important models for studying TTSSs and their role in pathogenicity. While many bacterial pathogens attach themselves to host cell proteins, these pathogenic *E. coli* are remarkable because they use a TTSS to deliver their own receptors into the host cells to which they attach. The bacterial protein receptor delivered into the host cell by the *E. coli* TTSS is called Tir for "translocated intimin receptor". Once inside the host cell, Tir binds to an *E. coli* protein "intimin" on the surface of the bacterium to give intimate attachment, a required prelude to

Received 23 June 2003. Published on the NRC Research Press Web site at http://canjchem.nrc.ca on 15 December 2003.

Dedicated to Professor Edward Piers, a wonderful colleague and friend.

K. Warabi, W.T. Zimmerman,[2] J. Shen, and R.J. Andersen.[3] Departments of Chemistry, Earth and Ocean Sciences, University of British Columbia, 2036 Main Mall, Vancouver, BC V6T 1Z1, Canada.
A. Gauthier, M. Robertson, and B.B. Finlay. Departments of Biochemistry and Microbiology, University of British Columbia, Vancouver, BC V6T 1Z3, Canada.
R. van Soest. Department of Coelenterates and Porifera, Zoologisch Museum, University of Amsterdam, Amsterdam, The Netherlands.

[1]This article is part of a Special Issue dedicated to Professor Ed Piers.
[2]On leave from DuPont Crop Protection, Stine-Haskell Research Center, Newark, DE, U.S.A.
[3]Corresponding author (email: randersn@interchange.ubc.ca).

 doi: 10.1139/V03-183

virulence. The TTSS, which forms a syringe-like pore that spans both bacterial and host cell membranes, consists of more than 20 proteins. Deleting or mutating any part of the TTSS significantly reduces the virulence of the affected *E. coli*. Non-pathogenic intestinal *E. coli* do not have TTSSs.

The TTSS is considered an attractive target for the development of new antibiotics that would selectively target pathogenic gram negative bacteria but would not harm the ordinary microflora that do not possess this virulence mechanism (1, 7–9). An added benefit of this type of hypothetical antibacterial agent is that by targeting a virulence mechanism, there would be little or no selective pressure for viability, potentially reducing the development of resistance. No TTSS inhibitors were known at the outset of our research. Consequently, there was a significant need to discover TTSS inhibitors that could be used to provide proof principle demonstrations and drug leads for this novel approach to treating important infectious diseases.

Marine sponges, which are the richest source of bioactive marine natural products (10, 11), are constantly exposed to bacteria in seawater and it has frequently been proposed that they contain chemical agents that have evolved to prevent pathogenic bacteria from infecting them (12, 13). We have screened a library of crude extracts from marine sponges for their ability to inhibit type III secretion of *E. coli* secreted proteins (Esps) by EPEC without affecting the growth or general secretion of the bacterium (1). Two extracts showed promising activity in the screen. Recently, we reported the structure of caminoside A, an antimicrobial glycolipid isolated from the Caribbean sponge *Caminus sphaeroconia* that was the first compound active in the TTSS inhibition screen (14, 15).[4] The second active extract came from the sponge *Pachymatisma johnstonia* (Bowerbank, 1842) collected along the coast of the Isle of Mann. Bioassay-guided fractionation of the *P. johnstonia* extract yielded a complex family of glycolipids that were active in the TTSS assay. The structure elucidation of pachymoside A (**1**), a novel glycolipid that is a representative member of this family, is described below.

Results and discussion

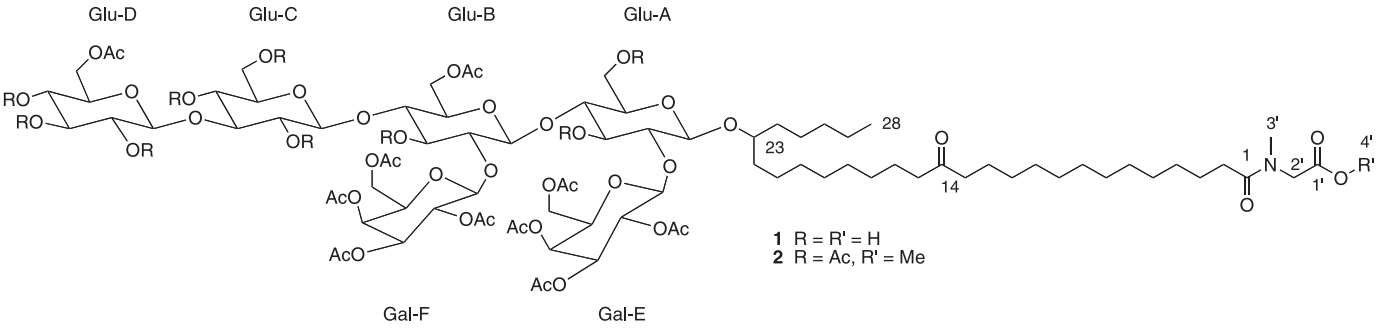
The TTSS inhibitory fraction obtained from an initial Sephadex LH20 chromatographic separation (eluent: MeOH) of the crude *P. johnstonia* extract gave spectroscopic data that was indicative of a complex mixture of closely related glycolipids. LR-FAB-MS analysis of the mixture revealed a series of clusters of molecular ions with centres at *m/z* 1918, 1932, 1944, 1961, 1973, 1988, 2003, 2015, and 2046. The ^{1}H and ^{13}C NMR spectra of the mixture showed evidence for the presence of a number of monosaccharide anomeric carbons, a large number of carbinol methines and methylenes, extensive acetylation, and a long aliphatic chain. All chromatographic attempts to resolve the complex mixture of natural products into individual components failed. To facilitate purification of the mixture, it was first methylated with (trimethylsilyl)diazomethane and then acetylated with Ac_2O–pyridine in the presence of DMAP. Reversed-phase HPLC separation of the mixture of derivatized natural products gave a pure peracetylpachymoside A methyl ester (**2**) as a colourless glass.

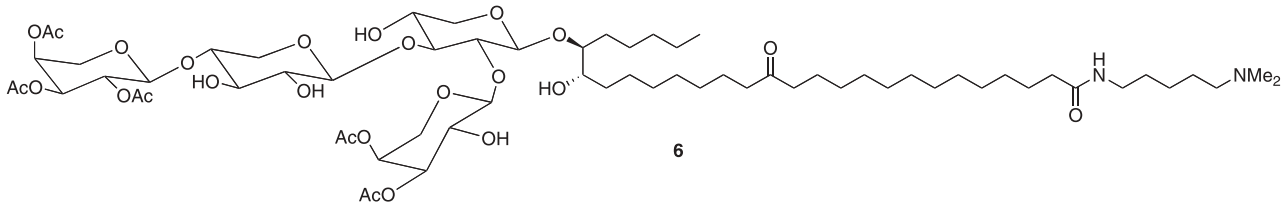
Peracetylpachymoside A methyl ester (**2**) gave a $[M + Na]^+$ ion at *m/z* 2333.9639 in the ESI-HR-MS consistent with a molecular formula of $C_{106}H_{159}NO_{54}$ (calcd. for $^{12}C_{105}{}^{13}CH_{159}NO_{54}Na$, 2333.9658), requiring 28 sites of unsaturation. The ^{13}C NMR spectrum of **2** (100 MHz) recorded in C_6D_6 showed the same general features of a glycolipid observed in the spectrum of the natural product mixture. Thus, it was possible to attribute clusters of peaks to ester or amide carbonyls (δ 168–170), monosaccharide anomeric carbons (δ 100–102), carbinol methines and methylenes (δ 67–81), methyl groups attached to O or N (δ 49–52), acetate methyl carbons (δ 29–31), and a significant saturated aliphatic fragment (δ 14–43). The ^{1}H NMR spectrum of **2** was very congested at 500 MHz so all of the ^{1}H 1D and 2D data used in the structure elucidation was recorded at 800 MHz in C_6D_6, which gave excellent dispersion.

