

NMR structural analysis of a novel triterpenoid saponin extracted from *Bellis perennis* L.

Michal Štujber^a, Didem Sohretoglu^b, Tibor Liptaj^a

^aDepartment of NMR Spectroscopy and Mass Spectrometry, Faculty of Chemical and Food Technology, Slovak University of Technology, Radlinského 9, 812 37, Bratislava, Slovakia

^bDepartment of Pharmacognosy, Faculty of Pharmacy, Hacettepe University, 06100 Ankara, Turkey
michal.stujber@stuba.sk

Abstract: One of the chromatographic fractions of an extract from *Bellis Perennis* has been analyzed using NMR. The major component of the fraction has been identified to be a novel triterpenoid saponin: 3-*O*- α -rhamnopyranosyl polygalacic acid 28-*O*-{ α -rhamnopyranosyl (1–3)- β -xylopyranosyl(1–4) α -rhamnopyranosyl(1–2)-[α -arabinofuranosyl(1–3)-4-*O*-acetyl- β -fucopyranoside]}.

Keywords: *Bellis perennis*, NMR, saponin

Introduction

Bellis perennis (common daisy) is a widespread plant. It has been used in traditional medicine for treating wounds, catarrh, rheumatism, arthritis, liver, kidney and respiratory tract disorders (Grieve 1982; Chevallier 1996; Dobelis 1989). The plant has recently been an object of several phytochemical studies (Yoshikawa et al. 2008; Morikawa et al. 2008; Li et al. 2005; Schöpke et al. 1991; Gudej and Nazaruk 2001; Avato and Tava 1995). These studies indicated the presence of a variety of triterpenoid compounds, especially saponins. The occurrence of saponins in the plant kingdom was summarized by Vincken et al. 2007.

An extraction followed by a chromatographic separation was performed by the group of Didem Sohretoglu. One of the fractions showed strong antitumor activities (Sohretoglu et al. 2012). This fraction was the object of our NMR structural analysis.

Materials and Methods

Approximately 3 mg of white powder were dissolved in CD₃OD. The sample was placed into a 5 mm NMR tube.

The sample was measured on two 600 MHz VN-MRS spectrometers, one equipped with a standard indirect detection triple-resonance probe (Faculty of Chemical and Food Technology, Slovak University of Technology) and the other equipped with a ¹³C enhanced indirect detection cold probe (Institute of Chemistry, Slovak Academy of Sciences).

A series of 1D and 2D spectra was measured including variations of the HSQC, HMBC, TOCSY, ROESY and DOSY experiments. All spectra were measured at the temperature of 25 °C. A broad-

band proton spectrum was measured with an acquisition time (at) of 1.7 s and a relaxation time (d1) of 1 second. A multi-frequency suppression (residual HDO and the CD₂H and OH group from methanol) was then performed with the PRESAT sequence with presaturation time d1 set to 2s. The carbon spectrum was measured with at = 1.1 s and d1 = 1 s. Backward linear prediction was used to eliminate the problem with pulse ringdown, which occurs on ¹³C enhanced cold probes because of a high Q-value of the resonant circuit. The spectral widths for 2D spectra were optimized from these 1D spectra. The relaxation and acquisition times were 1s long in all 2D experiments except of HSQC, where the acquisition time was shortened because of carbon decoupling. Forward linear prediction was used in the f1 dimension with 512 basis points and 512 predicted points. The Fourier number in the f1 dimension was 4096. The delays for antiphase creation in HSQC were optimized for a J-coupling of 140 Hz. Two HMBC spectra were measured with the delays optimized for 8 Hz and 3 Hz couplings. Two TOCSY spectra were measured with spin lock times of 20 and 100 ms. The ROESY spectrum was measured with a mixing time of 500 ms. The HMBC, ROESY and one of the TOCSY spectra were measured with solvent presaturation.

Results and Discussion

The proton spectrum of the sample itself (figure 1) was not amenable to the structural analysis, but was rich in information about the purity of the sample and the size and complexness of the molecule. Almost no peaks were directly identifiable because of signal overlap in the whole spectral range. Weak multiplets (peak intensities about

10 % of the peaks of the major component) with similar chemical shifts and multiplet structures as the dominant peaks were present in the spectrum indicating the presence of similar structures or isomers in the sample. The spectrum was very crowded in the aliphatic region with several singlets in the upfield region. The presence of a larger number of multiplets in the region of

3–4 ppm was a good indication of the presence of saccharide units.

