PERGAMON

Phytochemistry 57 (2001) 489–496

PHYTOCHEMISTRY

www.elsevier.com/locate/phytochem

Application of MS and NMR to the structure elucidation of complex sugar moieties of natural products: exemplified by the steroidal saponin from *Yucca filamentosa* L.

Annick Plock [a], Gabriele Beyer [a], Karl Hiller [a], Egon Gründemann [b], Eberhard Krause [c], Manfred Nimtz [d], Victor Wray [d,*]

[a]*Institut für Pharmazie, Humboldt-Universität zu Berlin, Goethestr. 54, 13086 Berlin, Germany*
[b]*Institut für Angewandte Chemie Berlin Adlershof e.V., Richard-Willstätter-Str. 12, 12489 Berlin, Germany*
[c]*Forschungsinstitut für Molekulare Pharmakologie, Alfred-Kowalke-Str. 4, 10315 Berlin, Germany*
[d]*GBF, Gesellschaft für Biotechnologische Forschung mbH, Mascheroder Weg 1, 38124 Braunschweig, Germany*

Received 18 September 2000; received in revised form 18 January 2001

Abstract

An approach, using well characterized procedures, is presented that should be of general applicability for the structural elucidation of complex sugar moieties of natural products. The methods used are exemplified by the structure elucidation of a new gitogenin-based steroidal saponin that has a strong leishmanicidal activity similar to preparations used in clinical practice and has been isolated by bioactivity-guided fractionation of the ethanolic extract of *Yucca filamentosa* L. leaves. The saponin has been characterized as 3-*O*-((ß-D-glucopyranosyl-(1→3)- ß-D-glucopyranosy-(1→2))(α-L-rhamnopyranosyl-(1→4)-ß-D-glucopyranosyl-(1→3))-ß-D-glucopyranosyl-(1→4)-ß-D-galactopyranosyl)-25R,5α-spirostan-2α,3ß-diol. © 2001 Published by Elsevier Science Ltd. All rights reserved.

Keywords: *Yucca filamentosa* L.; Agavaceae; Steroidal saponin; Gitogenin; Leishmaniasis; Sugar analysis

1. Introduction

The complexity of many bioactive natural products containing sugar residues dictates the use of a combination of modern techniques dominated by those involving mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy for structure elucidation. In more and more cases the use of a single technique precludes a satisfactory unambiguous determination. In the majority of cases the approach to the complete unambiguous structure can best be categorized as the elucidation of the sugar unit(s) and the aglycone. Here we outline a general approach shown in Scheme 1 and apply it to the structure elucidation of a complex steroidal saponin isolated from *Yucca filamentosa* L. as example. It should be emphasized here that the aim of the application of these techniques is to obtain an unambiguous structure making the most efficient use of modern instrumental techniques. As such the complete assignment of the NMR data is not a prerequisite and, as the molecules become more complex, often not feasible in a reasonable time span even using the highest field NMR spectrometers.

Y. filamentosa L. belongs to the Agavaceae family which has more than 580 species which are widely distributed in tropic and subtropic dry climate regions throughout the world. It is indigenous to northern and middle America, and in Europe it is cultivated as an ornamental plant. The plant is rich in steroidal saponins (Mahato et al., 1982; Voigt and Hiller, 1987). Our interest in these plants arose from their potential in the treatment of leishmaniasis, an increasing major global public-health problem. The pentavalent antimonials are the preferred drugs for the treatment of most forms of leishmaniasis, although treatment with these agents is not consistently effective and has serious toxic side effects. It is known that natural products have potential in the search for new and selective agents (Akendengue

* Corresponding author. Tel.: +531-6181-362; fax: +531-6181-355.
E-mail address: vwr@gbf.de (V. Wray).

PII: S0031-9422(01)00035-8

Sugar identification	**Aglycone identification**
a) Component monosaccharide identification and, when necessary, absolute configuration determination by chromatographic comparisons after acid hydrolysis b) Sugar-linkage analysis by GC-MS of partially methylated alditol acetates c) Sugar unit sequence from fragmentation ion identification in the MS-MS mode of ESI-MS of the native and permethylated molecule *Combined data from a) to c) affords the partial identity of sugar units and sugar sequence* d) Anomeric configuration and sugar conformation from NMR coupling constant (^{3}J(HH) and ^{1}J(CH)) data e) Identity and sequence confirmation from correlations in HMBC spectrum	a) Molecular weight from MS data b) 1D ^{1}H and ^{13}C data comparison with literature data for known aglycones c) Analysis of HMBC spectrum to confirm b) or to elucidate new structure. Most profitable approach involves the elucidation of correlations of the ^{1}H signals with neighbouring carbons two and three bonds removed in the order: i) Correlations from methyl groups ii) Correlations to and from substituents iii) Correlations from hydrogens bonded to heteroatom bonded carbons iv) Correlations to other hydrogens showing well defined ^{1}H signals assigned in the COSY spectrum d) Configuration analysis from "through-space" correlations in 1D NOE, and/or 2D NOESY/ROESY