The first step in the structure elucidation of **2** involved chemical degradation of the mixture of natural pachymosides to liberate the individual monosaccharides. Treatment of the pachymoside mixture with acetyl chloride in dry MeOH cleaved the glycosidic linkages and converted the individual monosaccharides into mixtures of α- and β-methylglycopyranosides. After removal of the reagents in vacuo, the resulting residue was acetylated with Ac_2O in pyridine and a catalytic amount of DMAP to give the peracetylated methylglycosides of the individual monosaccharides. Fractionation of the peracetylated methylglycosides via reversed-phase HPLC gave an inseparable mixture of methyl 2,3,4,6-tetra-*O*-acetyl-α-glucopyranoside (**3**) and methyl 2,3,4,6-tetra-*O*-acetyl-α-galactopyranoside (**4**) in a ratio of ≈2:1, respectively. The difference in intensities of the resonances for each of **3** and **4** made it possible to identify each of the component compounds by analysis of the 1D and 2D NMR data obtained for the mixture. To confirm the assigned structures of **3** and **4**, authentic samples of D-glucose and D-galactose were converted into their peractylated methyl α-glycopyranosides and the NMR chemical shifts of the authentic compounds were compared with the degradation products. Chiral GC analysis of the alditol acetates prepared from authentic L- and D-glucose and D-galactose and the alditol acetates prepared from the monosaccharides obtained by aqueous hydrolysis of the pachymoside mixture showed that the glucose and galactose residues in the pachymosides had the D configuration.

A very non-polar fraction, which contained a mixture of the aglycon portions of the pachymosides, was also obtained

[4] While the current study was underway, another group reported screening a chemical library for inhibition of YopE transcription with a luciferase reporter assay, yielding three compounds that inhibited the reporter signal expressed from the *yopE* promoter and effector protein secretion with little effect on bacterial growth (15). One compound appears to be affecting type III regulation, and this and one other compound are specific to the virulence-associated TTSS, and do not affect flagellum assembly and motility. Overall, these results are evidence in support of TTSS-specific inhibitors, although the bacterial target(s) and ability to inhibit TTSS in infected host cells awaits further study.





from the methanolysis reaction on the pachymoside mixture described above. HPLC fractionation of the mixture of aglycons yielded a number of closely related compounds. Methanolysis of pure peracetylpachymoside A methyl ester (**2**), followed by analysis of the non-polar product by HPLC, ^{1}H NMR, and EI-MS, revealed that the methyl ester **5** obtained from the large-scale methanolysis of the mixture of pachymosides corresponded to the aglycon fragment of **2**. The ester **5** gave a M^+ ion at *m/z* 539.4567 in the EI-HR-MS appropriate for a molecular formula of $C_{32}H_{61}NO_5$ (calcd. 539.4550), requiring three sites of unsaturation.

Three associated pairs of singlets were apparent in the ^{1}H NMR spectrum of **5**, each of which exhibited an intensity ratio of ca. 4:1. A singlet at δ 3.71 in the ^{1}H NMR spectrum (associated peak at δ 3.77), which showed a HMQC correlation to a carbon at δ 52.5 and a HMBC correlation to a carbon at δ 170.0, was assigned to a methyl ester. A second singlet at δ 3.05 (associated peak at δ 2.95), which showed an HMQC correlation to a carbon at δ 36.5 and HMBC correlations to carbons at δ 172.0 and 49.0, was assigned to a *N*-methyl amide. The third singlet at δ 4.11 (associated peak at δ 4.03), integrated for only two protons relative to the other singlets. It showed a HMQC correlation to a carbon at δ 49.0 and HMBC correlations to the ester carbonyl at δ 170.0, the amide carbonyl at δ 172.0, and the *N*-methyl carbon at δ 36.5, and was assigned to the α carbon of an *N*-acyl-*N*-methylglycine methyl ester fragment. As is typical for tertiary amides, slow rotation about the *N*-acyl-*N*-methylglycine amide bond is presumed to be responsible for the presence of associated pairs of peaks observed in the ^{1}H NMR spectrum for the *N*-methyl, methyl ester, and glycine α-methylene resonances.

Additional features in the ^{1}H NMR spectrum of **5** included a six proton multiplet at δ 2.36, a deshielded one proton multiplet at δ 3.55, and a three proton triplet at δ 0.87. The six proton multiplet (δ 2.36) showed HMQC correlations into carbon resonances at δ 38.1 and 42.8, and HMBC correlations into the amide carbonyl resonance at δ 172.0 and into a deshielded carbon resonance at δ 211.0, assigned to a saturated ketone. These correlations indicated that the six proton multiplet was comprised of three separate but coincidentally chemical shift-equivalent methylene resonances that were assigned to a methylene α to the amide carbonyl and two methylenes α to a saturated ketone. Consistent with these assignments, there was a small triplet at δ 2.20 in the ^{1}H NMR spectrum that was assigned to the methylene α to the amide carbonyl in the minor rotamer of **5**. The methine resonance at δ 3.55 showed a HMQC correlation to the carbon resonance at δ 72.5 consistent with the presence of a secondary alcohol in **5**, and the triplet at δ 0.87 was assigned to a methyl at the end of a linear aliphatic chain. Since the methyl ester, *N*-methylamide, and ketone carbonyls accounted for the three sites of unsaturation in **5**, and the only terminal residues identified in the NMR data were the methyl ester and aliphatic methyl groups, the aglycon had to have a 27-carbon linear chain extending from the *N*-acyl amide carbonyl to the terminal methyl group, and there had to be saturated ketone and secondary alcohol functionalities situated along the chain.

The positions of the ketone and secondary alcohol functionalities in **5** were determined by analysis of the EI-HR-MS data obtained for the compound as shown in Fig. 1. A major peak at *m/z* 145.0739 ($C_6H_{11}NO_3$) was attributed to a McLafferty rearrangement involving the C-4 amide carbonyl that results in cleavage of the C-5—C-6 bond, confirming the presence of the *N*-acyl-*N*-methylglycine methyl ester fragment at one terminus of the aglycon. A pair of peaks at *m/z* 284.2232 ($C_{16}H_{30}NO_3$) and 341.2557 ($C_{19}H_{35}NO_4$) were attributed to McLafferty rearrangements occurring on both sides of the ketone functionality, which located the ketone at C-14. Finally, a peak at *m/z* 468.3697 ($C_{27}H_{50}NO_5$) was assigned to an α cleavage adjacent to the secondary alcohol, which had to be at C-23. The base peak in the mass spectrum of **5** resulted from a cleavage of the amide bond to

Fig. 1. HR-EI-MS analysis of the aglycon methyl ester **5**.