The carbon spectrum (figure 2) carried the information about the size of the molecule in two ways – the number of signals was high and there was only a slight NOE enhancement of protonated carbon signals. This is caused by the distribution of inner motions of the molecule – precisely by the

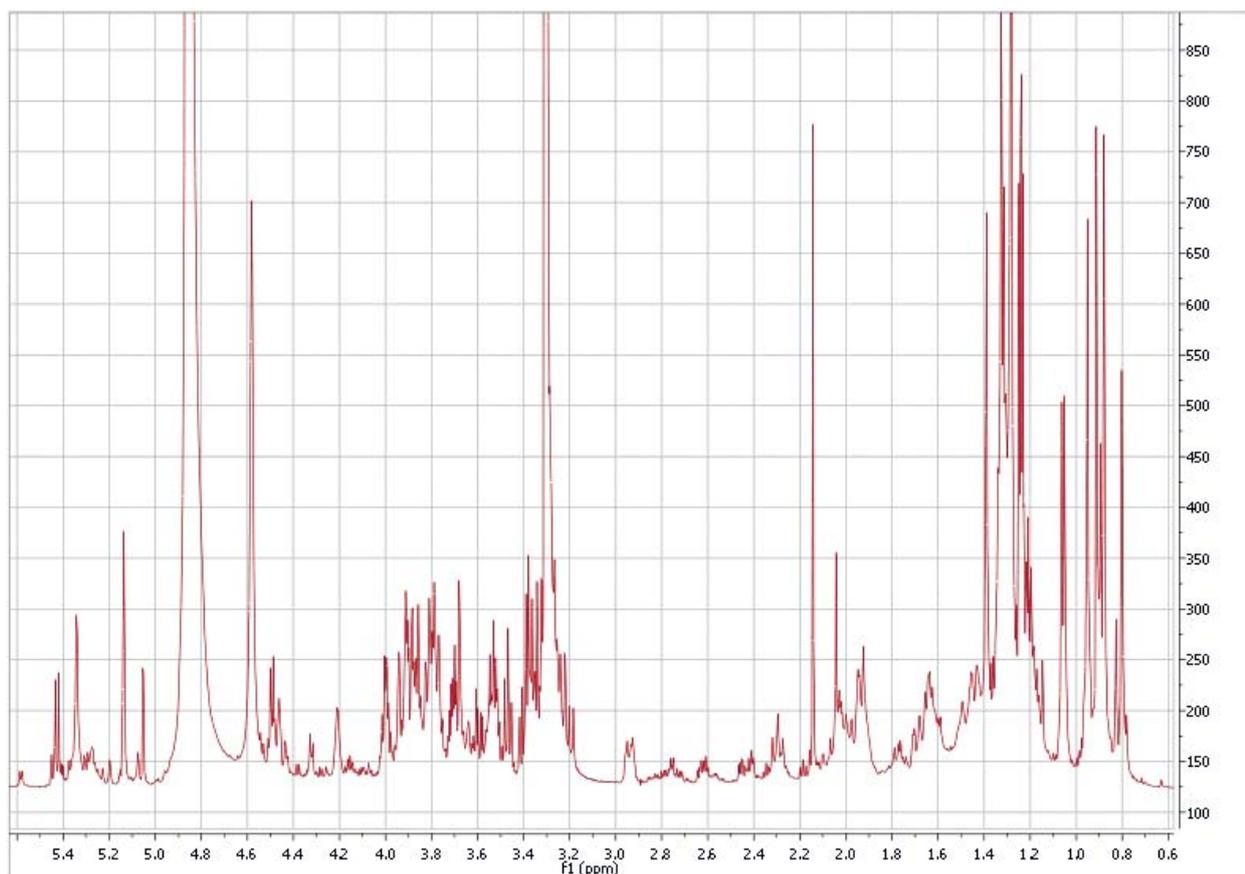


Fig. 1. The proton spectrum of the analyzed sample. The broad signals at 4.8, 4.6 and 3.3 belong to the solvent resonances.