Scheme 1. Steps required in the structure elucidation of sugar containing natural products.

et al., 1999). In a preliminary screening the ethanolic extract of *Y. filamentosa* L. showed a strong leishmanicidal activity against *Leishmania mexicana amazonensis*, and the bioactivity-guided fractionation of this extract led to the isolation of a new gitogenin-based steroidal saponin, 1. This activity has been studied by extracellular and intracellular assays with promastigote and amastigote forms of *Leishmania* spp. using microscopical and flow cytometric methods (Plock et al., 2001). Its extra cellular effect on promastigote cultures of *L. mexicana amazonensis* is as high as that of pentamidine which is a drug used in clinical practice. In an intracellular assay using macrophages infected with GFP-transfected *Leishmania major*, 1 showed a dose-related decline of the infection rate and of the number of parasites per cell. This leishmanicidal activity was shown to be at least partially due to the action on the parasite membrane. Clearly for a detailed understanding of such effects an unambiguous structure determination of this compound was desirable.

2. Results and discussion

As more and more modern NMR facilities become available it has become routine to use through-bond correlations, such as homonuclear ^{1}H–^{1}H COSY and TOCSY and inverse ^{1}H-detected heteronuclear ^{13}C–^{1}H HMQC and HMBC, for structure elucidation of sugar-containing molecules. Numerous publications report such determinations and in these a full assignment of the ^{1}H and ^{13}C data are usually attempted. However, as structures become more complex, particularly as the number of sugar units increases, this approach has its limitations and additional information from alternative sources is required. In particular information regarding the nature and linkage of the sugar can now be supplied from MS techniques that require only small amounts of material. We have already reported on the power of these techniques for the rapid, comparative analysis of saponin mixtures (Schöpke et al., 1996a). As an extension of this, here we show that a combination of MS and NMR techniques, in which MS data are initially used and are then supplemented with NMR data, afford an unambiguous structure elucidation without necessarily having recourse to a full, and often extremely time-consuming, complete assignment of the NMR data. Thus the step-by-step approach outlined in Scheme 1, and discussed in more detail below for a specific novel compound from *Y. filamentosa* L., has been used to afford structural information about the saponin in which the nature, position of attachment, and sequence of the sugar constituents as well as details of the aglycone have been umambiguously identified.

2.1. Component sugar identification

Initially a bioactivity-guided fractionation of the ethanolic extract of *Y. filamentosa* L. leaves by repeated column chromatography on silica gel and on reversed phase silica gel gave the leishmanicidal steroidal saponin **1**. Acidic hydrolysis yielded the aglycone and the sugars moieties, which by TLC comparison with authentic samples indicated the presence of galactose (Gal), glucose (Glc) and rhamnose (Rha). Although we have not determined the absolute configuration of the sugar units here they can be analysed using methods described by us previously (Schöpke et al., 1996b).

2.2. Sugar-linkage analysis

The sugar substitution pattern can be established after derivatization (methylation analysis) using micro-amounts of material (ca. 1 nmol) by methods originally developed by us for the oligosaccharide mapping of recombinant and natural glycoproteins (Grabenhorst et al., 1995). Thus the GC–MS of the partially methylated alditol acetates derived from **1** showed the presence of six partially methylated alditol acetates representing terminal rhamnose, terminal glucose, 3-substituted glucose, 4-substituted galactose, 4-substituted glucose and 2, 3-disubstituted glucose residues (Table 1). An aliquot of the intact permethylated compound was reserved for a MS–MS analysis reported below.