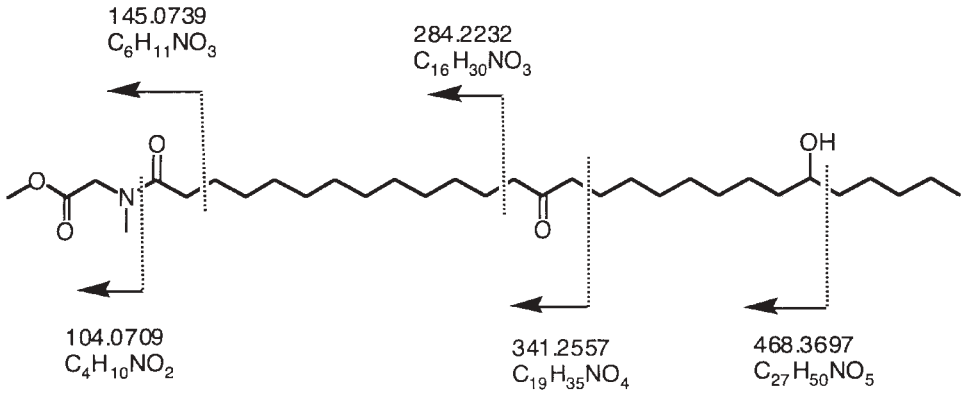
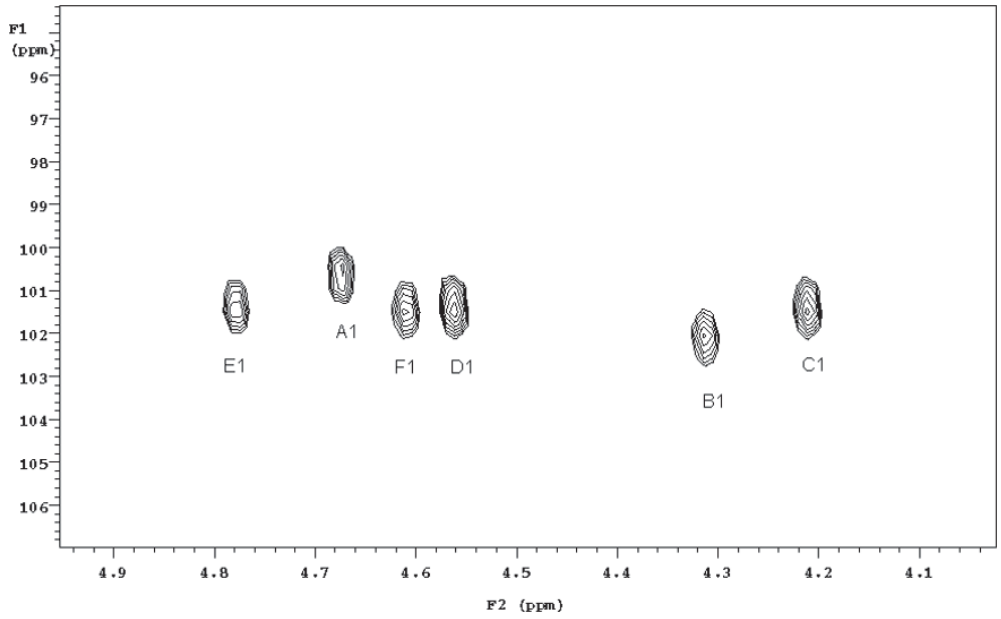


Fig. 2. Expansion of gHSQC spectrum of peracetylpachymoside A methylester (**2**) acquired in C_6D_6 at 800 MHz. Anomeric carbon correlations are shown.

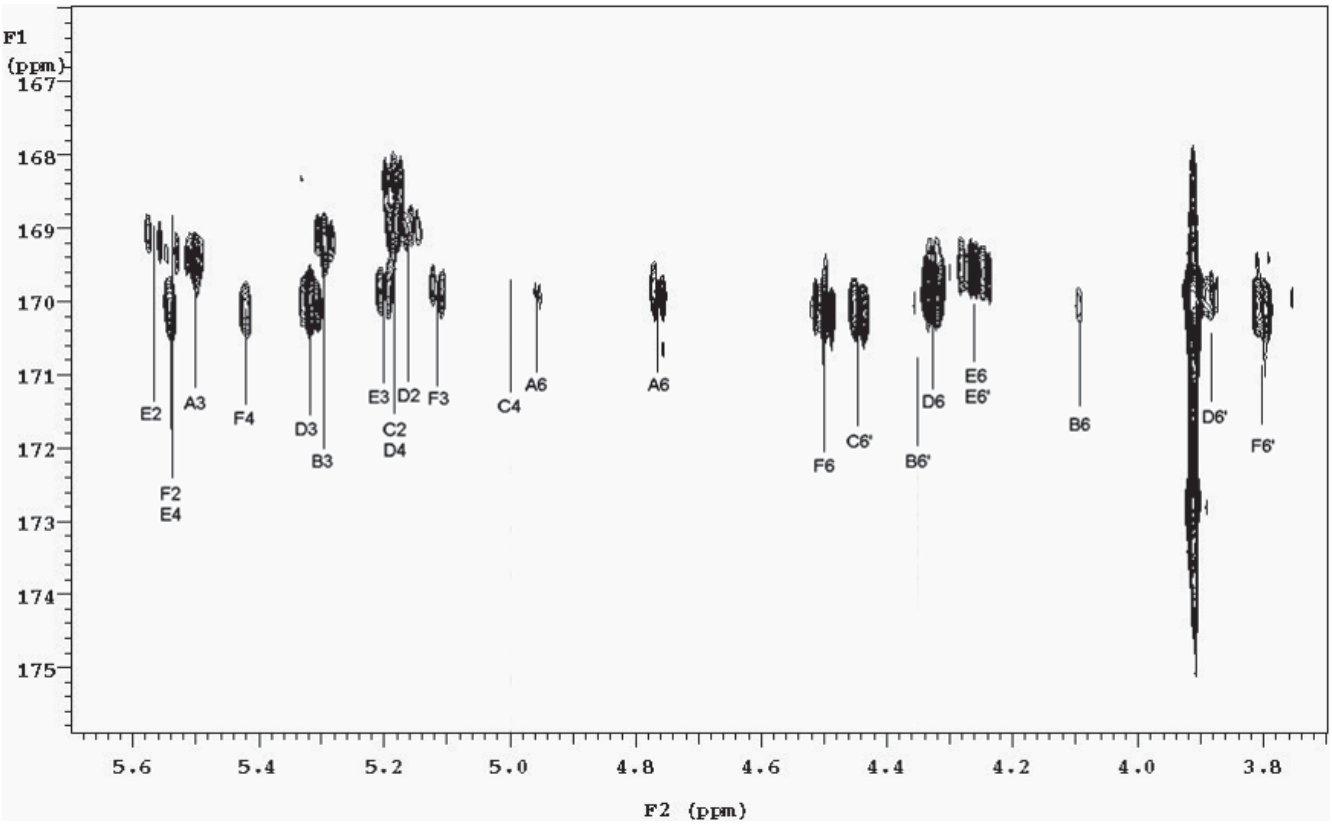


give a fragment ion at *m/z* 104.0709 ($C_4H_{10}NO_2$), which had the molecular formula of protonated *N*-methylglycine methyl ester.

Subtracting the atoms present in the aglycon ($C_{32}H_{60}NO_5$) from the molecular formula of peracetylpachymoside A methyl ester (**2**) ($C_{106}H_{159}NO_{54}$) left $C_{74}H_{99}O_{49}$ and 25 sites of unsaturation unaccounted for. Preliminary analysis of the NMR data obtained for **2**, in conjunction with the degradation studies, indicated the remaining atoms had to be part of glucose, galactose, or acetate residues. Six anomeric carbons and their attached protons (δ ^{13}C: ^{1}H; 100.7:4.67; 101.4:4.21; 101.4:4.56; 101.4:4.78; 101.5:4.61; 102.1:4.31) could be readily identified in the HSQC spectrum of **2** (Fig. 2). Subtracting the 36 carbons of six hexose monosaccharides (glucose and galactose) from the 74 carbons remaining to be accounted for, left 38 carbons, which corresponded to 19 acetate residues. The six sites of unsaturation accounted for by the putative monosaccharide rings, along with the 19 sites of unsaturation associated with the acetate residues and the three sites of unsaturation in the aglycon, accounted for all of the unsaturation required by the molecular formula of **2**. Therefore, what remained was to determine the connectivity among the aglycon, the six monosaccharide residues, and the 19 acetates, as well as determining the relative stereochemistry of the polysaccharide portion of the molecule. These features of **2** were elucidated by detailed analysis of its 800 MHz NMR data (Figs. 3 and 4).