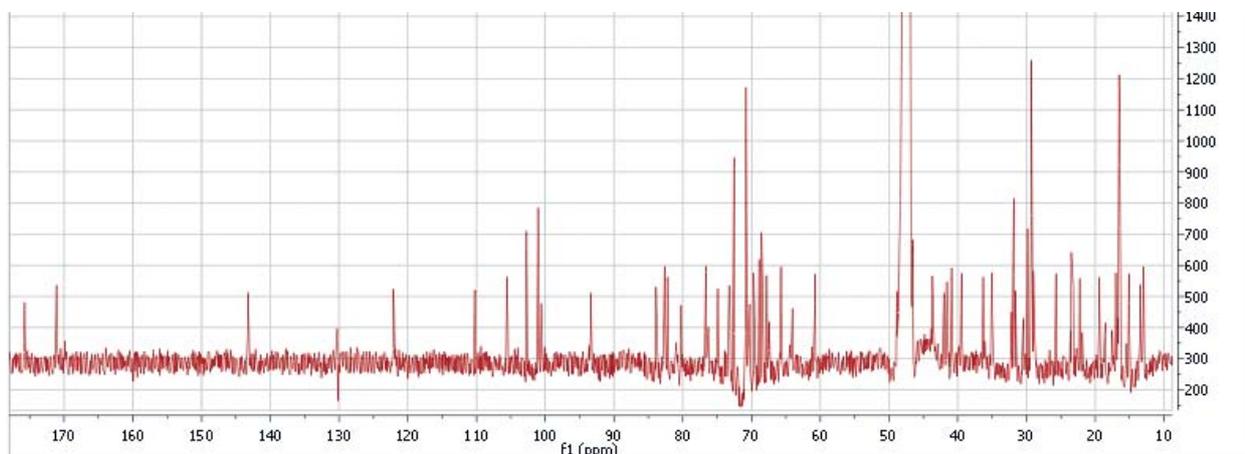


Fig. 2. The carbon spectrum of the sample. Most of the signals are well separated. The increased intensity of some peaks (around 70 ppm and 30 ppm) is caused mostly by signal overlap. Overlapping even in the carbon spectrum indicates the presence of repeating structural fragments.

lack of motions at the cross-relaxation frequencies, because of the high mass of the molecule.

The HSQC spectrum was one of the key steps in the analysis. It was possible to compare the chemical shifts of some one-bond C—H pairs and quarternary carbons with available literature data (Yoshikawa et al. 2008; Morikawa et al. 2008; Li et al. 2005; Schöpke et al. 1991) and get a basic idea about the possible structure of the molecule. Six methyl group resonances at $\delta_{\text{H}} = \{0.80, 0.88, 0.91, 0.95, 1.33 \text{ and } 1.39 \text{ ppm}\}$ and $\delta_{\text{C}} = \{16.5, 31.9, 13.4, 23.6, 17.0 \text{ and } 25.7 \text{ ppm}\}$, an olefin resonances at $\delta_{\text{H}} = 5.34 \text{ ppm}$ and $\delta_{\text{C}} = 122.1 \text{ ppm}$ and the presence of 8 CH₂ groups in the aliphatic region indicated, that the molecule is probably a substituted saponin.

Usually, the next step in the structure elucidation would be the identification of spin systems, but the signal overlap in the aliphatic region and the type of the structure (very short spin systems separated by quarternary carbons) rendered this step impossible. Therefore a different approach was used – to examine a larger number of possible backbones

and find out, which of them (if any) is in agreement with our structure. Some of the chosen structures differing with only a few groups are depicted on figure 3.

The double bond present in all structures was chosen as a well defined starting point for structural analysis. The backbone elucidation was then done by combining all 2D methods. It consisted of the identification of a coupled spin system, then “jumping” to the next spin system, identifying it, etc. The methyl groups were distinguished mostly from the HMBC spectra. An example of the identification of two methyl groups is shown on figures 4 and 5.

The next step would be identifying the methyl group 25 in the same fashion – finding which singlet in the proton dimension shares the correlation with carbon 9. After repeating several similar steps the backbone structure of the aglycone part of the molecule was eventually identified and it was found, that it is identical with the structure shown in the figure 3c. The next problem was the identification of the substituents attached to the aglycone part of the molecule. Several correlation “networks”