2.3. Sugar sequence analysis

Electrospray ionization in the positive as well as the negative ion mode, especially in combination with MS–MS techniques, has quickly become the method of choice for sugar sequence analysis. The methodology has the advantage of requiring only minute amounts of substance (< 1 μg), and affords a direct analysis of both purified materials and mixtures without the necessity of derivatization. We have used this method previously for comparative analysis of saponin mixtures and have shown that the introduction of sodium salts has the advantage of producing sodium-adduct molecular ions, which resolves ambiguities arising in protonated molecular-ion spectra and afford readily interpretable spectra (Schöpke et al., 1996a). Thus the positive ion mode ESI-mass spectrum of **1** afforded a molecular ion at *m/z* 1412 $[M+Na]^+$. This parent ion was selected using the quadrupole mass analyser and collision-induced decomposition gave daughter ions that were separated in the time-of-flight mass analyser. The resulting MS–MS spectrum showed fragment ions at *m/z* 979 $[hex+hex+hex+hex+hex+dhex+Na]^+$, *m/z* 817 $[hex+hex+hex+hex+dhex+Na]^+$, *m/z* 671 $[hex+hex+hex+hex+Na]^+$, *m/z* 655 $[hex+hex+hex+dhex+Na]^+$, *m/z* 509 $[hex+hex+hex+Na]^+$ and at *m/z* 347 $[hex+hex+Na]^+$. Peaks at *m/z* 1266 $[M\text{-dhex}+Na]^+$, *m/z* 1250 $[M\text{-hex}+Na]^+$, *m/z* 1104 $[M\text{-dhex-hex}+Na]^+$, *m/z* 1088 $[M\text{-hex-hex}+Na]^+$, *m/z* 942 $[M\text{-dhex-hex-hex}+Na]^+$, *m/z* 780 $[M\text{-dhex-hex-hex-hex}+Na]^+$ and at *m/z* 618 $[M\text{-dhex-hex-hex-hex-hex}+Na]^+$ were generated through loss of the respective saccharides. The appearance of a peak at *m/z* 979 which is generated by a hexasaccharide + Na-fragment and the presence of two terminal sugar units in the GS–MS analysis above indicates **1** is a monodesmoside with a branched sugar chain.

The above MS–MS analysis can of course be applied to the permethylated sample, which has to be produced during the sugar-linkage analysis. Such an analysis has the advantage of avoiding ambiguities in the fragmentation pattern of the native molecule. Thus in the present example the sequential elimination of two terminal monosaccharide units yields the same fragment ion mass as the loss of a single disaccharide unit from the native molecular ion, whereas from the permethylated sample these two fragments differ by one methyl group and therefore can be unequivocally distinguished. In addition the number of acidic protons in the aglycone, which are also permethylated, can be determined. The MS–MS analysis of the permethylated derivative of **1** is summarized in Fig. 1.

Thus the combined data indicated that the sugar sequence in 1 is {Glc-hex}{Rha-hex}-Glc-hex-R and are compatible with a molecular formula of $C_{63}H_{104}O_{33}$, Scheme 2.

At this stage in the analysis the combined methods afford information on the number, the nature and linkage positions, but not of the anomeric configuration or the exact sequence, of the constituent monosaccharides in the glycoside. The molecular weight of the agylcone is evident by default. NMR data are now necessary to afford the remaining stereochemical details of the sugars and identification of the aglycone.

In contrast to MS analyses considerably more material is required for NMR analyses. Thus in reality satisfactory 1D and 2D ^{1}H homonuclear spectra can be recorded of amounts > 100–500 μg, while the less sensitive 2D inverse ^{1}H-dectected heteronuclear ^{13}C–^{1}H correlations require > 1–2 mg of homogeneous material.

2.4. Anomeric configuration and sugar conformation

The anomeric configurations of the various sugar moieties followed from the magnitudes of the vicinal ^{1}H–^{1}H and one-bond ^{13}C–^{1}H coupling constants shown in Tables 2 and 3, respectively. All glucoses and galactose had a ß-configuration and the α-configuration of the rhamnose was unambiguously established from the ^1J(CH) of 169.6 Hz (Hansen, 1981). The latter is the most reliable coupling for establishing the rhamnose configuration as the similarity of the gauche dihedral angle for both α and ß anomers leads to similar small vicinal ^{1}H–^{1}H coupling constants. The linkage position of the sugar chain to the aglycone follows from either a

Table 1
Methylation analyses of **1** (derivatives are listed in the order of elution)

Per-*O*-acetylated derivatives with *O*-Me in position	Type	Interpretation
2,3,4	Rhamnitol	Terminal rhamnose
2,3,4,6	Glucitol	Terminal glucose
2,4,6	Glucitol	3-Substituted glucose
2,3,6	Galacitol	4-Substituted galactose
2,3,6	Glucitol	4-Substituted glucose
4,6	Glucitol	2,3-Disubstituted glucose