Starting with the six well-resolved anomeric proton resonances, it was possible by careful analysis of the COSY, 1D-TOCSY, 2D-TOCSY, HSQC, and HMBC data to assign resonances to all of the hydrogen and carbon atoms in each of the six monosaccharide residues as listed in Table 1. Degradation had shown that the pachymosides contained only D-glucose and D-galactose monomers. Analysis of the ^{1}H vicinal NMR coupling constants for each of the monosaccharide residues A to F confirmed that there were four glucose units (A–D) and two galactose units (F and E) in **2**. In addition, all six of the anomeric proton resonances appeared as doublets with coupling constants between 7.0 and 8.2 Hz, indi-

Fig. 3. Expansion of gHMBC of peracetylpachymoside A methyl ester (**2**), (natural abundance) acquired at 800 MHz in C_6D_6. Assignments for carbinol methine and methylene HMBC correlations to acetate carbonyls are indicated. No coupling for C-4 to its attached acetate carbonyl is observed for the unlabeled sample.



cating that all of the glycosidic linkages were β. Examination of the proton chemical shifts (Table 1) in each of the monosaccharide residues for acyl shifts predicted that glucose-A was acetylated at C-3 and C-6, glucose-B was acetylated at C-3 and C-6, glucose-C was acetylated at C-2, C-4, and C-6, glucose-D was acetylated at C-2, C-3, C-4, and C-6, galactose-E was acetylated at C-2, C-3, C-4, and C-6, and galactose-F was also acetylated at C-2, C-3, C-4, and C-6. This acetylation pattern, which was confirmed by HMBC correlations from the corresponding monosaccharide protons to the acetate carbonyls (Fig. 3), was consistent with the presence of 19 acetate residues in **2**. Finally, HMBC and ROESY correlations (Table 1 and Fig. 4) provided the information required to identify the linkages between the six monosaccharide residues A–F and the aglycon as shown in Fig. 4.

NMR analysis of the crude mixture of pachymosides showed that the natural products are partially acetylated. Information about the natural acetylation pattern was lost during the peracetylation step required for isolation of pure peracetylpachymoside A methyl ester (**2**). To locate the sites of introduced acetylation in **2**, a second portion of the crude pachymoside mixture was methylated with (trimethylsilyl)diazomethane and then acetylated with $(CH_3{}^{13}CO)_2O$, pyridine, and DMAP as before. The crude reaction mixture was fractionated to give peracetylpachymoside A methyl ester (**2**) labelled at the sites of synthetic acetylation with $CH_3{}^{13}CO$. The ESI-LR-MS of the unlabeled peracetylpachymoside A methyl ester (**2**) gave a $[M + Na]^+$ at *m*/*z* 2335, while the corresponding ESI-LR-MS of the ^{13}C-acetate labelled compound gave two $[M + Na]^+$ major peaks at *m*/*z* 2343/2344, indicating the addition of eight and nine ^{13}C-acetyl groups. An HMBC spectrum collected on the ^{13}C-labelled sample of **2** showed intense peaks for the monosaccharide proton to acetyl carbonyl correlations involving ^{13}C-labelled acetate residues and essentially no correlations for the natural abundance acetates (Fig. 5). Using this intensity difference as a marker of introduced acetate, it was possible to show that ^{13}C-acetate had been introduced at 10 sites in the molecule (A-3, A-6, B-3, B-6, C-2, C-6, D-2, D-3, D-4, D-6; see Fig. 5). Evidence for labelled acetate introduction at C-4 came from the observation of additional scalar coupling in the glucose-C H-4 resonance observed at 400 MHz, which was attributed to three bond ^{1}H/^{13}C coupling. The absence of a corresponding HMBC cross peak in the 800 MHz spectrum is an anomaly produced by the line broadening of the C residue H-4 resonance at high field (vide infra). There was no evidence for introduction of ^{13}C-labelled acetate at any of the hydroxyl positions of the galactose-E and galactose-F residues, indicating that they are fully acetylated in all of the natural pachymosides. Since there is evidence for introduction of labelled acetate at 11 hydroxyls in **2** and the MS data indicates that only eight or nine labelled acetates have been introduced in individual molecules, the natural pachymoside mixture must consist in part of a major group of compounds that have the same aglycon **5**, are peracetylated on the galactose residues E and F, and have either two or three acetates distributed among

Table 1. 1D- and 2D-NMR data for pachymoside A peracetate methyl ester (**2**).[a]

Position	δ ^{13}C NMR	δ ^{1}H NMR	Multiplicity	*J* (Hz)	COSY (H → H)	HMBC (H → C)	TOCSY (H → H)	ROESY (H → H)
Aglycon[b]								
1	172.8	—						
2	33.0/32.8	2.04/1.98	t	7.5		1	3,4	
3	25.3	1.72/1.74			2, 4	2, 4	2,4	
4	29.7	1.30			3	3	2,3	
5–10		1.26–1.32						
11	29.7	1.23			12	12, 13	12,13	
12	24.2	1.58	qn	7.5	11,13	13, 14	11,13	
13	42.7	2.09	t	7.5	12	12, 14	11,12	
14	208.9	—						
15	42.7	2.12	t	7.5	16	14, 16	16,17,18,19,20,21	
16	24.2	1.64	qn	7.5	15, 17	14, 15	15,17,18,19,20,20′, 21	
17	30	1.33			16	15, 16	15,16,18	
18	30	1.44				17/19(vw)	15,16,17	
19	30	1.32				16,17,20,21	18,20	
20	30	1.44			20′,21	18/19(vw)	20′,21,21′	
20′	—	1.38			20	18/19	19,20,21,22,23,24	
21	25.3	1.55			20	20(w)	20,22	
22	34.5	1.70			22′	21,23	20,20′,21,22′,23	
22′	—	1.57			22	21,23	20,22	
23	80.2	3.71			22,22′, 24,24′	22, 24, A1	20-22/20′-22′,24-26/24′26′	A1,22,22′,24,24′
24	35.3	1.70			24′	23,25	23,25,25′,26	
24′	—	1.63			24	25(w)	24,25,26	
25	25.3	1.52					24,24′(w),26,26′	
25′	—	1.54				27(vw)	24,24′,26	
26	32.4	1.39			28	25, 27	23,24,24′,25,25′	
26′	—	1.34				27	25,25′,26,27	
27	23.1	1.40			28	26, 28	28,26,25,24,23	
28	14.4	0.97	t	7.4	27	26, 27	27,26,25,24,23	
1′	169.9	—						
2′	49.2/51.0	3.91/3.42	s			1, 1′		
3′	35.7/34.5	2.44/2.79	s			1, 2′		
4′	51.3/*	3.27/3.24	s			1′		
Glu-A[c]								
A-1	100.7	4.67	d	7.2	A2	23, A5		23
A-2	78.7	3.79	dd	7.2, 9.0	A1, A3	A1, A3, E1		E1
A-3	75.1	5.50	t	9.0	A2, A4	A2, A4		
A-4	78.8	3.73	dd	9.0, 9.6	A3, A5	A3, A6, B1		B1
A-5	72.4	3.82	ddd	2.6,4.5,9.6	A4, A6, A6′	A4		
A-6	62.2	4.76	dd	4.3, 12.0	A5, A6′	A4		
A-6′	—	4.95	dd	2.8, 12.0	A5, A6′	A4, A5		
Glu-B								
B-1	102.1	4.31	d	7.0	B2	A4		A4
B-2	78.1	3.62	dd	7.0, 8.6	B1, B3	B1, B3, F1		F1
B-3	74.0	5.29	dd	8.6, 9.6	B2, B4	B2, B4		
B-4	76.6	3.50	dd	9.6, 11.2	B3, B5	B3, B5, C1		
B-5	72.2	3.08	ddd	1.6,5.3,11.3	B4, B6, B6′	B1		B1
B-6	62.8	4.10	dd	5.2, 11.9	B5,B6′	B5		
B-6′	—	4.35	dd	1.9, 11.9	B5, B6	B4, B5		
Glu-C								
C-1	101.4	4.21	d	8.2	C2	B4	C(2,3,4,5)[d]	B4, C3, C6(w)[e]
C-2	73.2	5.18	dd	8.2, 9.4	C1, C3	C1, C3	C(1,3,4,5)	C4(vw)
C-3	79.1	3.89	nco[f]		C2, C4	(nco)	C1, C2	D1, C1(w)
C-4	68.4	5.00	t[f]	9.6	C3, C5	(nco)	C1	C6′(w)

Table 1 (*concluded*).