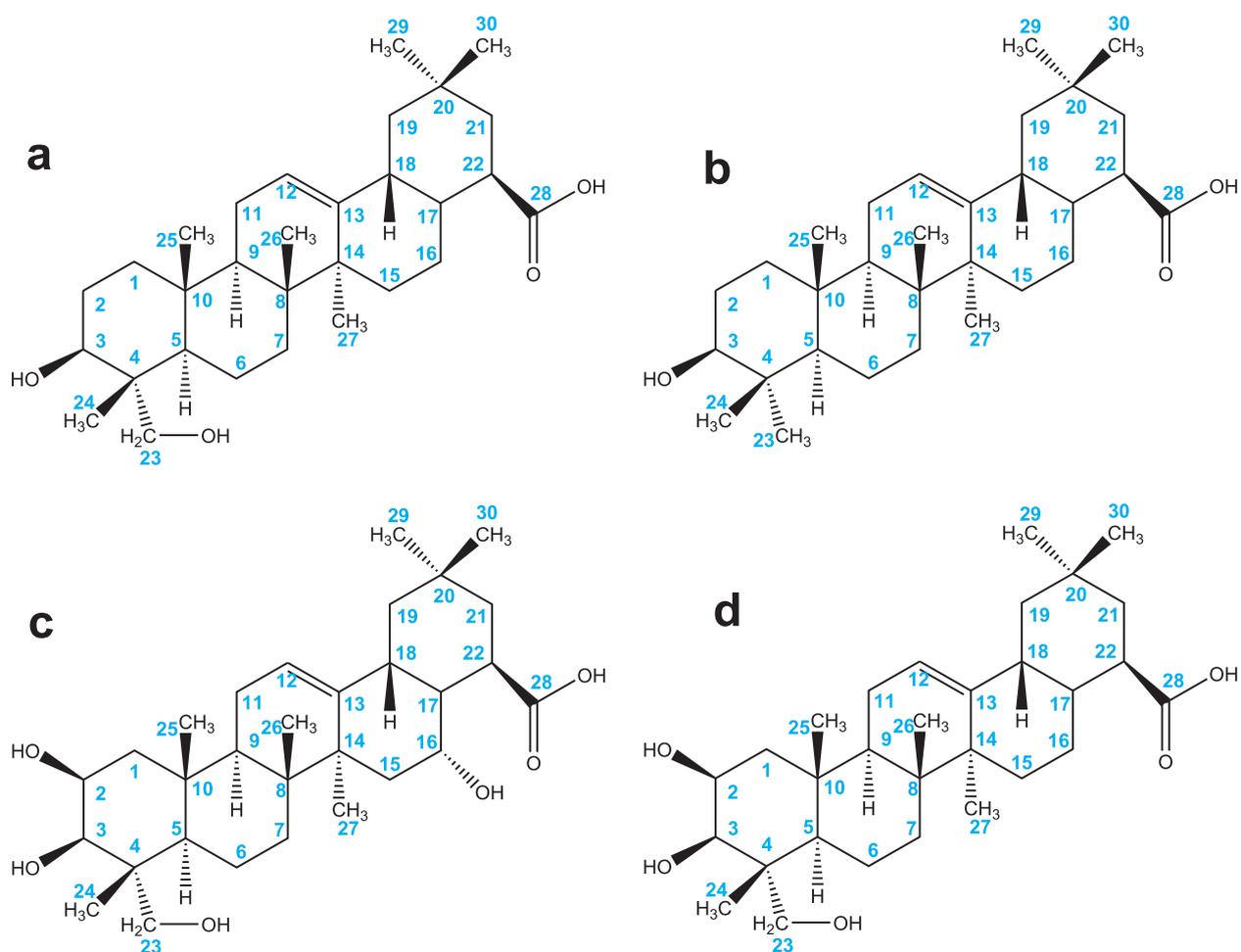


Fig. 3. Structural backbones as candidates for the analyzed structure selected after the analysis of the HSQC spectrum.