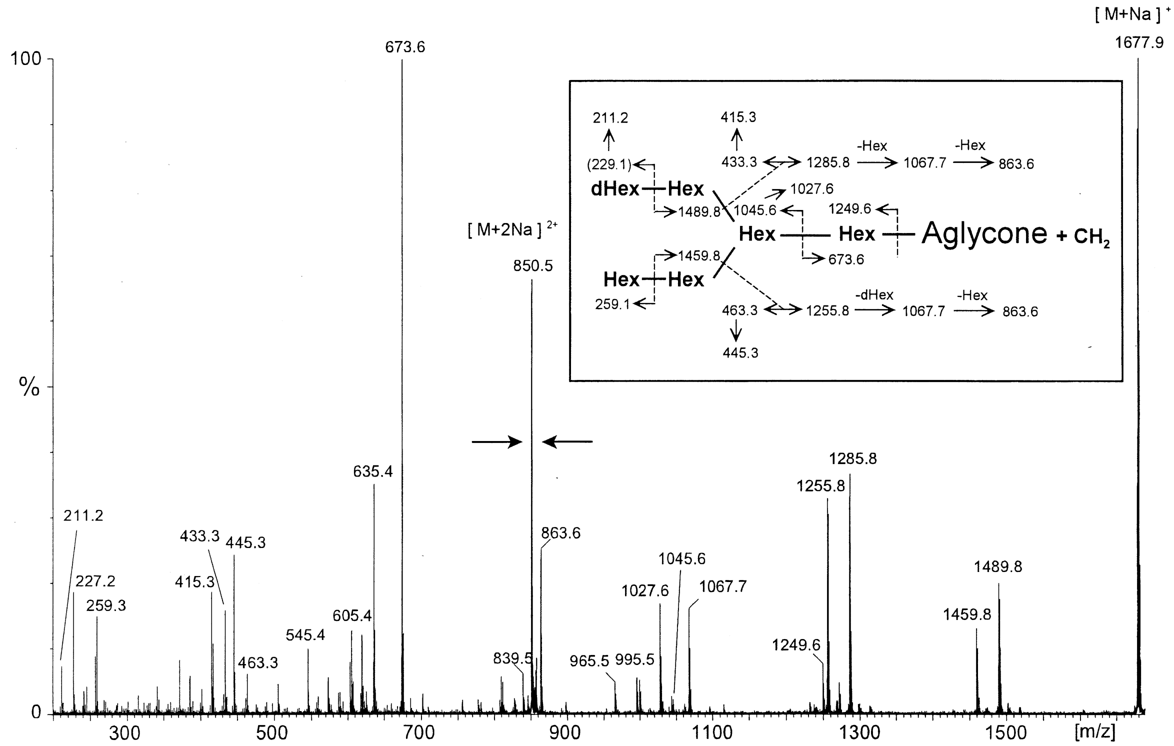

Fig. 1. ESI daughter ion spectrum of the doubly charged $[M+2Na]^{2+}$ permethylated molecular ion of the saponin isolated from *Yucca filamentosa* L. with inserted fragmentation scheme. The fragmentation pattern indicates a branched hexasaccharide glucan structure consisting of five hexose and one deoxyhexose residues.The aglycone has one acidic proton, as is indicated by its derivatisation with one methyl group upon permethylation. Hex = hexose, dHex = deoxyhexose.

a) Component identification: D-galactose (Gal), D-glucose (Glu), L-rhamnose (Rha)

b) Sugar-linkage analysis: 4-Gal, 3-Glu, 4-Glu, 2,3-Glu, terminal Glu, and terminal Rha in a ratio of ca. 1:1:1:1:1:1

c) Sugar sequence: deoxyhexose(Rha) - hexose - hexose(Glu) - hexose
|
hexose(Glu) - hexose

d) NMR coupling data: 1α-deoxyhexose (Rha) - 1ß-hexose - 1ß-hexose(Glu) - 1ß-hexose
|
1ß-hexose(Glu) - 1ß-hexose

e) 2D NMR data: Rha- (1α-4) -Glu- (1ß-3) -Glu- (1ß-4) -Gal- 1ß -Aglycone
|
Glu- (1ß-3) -Glu- (1ß-2)

Scheme 2. Structure elucidation of sugar chain in **1**.

comparison of ^{13}C data with that of the unsubstituted aglycone or more explicitly from vicinal ^{13}C–^{1}H correlations in the HMBC spectrum (see below, Table 4 and Scheme 2).

2.5. Exact sugar sequence

The identity and exact sugar sequence may already be apparent from the data accumulated above. If this is not the case, as here, then a partial interpretation of the 2D homonuclear and heteronuclear NMR correlations is necessary. The extent of such an analysis depends on the molecule under investigation, but in general will be of the form used for **1**.

The observation of a correlation of the anomeric proton of Gal (Table 2) to a low field carbon (C-3) of the aglycone, evident from further long-range correlations to aliphatic, non-sugar protons, identified the chain linkage to the aglycone. The H1–H4 spin system of the galactose unit was readily identified from the 2D COSY