Position	δ ^{13}C NMR	δ ^{1}H NMR	Multiplicity	*J* (Hz)	COSY (H → H)	HMBC (H → C)	TOCSY (H → H)	ROESY (H → H)
C-5	72.7	3.51	nco[f]		C4, C6, C6′	C1[g]	C6, C6′	C2, C6
C-6	61.9	4.11	dd	2.5, 12.7	C5, C6′	C4, C5	C5, C6′	C5, C6′
C-6′	—	4.44	dd	4.2, 12.3	C5, C6	C5(w)	C5, C6	C6
Glu-D								
D-1	101.4	4.56	d	8.2	D2	C3		C3
D-2	71.6	5.15	dd	8.2, 9.8	D1, D3	D1, D3		
D-3	73.3	5.32	dd	9.4, 9.8	D2, D4	D2, D4		
D-4	68.1	5.18	dd	9.4, 10.4	D3, D5	D3, D5		
D-5	72.2	3.05	ddd	2.1,3.9,10.2	D4, D6, D6′	D1(w)		
D-6	61.5	4.32	dd	3.9, 12.5	D5, D6′	(nco)		
D-6′	—	3.88	dd	2.1, 12.5	D5, D6	D4, D5		
Gal-E								
E-1	101.4	4.78	d	8.0	E2	A2		A2
E-2	69.7	5.56	dd	8.0, 10.5	E1, E3	E1, E3		
E-3	71.6	5.20	dd	3.4, 10.5	E2, E4	E2, E4		
E-4	67.3	5.54	dd	0.8, 3.4	E3, E5(w)	E2, E3		
E-5	70.9	3.49	ddd	0.8,6.2,7.8	E4(w),E(6/6′)	E4, E6		
E-6	60.9	4.25	dd[h]	6.2, 11.0	E5	E4, E5		
E-6′	—	4.27	dd[h]	7.7, 11.0	E5	E4, E5		
Gal-F								
F-1	101.5	4.61	d	8.0	F2	B2		B2
F-2	68.9	5.54	dd	8.0, 10.6	F1, F3	F1, F3		
F-3	71.4	5.11	dd	3.4, 10.6	F2, F4	F2, F4(w)		
F-4	67.5	5.42	dd	0.8, 3.4	F3, F5(w)	F2, F3		
F-5	71.2	3.33	ddd	0.8,5.3,8.1	F4(w), F6, F6′	F4, F6		
F-6	61.3	4.50	dd	7.7, 11.3	F5, F6′	F4, F5		
F-6′	—	3.80	dd	5.3, 11.3	F5, F6	F4, F5		

Note: (w) denotes a weak but detectable correlation, (vw) very weak; singlet (s), doublet (d), triplet (t), quartet (q), quintet (qn), unresolved multiplet (m); no coupling observed (nco).

[a]Spectra recorded in C_6D_6 at 800 MHz.

[b]Major and (or) minor rotomeric forms indicated where observed.

[c]Acetate resonances not indicated.

[d]1D-TOCSY shows correlations: C1-C2, C3, C4, C5 and C6-C2, C4, C5, C6′.

[e]Possible overlap with B-4 → C1 in HMBC.

[f]Peaks broadened at 800 MHz, C4 appears as a triplet at 400 MHz; (nco)

[g]1D-NOE (400 MHz) shows C1-C3-C5 (negative enhancement).

[h]Non-first-order coupling (see *J* resolved in *Supplementary information*[5]), COSY cross peaks coincident.

the 11 hydroxyl functionalities of the four glucose residues A–D.

The mass of a pachymoside having the aglycon **5** and acetylation at 10 sites is 1918, a mass found in the FAB-MS of the crude mixture of natural pachymosides prior to derivatization. Adding an additional acetate to the sugar alcohols generates a natural pachymoside with a mass of 1960, which was also represented in the FAB-MS of the crude mixture. Therefore, it appears that a major part of the structural diversity in this family of glycolipids involves increasing degrees of acetylation and differing positions of acetylation on the pachymoside template. All of these variations in acetylation pattern would have been converted into the same peracetylpachymoside A methyl ester (**2**) upon derivatization for purification. We have arbitrarily chosen to define pachymoside A (**1**) as the component that has natural acetylation at the eight galactose hydroxyls and at B-6 and D-6, since the HMBC correlation intensities in the ^{13}C-labelled compound indicate that these latter two positions have the greatest degree of natural acetylation.

It was clear from the NMR data obtained for the mixture of natural pachymosides and the aglycons liberated by methanolysis that there are also natural variations in the aglycon. There was evidence for aliphatic methyl doublets in the ^{1}H NMR spectrum of both the mixture of aglycon methyl esters and the natural pachymoside mixture indicating the presence of internal methyl branches on the aglycon fatty acyl chain. The detailed structures of the various aglycons other than **5** have not been determined due to the difficulties of locating branching methyls in a long aliphatic chain.

During the course of analyzing the NMR data for peracetylpachymoside A methyl ester (**2**), an unexpected difference in the data obtained at 400, 500, and 800 MHz in

[5] Supplementary data may be purchased from the Directory of Unpublished Data, Document Delivery, CISTI, National Research Council Canada, Ottawa, ON K1A 0S2, Canada (http://www.nrc.ca/cisti/irm/unpub_e.shtml for information on ordering electronically).

Fig. 4. A minimal set of unambiguous ROESY and HMBC correlations linking the substructures of **2**.

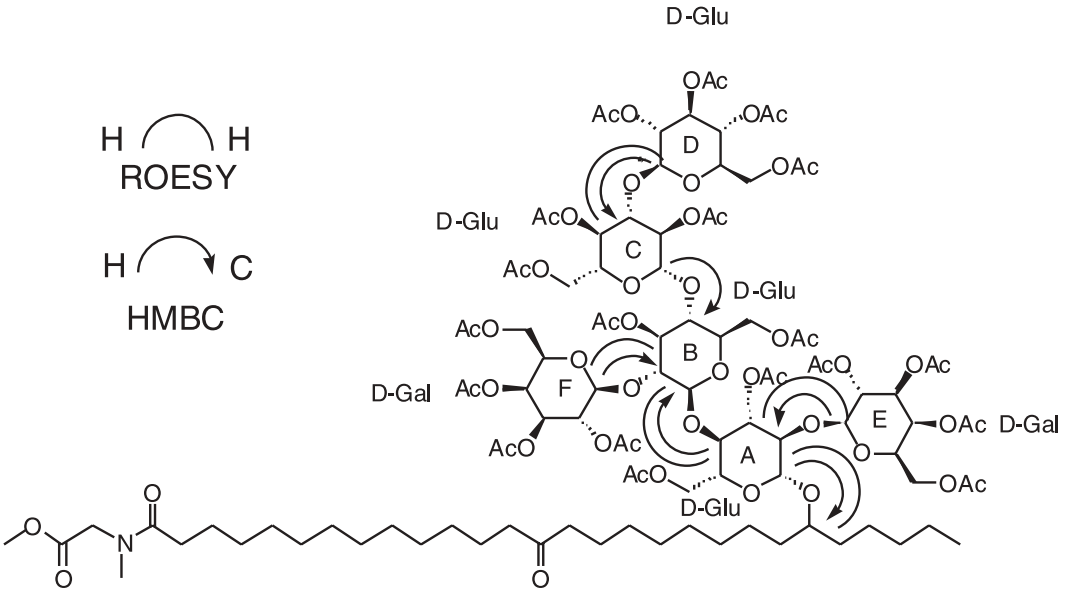
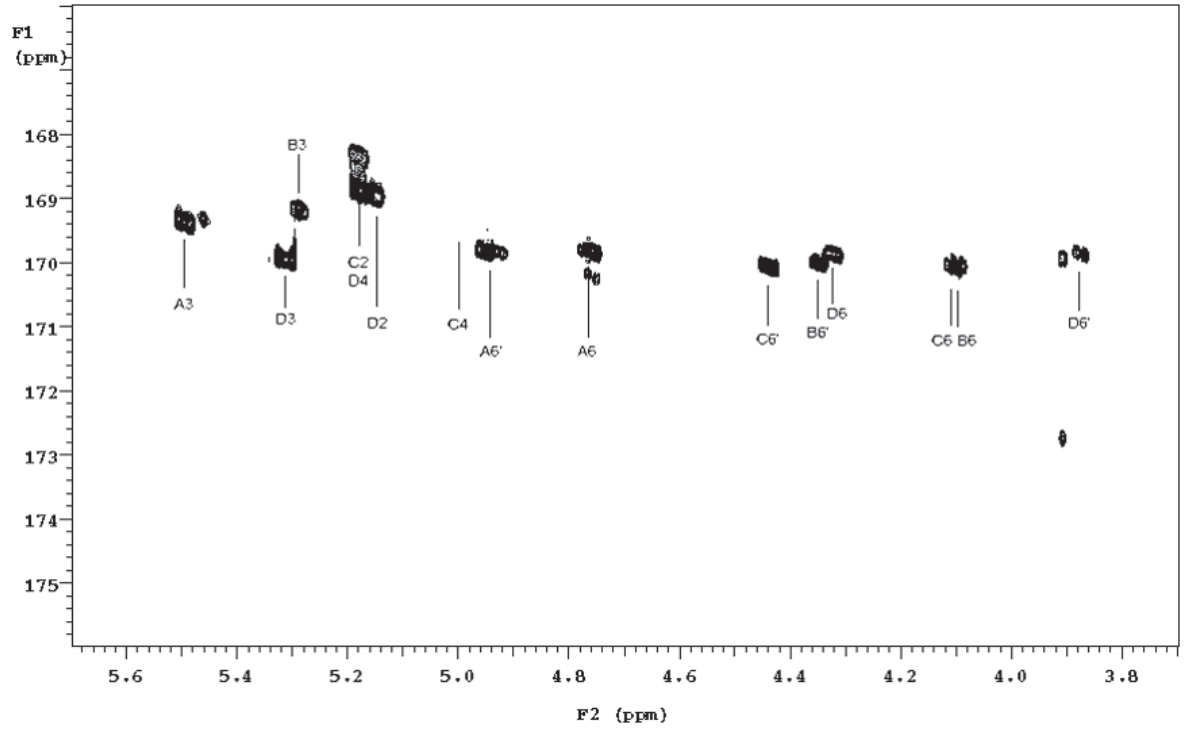
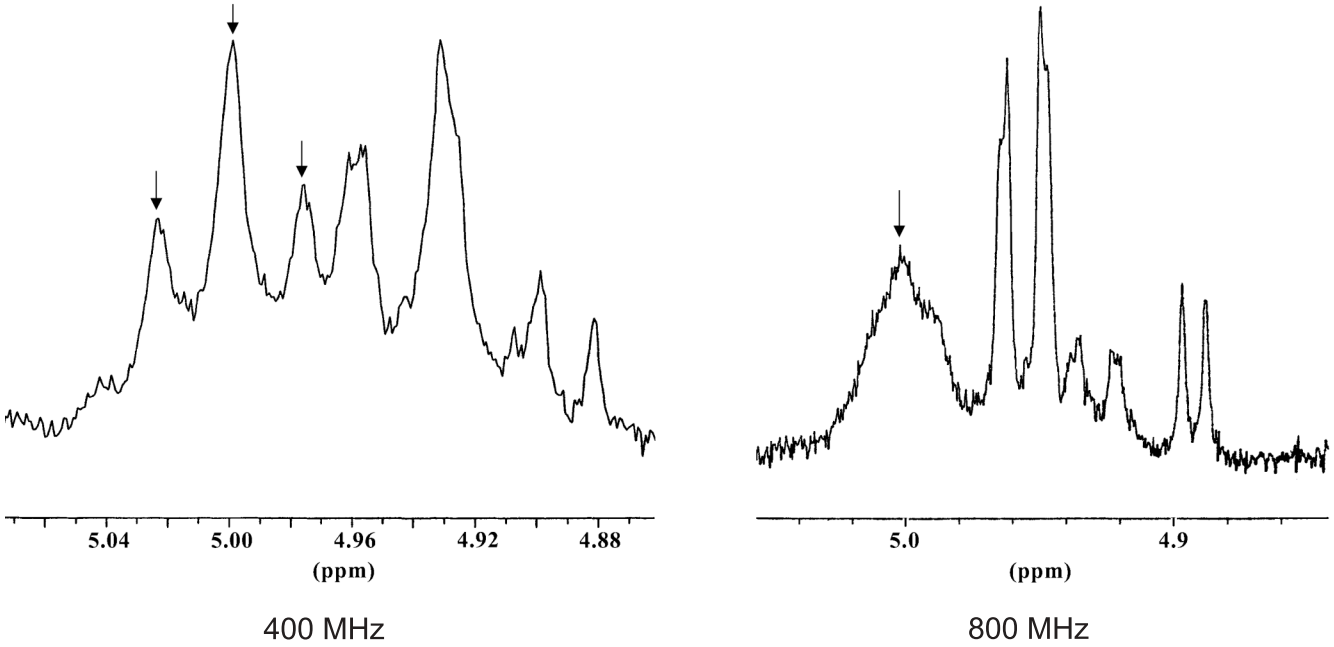


Fig. 5. Expansion of gHMBC spectrum of peracetylpachymoside A methyl ester (**2**) ($CH_3{}^{13}CO$-labeled) acquired at 800 MHz in C_6D_6. Assignments of methines and methylenes coupled to ^{13}C-labeled acetates are indicated. The methine at C-4 exhibits weak coupling at 500 MHz (2D data), while none is observed at 800 MHz. 1D proton data (400 MHz) for this methine exhibits additional splitting (unlabeled appears as triplet, ^{13}C-labeled appears as triplet of doublets). Cross peaks for the methylene to carbonyl coupling are of generally weaker intensity than the methines and vary in order of decreasing intensity as follows: C6 > A6 > B6 > D6.



C_6D_6 was encountered. As shown in Fig. 6, the resonance assigned to the glucose-C H-4 resonance appeared as a clear dd (J = 9.9, 9.6 Hz) at 400 and 500 MHz as expected, but appeared as a broad unresolved multiplet at 800 MHz. Normally the increased dispersion at 800 MHz simplifies multiplets by removing many second-order effects and signal overlap. Therefore, the deterioration of the multiplet structure on going from 400 to 800 MHz was contrary to our usual experience with 800 MHz data. One possible explanation for the observation is restricted conformational mobility at C-4 of glucose-C that is fast on the NMR timescale at 400 MHz leading to a single sharp resonance, but slow on

Fig. 6. NMR spectra (400 and 800 MHz) (C_6D_6) of the glucose-C residue H-4 resonance.



the NMR timescale at 800 MHz leading to a broadening of the resonance. Since the H-1, H-2, and H-6/H-6′ resonances in residue C are sharp at 800 MHz, the conformational equilibrium probably does not involve a change in the pyranose conformation. The existence of at least two conformers for rotation about the C-4 oxygen to acetate carbonyl bond that interconvert slowly on the 800 MHz NMR timescale, is a probable explanation for the observed field-dependent NMR behaviour of H-4 on glucose-C. We are not aware of any literature reports of similar observations of peak shape change as a function of increasing field strength.