Table 2
^{1}H NMR data for the sugar and aglycone moieties of **1**

	Sugar moiety		
	^{1}H shift, J in Hz		^{1}H shift, J in Hz
3-*O*-Gal		Glc (C)	
H-1	4.39, J(1–2) 7.8	H-1	4.61, J(1–2) 7.8
H-2	3.74, J(2–3) 9.7	H-2	3.34, J(2–3) 9.1
H-3	3.58, J(3–4) 3.2	H-3	3.44, J(3–4) 9.1
H-4	4.08, J(4–5) < 2	H-4	#[c]
H-5	3.59, J(5–6a) 7.2	H-5	*
H-6a	3.90	H-6a	*
H-6b	3.71, J(6a–6b) 11.2	H-6b	*
Glc (A)		Glc (D)	
H-1	4.66, J(1–2) 7.7	H-1	4.81, J(1–2) 8.0
H-2	3.80	H-2	3.35, J(2–3) 8.9
H-3	3.81	H-3	3.53, J(3–4) 9.0
H-4	3.39	H-4	3.58, J(4–5) 9.2
H-5	*[b]	H-5	3.47
H-6a	*	H-6a	3.87
H-6b	*	H-6b	3.69
Glc (B)		Rha	
H-1	5.07, J(1–2) 8.0	H-1	4.84[a] J(1–2) small
H-2	3.45, J(2–3) 9.0	H-2	3.87, J(2–3) small
H-3	3.63, J(3–4) 9.0	H-3	3.66, J(3–4) 9.6
H-4	3.49	H-4	3.45, J(4–5) 9.6
H-5	3.42, J(5–6a) 2.3	H-5	4.02, J(5–6) 6.2
H-6a	3.95	H-6	1.31
H-6b	3.87, J(6a–6b) 12.2		
Aglycone (only unambiguous assignments are shown)			
H-1a	2.03, (1a–1b) 12.6, (1a–2) 4.9		
H-1b	0.98, (1b–2) 12.4		
H-2	3.70		
H-3	3.52		
H-4a	1.77		
H-4b	1.45		
H-14	1.20		
H-15a	2.01		
H-15b	1.29		
H-17	1.78, (17–20) 6.9, (17–16) 6.7		
Me-18	0.93		
Me-19	0.83		
H-20	1.94, (20–21) 7.0		
Me-21	0.99		
H-23a	1.74, (23a–23b) 13.4, (23a–24) 4.7		
H-23b	1.60		
H-25	1.63		
Me-27	0.83, (25–27) 6.4		

[a] Under residual water signal.
[b] *There are two interchangeable sets of signals for Glc(A) and Glc(C) for H-6a/H-6b/H-5 at 3.96/3.64/3.39 and 3.93/3.68/3.38 with J(6a–6b)/J(6a–5) of ~12/~2 and 11.9/2.5 respectively.
[c] # Not assignable.

Table 3
^{13}C NMR chemical shifts and ^{1}J(CH)[a] of compound **1** (CD_3OD)

	Sugar moiety		
	^{13}C shift (^{1}J(CH))		^{13}C shift (^{1}J(CH))
3-0-Gal 1	102.9 (160.2)	Glc (C) 1	105.3 (163.0)
2	72.9	2	$
3	$[c]	3	#[d]
4	79.8	4	70.3
5	$	5	#
6	61.4	6	§[b]
Glc (A) 1	104.3 (163.6)	Glc (D) 1	103.9 (163.3)
2	81.0	2	$
3	87.5	3	76.9
4	71.6	4	79.5
5	#	5	#
6	§	6	61.8
Glc (B) 1	103.8 (166.4)	Rha 1	102.9 (169.6)
2	75.0	2	72.5
3	87.6	3	72.2
4	70.7 or 72.5	5	73.8
5	#	5	70.7
6	62.9	6	17.9
Aglycone			
	^{13}C shift		^{13}C shift
1	45.8	15	32.4
2	71.0	16	82.2
3	85.0	17	63.9
4	35.8	18	16.9
5	45.8	19	13.7
6	28.8	20	42.9
7	32.8	21	14.8
8	34.0	22	110.5
9	55.7	23	33.5
10	37.9	24	29.8
11	22.3	25	31.4
12	41.1	26	67.8
13	41.7	27	17.4
14	57.4		

[a] The magnitude of ^{1}J(CH) of the anomeric carbons of the sugar moieties, given in parentheses, were determined from the residual cross peaks in the inverse long-range ^{13}C–^{1}H correlation (HMBC).
[b] § The C-6 shifts at 62.7 and 63.1 ppm correspond to ^{1}H shifts at 3.93/3.68 and 3.96/3.64 ppm, respectively (see footnote of Table 2).
[c] $ Shifts at 75.5, 75.6, 75.6, 75.7 ppm.
[d] # Shifts at 77.1, 77.5, 77.8, 78.0, 78.1 ppm.