Glycolipids are emerging as an important class of structurally diverse sponge metabolites that frequently have interesting biological activities (14, 16–21). The pachymosides represent a novel family of sponge glycolipids that have unprecedented glycon and aglycon fragments. The fatty acyl component of their aglycon has the same number of carbons and identical positioning of ketone and glycon attachments as the fatty acyl component of erylusamine E (**6**), one of a family of IL-6 receptor antagonists isolated from sponges in the genus *Erylus* (19, 20). However, the sugar residues in **1** (D-glucose and D-galactose) and **6** (D-arabinose and L-xylose) are different, the number and connectivity of the sugars are different, and the amino components of the aglycon amides are different. The pachymosides have an acidic aglycon while the erylusamines all have basic aglycons. One interesting structural feature of the pachymosides is the complete acetylation of the galactose residues and only partial acetylation of the glucose residues. Further investigation of the biological activity of the pachymosides revealed that although they showed promising activity in the TTSS inhibition assay, they are not true TTSS inhibitors. Rather, their observed activity appears to stem from their ability to effectively activate extracellular bacterial proteases that rapidly degrade the excreted Esps, resulting in a false indication of TTSS inhibition in the assay (1).

Experimental

General experimental procedures

^{1}H, COSY-gr, HMBC-gr (optimized for $^{2,3}J$ = 8 Hz), and HMQC-gr for (peracetyl pachymoside A methyl ester) were initially recorded on a Bruker AMX500 NMR spectrometer. These experiments plus 1D- and 2D-TOCSY, ROESY, HSQC, eCOSY, and *J*-resolved experiments were repeated on a Varian INOVA 800 MHz spectrometer at the NANUC facility at the University of Alberta. ^{13}C NMR data were recorded on a Bruker AM400 NMR spectrometer. Chemical shifts were referenced to solvent peaks (δ_H 7.15 ppm, δ_C 128.0 ppm for C_6D_6). Low-resolution electrospray (ES) mass spectra were recorded on a Bruker Esquire LC mass spectrometer, operated by direct infusion. High-resolution ES mass spectra were recorded on a Micromass LCT mass spectrometer. Both low- and high-resolution EI mass spectra were recorded on an AEI MS-50 mass spectrometer. Flash silica gel column chromatography was performed using 230–400 mesh silica gel 60 (Silicycle). Preparative HPLC separations were performed using a Waters 515 pump and a Waters 2487 dual tunable absorbance detector. Solvents were all HPLC grade (Fisher) and filtered prior to use. Thin layer chromatography (TLC) was performed with Merck silica gel 60 F_{254} aluminum sheets, and visualized with vanillin – sulfuric acid spray reagent. Pyridine, methanol, benzene, acetic anhydride, 4-(dimethylamino)pyridine (DMAP), and (trimethylsilyl)diazomethane (solution in hexanes) used in derivatization reactions were anhydrous and (or) reagent grade (Aldrich) and used without further purification. Acetic anhydride (1,1′-$^{13}C_2$, 99%) was purchased from Cambridge Isotope Laboratories, Inc.

Biological material and isolation of pachymosides

Specimens of *P. johnstonia* (650 g wet weight) were collected by hand using SCUBA from rocky subtidal marine

habitats along the coast of the Isle of Mann. A voucher sample (ZMA 17065) has been deposited at the Zoologisch Museum, University of Amsterdam. Freshly collected sponge material was frozen immediately and stored at –15 °C until workup. A portion of the frozen sponge (wet wt 111 g) was cut into pieces and extracted twice with MeOH over a six day period. The combined MeOH extracts were filtered and concentrated by rotary evaporation to an amber semisolid. This residue was partitioned between water (50 mL) and ethyl acetate (ea) (100 mL), and the aqueous layer extracted twice more with ethyl acetate (100 mL ea). The yellow-colored ethyl acetate solution was evaporated to dryness in vacuo to afford 1.60 g of a greenish amber amorphous solid. A portion of this material (0.16 g) was purified by Sephadex™ LH-20 size exclusion chromatography (28 mm ID × 100 cm bed length) eluting with 100% MeOH. Early eluting like fractions (as detected by TLC developed with 15% MeOH–CH_2Cl_2) were combined and concentrated in vacuo to give a mixture of pachymosides as a white amorphous solid (0.106 g). This material exhibited a distribution of molecular ion clusters centred at *m/z* 1973/1974 in the LR-FAB-MS.

Methanolysis of the pachymoside mixture

Pachymosides (60 mg) were dissolved in 10 mL of anhydrous methanol and 0.2 mL of CH_3COCl was slowly added. The reaction mixture was stirred at 70 °C with refluxing for 2 h. MeOH was removed from the reaction mixture in vacuo and the resulting residue was partitioned between H_2O and Et_2O. The aglycon methyl esters (25 mg) went into the organic phase and the methyl glycosides (30mg) went into the aqueous phase.

Acetylation of monosaccharide methylglycosides

The mixture of monosaccharide methyl glycosides (20 mg) obtained from methanolysis of the pachymosides was added to 9 mL of anhydrous pyridine, 3 mL of acetic anhydride, and ~1 mg of DMAP as catalyst. The reaction mixture was stirred overnight at room temperature. After removal of the pyridine and acetic anhydride under high vacuum, the crude product was partitioned between H_2O and Et_2O. The organic extract soluble materials were purified by isocratic normal phase HPLC (eluent: 31% EtOAc:69% hexanes) monitored with a RI detector to give one fraction containing peracetylated glucose/galactose α-methyl glycopyranosides and a second fraction containing peracetylated glucose/galactose β-methyl glycopyranosides.

Preparation of monosaccharide standards

Commercial D-glucose or D-galactose (20 mg, 0.11 mmol) were dissolved in 10 mL MeOH with 0.2 mL of CH_3COCl (2 mmol) and stirred with refluxing at 70 °C for 2 h. The solvent was then removed in vacuo to give a residue that was partitioned between H_2O and Et_2O. The aqueous fraction was dried and added to 9 mL (100 mmol) of anhydrous pyridine and 3 mL of acetic anhydride (30 mmol) with ~1 mg of DMAP as a catalyst. The acetylation mixture was stirred at room temperature overnight. After removal of the pyridine and acetic anhydride under high vacuum, the crude product was partitioned between H_2O and Et_2O and the organic extract was purified by normal-phase HPLC as described above to give pure samples of methyl 2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranoside (**3**) and methyl 2,3,4,6-tetra-*O*-acetyl-α-D-galactopyranoside (**4**).

Chiral GC analysis of alditol acetates

The pachymoside mixture was hydrolyzed for 1 h at reflux in H_2O (20 mL) and 0.25 mL of 6 N HCl. The reaction was worked up by evaporating the H_2O to dryness under high vacuum and the residue was partitioned between H_2O and Et_2O. The water soluble materials were concentrated to dryness via lyophilization and reduced with $NaBH_4$ in MeOH at rt. After removing the MeOH in vacuo, the residue was treated with acetic anhydride and pyridine and a catalytic amount of DMAP to give a mixture of alditol acetates. Normal-phase flash chromatography separated the mixture into glucitol peracetate and galactitol peracetate. Commercial samples of D-glucose, L-glucose, and D-galactose were also converted to their alditol peracetates using the same reaction conditions. The alditol peracetates were analyzed by GC on a Chrial Select 1000 β-Dex-390 column using flame ionization detection. The GC analysis showed that the glucose and galactose obtained from hydrolysis of the pachymoside mixture both had the D configuration.