correlations in which H-4 was to low field and had two small vicinal ^{1}H–^{1}H couplings that are characteristic of this system. These observations implied the galactose unit was adjacent to the aglycone and the other four hexose units were glucose. The anomeric proton of the terminal rhamnose correlated with a sugar carbon at 79.5 ppm, which from the chemical shift must be that of a 4-substituted glucose unit, Glc (D) in Table 3 (C-3 of a 3-substituted glucose unit would be at ca. 87 ppm). Consequently these data and the linkage analysis above indicates the terminal glucose must be bound to the 3-substituted glucose, Glc (B) in Table 3, by default. The only question that now remains is the relative position of these two dissaccharides (glucose 4-1α rhamnose and glucose 3-1ß glucose) on the branching trisubstituted glucose unit, Glc (A) (Scheme 2). Again a partial analysis of the COSY spectrum and the considerable differences in

Table 4
Characteristic long-range correlations in the HMBC spectrum of the aglycone of **1**

Hydrogen	Carbons
Correlations from methyl groups	
Me-18	C-12, C-13, C-14, C-17
Me-19	C-1, C-5, C-10, C-9
Me-21	C-17, C-20, C-22
Me-27	C-24, C-25, C-26
Correlation from sugar substituent	
H-1 Gal	C-3
Correlations from hydrogens bonded to heteroatom carrying carbons[a]	
H-26A/B	C-22, C-24, C-25, C-27
Correlations from selective well defined ^{1}H signals assigned from COSY spectrum	
H-1A[b]	C-2, C-3, C-5, C-10, C-19
H-1B[b]	C-2, C-3, C-9, C-10, C-19
H-15A	C-13, C-14, C-16, C-17
H-20	C-13, C-17, C-21, C-22, C-23
H-17	C-12, C-13, C-16, C-18, C-20, C-21

[a] In the present case there are no long-range correlations apparent from H-2 and H-3.
[b] There are some differences in the carbons showing correlations due to the stereochemical dependence of the appropriate long-range coupling constants.

chemical shifts of C-2 and C-3 of Glc (A) together with the observation of the respective long-range correlations in the HMBC spectrum from the anomeric protons of Glc (B) and Glc (D), respectively, to these carbons established the relative chain orientations shown in Scheme 2 (e).

Again we emphasize that the unambiguous structure of the sugar unit was established without a complete assignment of the NMR data. Even on close inspection of the 2D correlations a number of assignments both in the ^{1}H and ^{13}C spectra (Tables 2 and 3, respectively) could not be unambiguously establish even at the high magnetic fields employed here (14 Tesla, 600 MHz: ^{1}H; 150 MHz: ^{13}C) because of signal overlap. As the complexity of the sugar chains increases full assignments will be even more difficult although it can be anticipated that correlations to the low-field carbons, i.e. those carrying the substituents and hence the information regarding the exact linkage, will still be determinable.

2.6. Aglycone identification

In many cases the aglycone can be readily identified from its molecular weight deduced from the MS data and from careful comparison of the ^{1}H and ^{13}C data with known literature data. In the present case the aglycone gitogenin was unambiguously identified (Agrawal et al., 1985; Nakano et al., 1988). When this is not possible then a detailed analysis of the heteronuclear and homonuclear "through-bond" NMR correlations is the most rigorous approach, followed by "through-space" data from 1D NOE and/or 2D/NOESY/ROESY spectra to determine the relative configurations of asymmetric centres.

In the present case a systematic inspection of the HMBC data detailed in Scheme 1 (right, c) for **1** yields the sets of correlations given in Table 4. The strongest correlations are from the methyl groups and provide correlations to all adjacent carbons two and three bonds from the respective protons. Additional correlations from the galactose (above) defines the sugar chain position while the correlations from H-26A/B and from a number of the well defined protons identified in the COSY spectrum are sufficient in combination to define the molecular fragments incorporating C-9/C-10(C-19)/C-1→C-5 and C-12/C-13(C-18)/C-14→C-27 of Fig. 2. These data provide unambiguous confirmation of the structure of the known aglycone. They would not, however, be sufficient to unambiguously define an unknown aglycone as the central section C-6/C-7/C-8 and C-11 together with their attached protons remain undefined. In such cases further 2D spectra would be necessary, in particular HMQC and "through-space" correlations.

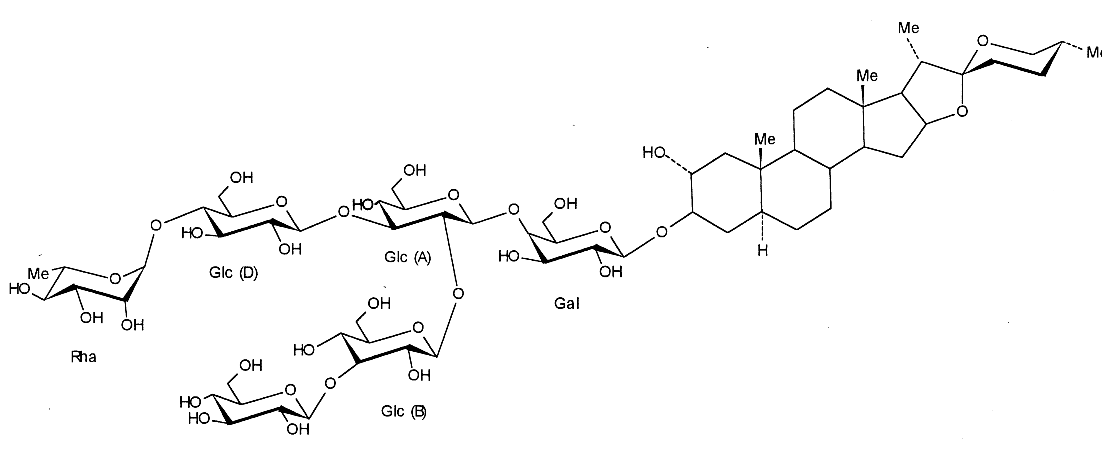

Fig. 2. Structure of **1** from *Yucca filamentosa* L.

3. Conclusion

Adherence to the procedures outlined in Scheme 1 affords an unambiguous identification of complex sugar moieties without recourse to full signal assignments. Similarly known aglycone structures can be readily identified and their structures confirmed with data already present in spectra recorded for the sugar analysis. In those cases where a novel aglycone is present additional NMR data will probably be necessary. In the present case the combined chemical and spectroscopic data, together with reasonable assumptions for the absolute stereochemistries of the sugar moieties, afforded unambiguous evidence for the novel structure for **1** of 3-*O*-((ß-D-glucopyranosyl-(1→3)-ß-D-glucopyranosyl-(1→2)) (α-L-rhamnopyranosyl-(1→4)-ß-D-glucopyranosyl-(1→3))-ß-D-glucopyranosyl-(1→4)-ß-D-galactopyranosyl)-25*R* 5α -spirostan-2α, 3ß-diol, Fig. 2.

4. Experimental

4.1. General experimental procedures

1D (^{1}H, ^{13}C and DEPT-135) and 2D (COSY and HMBC) NMR spectra of compound **1** were recorded in CD_3OD at 300 K on a Bruker AVANCE DMX-600 NMR spectrometer (^{1}H: 600.13 MHz, ^{13}C: 150.92 MHz) using standard Bruker software. Chemical shifts are reported in ppm relative to TMS but were recorded relative to the residual signal of the solvent (^{1}H: 3.35 ppm, ^{13}C: 49.00 ppm) and coupling constants in Hz. The NMR spectra of the aglycone of **1** were recorded in $CDCl_3$ on a Varian 500 MHz spectrometer. A Micromass QTOF2 mass spectrometer equipped with a non-spray ion source was used for ESI–MS and MS–MS. TLC was performed on silica gel 60 F_{254} or on reversed phase 18 silica gel plates or foils (Merck); solvent system A: $CHCl_3$–MeOH–H_2O (6:4:1) and B: MeOH–H_2O (8.5:1.5). The spots of the TLC were detected by 10% H_2SO_4 followed by heating for saponin and sapogenin, and with thymol/H_2SO_4 for sugars.

4.2. Plant material

The leaves of *Y. filamentosa* L. were provided by Dr. Paul Brueckner, Arboretum, Institute of Biology, Humboldt-University of Berlin, Germany, in July 1996. A plant specimen is on file in the Institut für Pharmazie, Humboldt-Unversität, Berlin.

4.3. Extraction and isolation

The air-dried pulverized leaves (500 g) were first macerated with petroleum ether, then with ethyl acetate, each for 90 min. Finally the dried material was extracted with ethanol/water 7:1 in a Soxhlet apparatus for 120 min. After filtration, the solvent was evaporated under reduced pressure to yield 31 g of extract. The isolation of the active component of this extract was performed by bioactivity-guided fractionation. The extract was chromatographed on 2000 g silica gel 60, 0.063–0.2 µm (Merck) with solvent system A as eluent to give 150 fractions. These fractions were combined into pools A–D according to their similar TLC patterns (silica gel, solvent system A). Pool A (6.5 g) showed high leishmanicidal activity and was chromatographed on 600 g reversed phase 18 silica gel material (Merck) with solvent system B as eluent to give pools Aa–Ad. The chromatography of Ac (1.3 g) on 100 g silica gel 60 with $CHCl_3$–MeOH–H_2O (10:3:1, lower phase) led to pure **1** (25.7 mg), (TLC solvent system A; R_f 0,4). **1**: Colorless needles (from MeOH), mp 292–293°C; ESI–MS positive mode *m*/*z*: 1412 $[M+Na]^+$, 1266, 1250, 1104, 1088, 942, 780, 618 (see text for peak assignments). ^{1}H and ^{13}C NMR data are given in Tables 2–4.