Aglycon (5)

The Et_2O soluble fraction (25mg) from methanolysis of the pachymoside mixture was loaded on a normal phase silica gel flash column and eluted with a gradient of hexanes–EtOAc (5% EtOAc:95% hexanes – 100%EtOAc). The fractions eluting with 20% EtOAc:80% hexanes contained one major component and one minor component by TLC analysis. The more polar major component was fractionated further by isocratic reversed-phase HPLC (eluent: 88% MeOH:12% H_2O (monitored at 206.0 nm)) to give pure methyl ester **5** (see *Results and discussion* for spectroscopic data).

Esterification of the pachymoside mixture

A 45 mg (ca. 0.024 mmol) sample of the pachymoside mixture was dissolved in a mixture of 0.5 mL each of MeOH and benzene and treated with 0.080 mL of (trimethylsilyl)diazomethane (2.0 mol L^{-1} solution in hexanes) and stirred at ambient temperature for 24 min. The reaction mixture was diluted with 5 mL EtOAc, evaporated to dryness (aspirator), then redissolved in EtOAc (6 mL), washed with sat. NaCl solution, dried (Na_2 SO_4) and again evaporated in vacuo to a film (38 mg). A short reaction time was found to be critical in this step. If left to react overnight, most of the acetate groups appeared to be lost, as determined by NMR. Mass spectroscopy (ES) on this sample again exhibited a distribution of molecular ion clusters centered at 2012 (Na^+ adducts) with a nearly identical pattern to that found for the underivatized material. This represents a shift in molecular weight of plus 14 amu, which is correct for the methyl ester derivative, and confirms that the acetate groups of the pachymosides remained fully intact in this step.

LR-ESI-MS data[a]: (Na^+ adducts) *1956* (55), 1969 (75), *1998* (70), 2012 (100), *2040* (65), 2053 (80); exchanged with CD_3OD: *1965* (50), 1979 (75), *2006* (55), 2018 (100), *2046* (50), 2060 (70). (Italic indicates homolog ions that

would afford pachymoside A peracetate upon full acetylation).

Acetylation of the pachymoside methyl ester mixture

The sample of mixed pachymoside methyl esters (38 mg) was dissolved in anhydrous pyridine (3.0 mL) under argon and treated with 1.3 mg DMAP, then with 1 g of acetic anhydride (or 1,1′-$^{13}C_2$ acetic anhydride) and stirred for 15 h at ambient temperature. After evaporation under high vacuum, the residue was taken up in EtOAc–hexanes (1:1) and subjected to flash column chromatography (12 × 210 mm) eluting with EtOAc–hexanes (4:1) to afford 35 mg of a white solid. TLC in this solvent system showed a spot at 0.18 R_f for the desired derivative. This was taken up in MeOH (2.0 mL) and fractionated by reverse-phase HPLC (Inertsil C_{18} 9.4 × 250 mm, eluting with H_2O–1-PrOH (7:4). UV detection at 202 and 280 nm, up to 3 mg per injection) to give pachymoside A methyl ester (**2**), peracetate (4.0 mg) (or its corresponding 1-^{13}C-acetylated analog).

Data for unlabeled analog

HR-ESI-MS: *m*/*z* 2333.9639 (calcd. for $^{12}C_{105}{}^{13}CH_{159}NO_{54}Na$ 2333.9658); ^{1}H and ^{13}C NMR (800 and 100 MHz, respectively) see Table 1.

LR-ESI-MS: Unlabeled acetate derivative: 2335 (100, Na^+), 1178 ($2Na^+$). ^{13}C-acetate derivative: 2343/2344 (100, Na^+), 1183 (10, $2Na^+$), indicating the presence of eight to nine ^{13}C-acetyl groups.

(Note: LR-ESI-MS data *m*/*z* absolute values estimated at ±1 amu at these values; relative values of the weighted cluster center are reported. Each molecular ion cluster generally shows three or more sequential peaks owing to naturally abundant ^{13}C).

Acknowledgements

Financial support was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) (RJA), Canadian Institutes for Health Research (CIHR) (AG, BBF), and the Canadian Bacterial Disease Network Center of Excellence (BBF). The authors thank M. LeBlanc and D. Williams for collecting *P. johnstonia* and the Canadian National High Field NMR Centre (NANUC) for their assistance and use of the facilities. NANUC is funded by CIHR, NSERC, and the University of Alberta.

References

1. A. Gauthier and B.B. Finlay. ASM News, **68**, 383 (2002).
2. B.A. Vallance and B.B. Finlay. Proc. Natl. Acad. Sci. U.S.A. **97**, 8799 (2000).
3. B. Kenny, R. DeVinney, M. Stein, D. Reinscheid, E.A. Frey, and B.B. Finlay. Cell (Cambridge, Mass.), **91**, 511 (1997).
4. C. Deibel, S. Kramer, T. Chakraborty, and F. Ebel. Mol. Microbiol. **28**, 463 (1998).
5. S.S. Park, R.W. Worobo, and R.A. Durst. Crit. Rev. Biotechnol. **21**, 27 (2001).
6. J. Nataro and J. Kaper. Clin. Microbiol. Rev. **11**, 142 (1998).
7. J. Mecsas and E.J. Strauss. Emerging Infect. Dis. **2**, 271 (1996).
8. C.J. Hueck. Microbiol. Mol. Biol Rev **62**, 379 (1998).
9. G.R. Cornelis and F. Van Gijsegem. Ann. Rev. Microbiol. **54**, 735 (2000).
10. D.J. Faulkner. Nat. Prod. Rep. **19**, 1 (2002), and previous reviews in this series.
11. J.W. Blunt, B.R. Copp, M.H.G. Munro, P.T. Northcote, and M.R. Prinsep. Nat. Prod. Rep. **20**, 1 (2003).
12. R.W. Newbold, P.R. Jensen, W. Fenical, and J.R. Pawlik. Aquat. Microb. Ecol. **19**, 279 (1999).
13. J.E. Thompson, R.P. Walker, and D.J. Faulkner. Mar. Biol. (Berlin), **88**, 11 (1985).
14. R.G. Linington, M. Robertson, A. Gauthier, B.B. Finlay, R. Van Soest, and R.J. Andersen. Org. Lett. **4**, 4089 (2002).
15. A. Kauppi, R. Nordfelth, H. Uvell, H. Wolf-Watz, and M. Elofsson. Chem. Biol. **10**, 241 (2003)
16. E. Fattorusso and A. Mangoni. Marine glycolipids. *In* Progress in the chemistry of organic natural products. *Edited by* W. Hertz, G.W. Kirby, R.E. Moore, W. Steiglich, and Ch. Tamm. Springer, Wein. 1997. pp. 215–301.
17. V. Costantino, E. Fattorusso, and A. Mangoni. Tetrahedron, **56**, 5953 (2000).
18. V. Costantino, E. Fattorusso, C. Imperatore, and A. Mangoni. Eur. J. Org. Chem. 4457 (2001).
19. V. Costantino, E. Fattorusso, A. Mangoni, M. Di Rosa, and A. Ianaro. J. Am. Chem. Soc. **119**, 12 465 (1997).
20. N. Sata, N. Asai, S. Matsunaga, and N. Fusetani. Tetrahedron, **50**, 1105 (1994).
21. R. Goobes, A. Rudi, Y. Kashman, M. Ilan, and Y. Loya. Tetrahedron, **52**, 7921 (1996).