4.4. Acidic hydrolysis and aglycone identification

1 was hydrolysed with 1N HCl by refluxing for 90 min on a water bath. The reaction mixture was neutralized with 1 N NaOH and submitted to liquid-liquid partition with diethyl ether. The ether phase was evaporated to dryness under reduced pressure. From the residue gitogenin was identified: colorless needles (from $CHCl_3$), mp 268–270°C; ms 432 $[M]^+$, 417, 318, 289, 139; ^{1}H NMR ($CDCl_3$): $\delta=0.76$ (3H, *s*), 0.78 (3H, *d*, $J=6.3$ Hz), 0.86 (3H, *s*), 0.95 (3H, *d*, $J=6.8$ Hz), 3.30–3.63 (4H, *m*), 4.36 *m* (1H, *m*); ^{13}C NMR ($CDCl_3$, C-1–C-27): 45.0, 73.0, 76.4, 35.6, 44.8, 27.8, 31.8, 34.4, 54.3, 37.6, 21.2, 39.9, 40.6, 56.1, 31.4, 80.8, 62.1, 16.5, 13.6, 41.5, 14.5, 109.2, 32.0, 28.8, 30.3, 66.8, 17.1

4.5. Component monosaccharide identification

1 was hydrolysed at 100°C for 1 h on TLC silica gel plates in a chamber filled with conc. HCl. After separation with solvent system A, galactose (R_f 0.36), glucose (R_f 0.39) and rhamnose (R_f 0.57) were identified by comparison with authentic samples.

4.6. Methylation analysis of the sugar constituents

For linkage analysis, the sample was permethylated as described (Anumula and Taylor, 1992). Purification of the permethylated sample, hydrolysis, reduction and acetylation were performed as described previously (Nimtz et al., 1990). The resulting partially methylated alditol acetates were analysed on a ThermoQuest GCQ ion trap GC/MS system (MS: EI mode, GC: 30 m DB 5 column).

Acknowledgements

We gratefully acknowledge the technical assistance of B. Jaschok-Kentner, C. Kakoschke, E. Surges and A. Tiepold (all GBF), and a grant from the Deutsche Forschungsgemeinschaft (Hi 505/3-2).

References

Agrawal, P.K., Jain, D.C., Gupta, R.K., Thakur, R.S., 1985. Carbon-13 NMR spectroscopy of steroidal sapogenins and steroidal saponins. Phytochemistry 24, 2479–2496.

Akendengue, B., Ngou-Milama, E., Laurens, A., Hocquemiller, R., 1999. Recent advances in the fight against leishmaniasis with natural products. Parasite 6, 3–8.

Anumula, K.R., Taylor, P.B., 1992. A comprehensive procedure for preparation of partially methylated alditol acetates from glycoprotein carbohydrates. Anal. Biochem. 203, 101–108.

Grabenhorst, E., Hoffmann, A., Nimtz, M., Zettlmeissl, G., Conradt, H., 1995. Construction of stable BHK-21 cells coexpressing human secretory glycoproteins and human Gal(ß1-4)GlcNAc-R α2,6-sialyltranferase. Eur. J. Biochem. 232, 718–725.

Hansen, P.O., 1981. Carbon-hydrogen spin-spin coupling constants. Prog. NMR Spectrosc. 14, 175–296.

Mahato, S.B., Ganguly, A.N., Sahu, N.P., 1982. Steroid saponins. Phytochemistry 21, 959–978.

Nakano, K., Matsuda, E., Tsurumi, K., Yamasaki, T., Murakami, K., Takaishi, Y., Tomimatsu, T., 1988. The steroidal glycosides of the flowers of *Yucca gloriosa*. Phytochemistry 27, 3235–3239.

Nimtz, M., Noll, G., Pâques, E.-P., Conradt, H.S., 1990. Carbohydrate structures of a human tissue plasminogen activator variant expressed in recombinant Chinese hamster ovary cells. FEBS 271, 14–18.

Plock, A., Sokolowska-Köhler W., Presber W., 2001. Application of flow cytometry and microscopical methods to characterize the effect of herbal drugs on *Leishmania* spp. Exp. Parasitol., in press.

Schöpke, T., Hiller, K., Wray, V., Nimtz, M., 1996a. Application of MS–MS for the rapid, comparative analysis of saponin mixtures as exemplified by the deacylated and partially deacylated triterpenoid saponins of *Bellis annua*. Planta Med. 62, 336–340.

Schöpke, T., Wray, V., Nimtz, M., Hiller, K., 1996b. Triterpenoid saponins from *Bellium bellidioides*. Structure of the minor deacylsaponins. Phytochemistry 41, 1399–1403.

Voigt, G., Hiller, K., 1987. Neuere Ergebnisse zur Chemie und Biologie der Steroidsaponine. Scientia Pharmaceutica 55, 201–227.