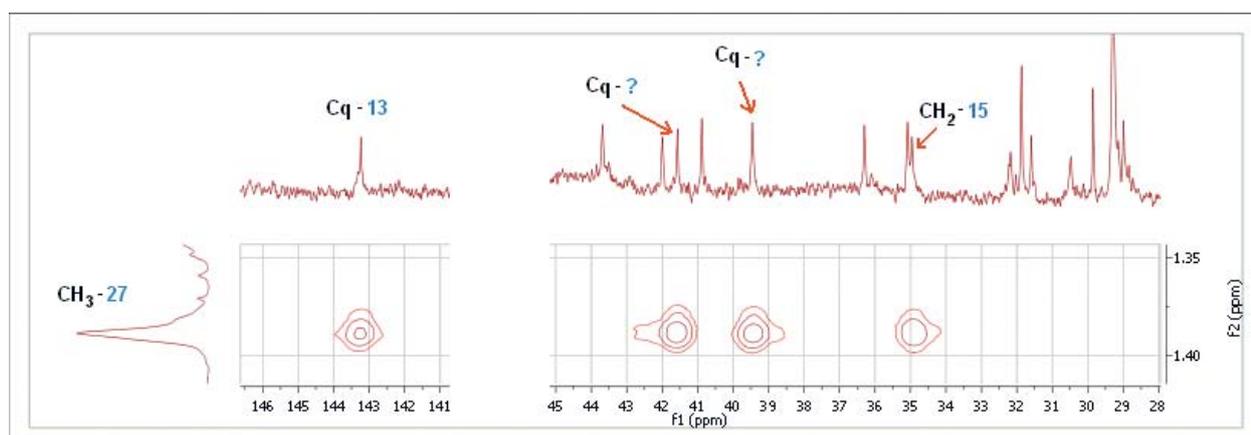


Fig. 4. A slice of the HMBC spectrum capturing the long range heteronuclear correlations of the protons of the methyl group 27 H/C δ 1.39/25.7. It was the only methyl group capable of having a three bond correlation with the quarternary carbon of the double bond. Since the quarternary carbons 8 and 14 are indistinguishable from the HMBC, the carbons at 41.6 and 39.5 ppm are marked by a question mark at this moment.

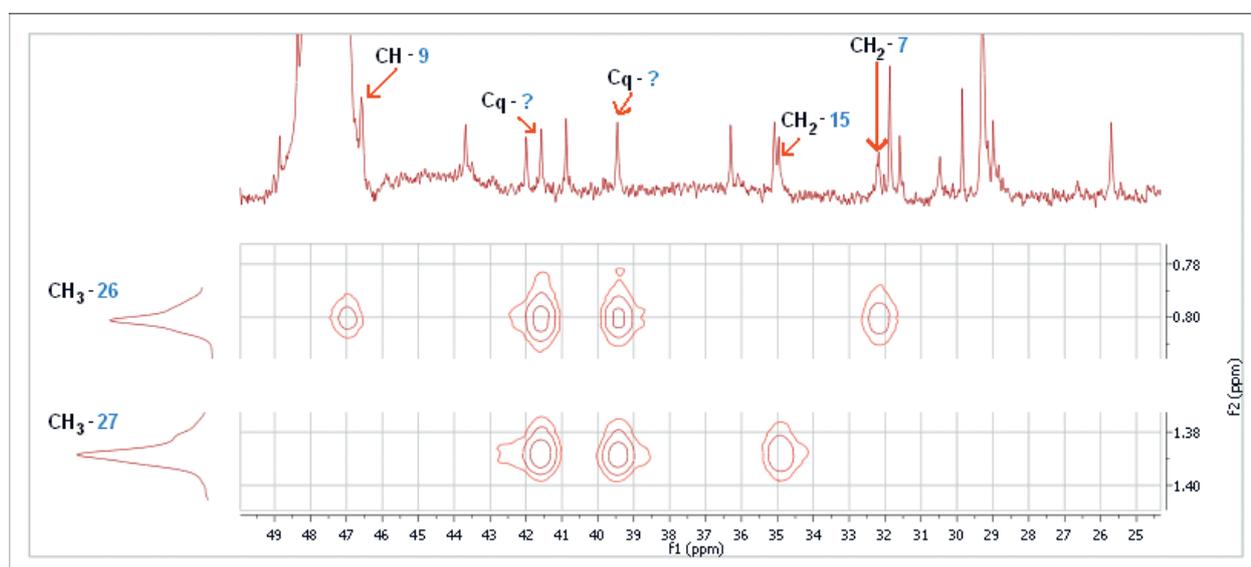


Fig. 5. Comparison of two HMBC slices – the methyl group 27 (mentioned in figure 4) can “share” the same correlations with the quarternary carbons only with the methyl group on the position 26 H/C δ 0.80/16.5.

were visible in the TOCSY spectra. The correlation multiplets of the cross-peaks were resolved good enough to determine the coupling constants. For the sugar units in their pyranose form, these are usually sufficient to identify monosaccharide units unambiguously. An example of a TOCSY network is shown on figure 6.

In the TOCSY spectrum it was possible to identify the whole sugar spin system even for the sugar unit with its anomeric proton completely overlapped by the strong signal of water, which was suppressed by presaturation. It was possible because of the coherence transfer from the neighboring protons to this anomeric proton).

Using the information from the TOCSY spectra straightforward and unambiguous identification of 5 sugar units in the pyranosyl form was possible. However the identification of the last sixth was trickier as the unit exhibited ^1H — ^1H coupling constants, which were difficult to interpret, and unusually high proton and carbon chemical shifts and a CH_2 group with only small differences of proton chemical shifts. These data suggested that this unit is a pentose sugar in furanose form. Our conclusion is in accord with the literature data (Mizutani et al. 1989).

The remaining step in the structural analysis was to connect the monosaccharide units and the aglycone. The connections were mostly found by the HMBC

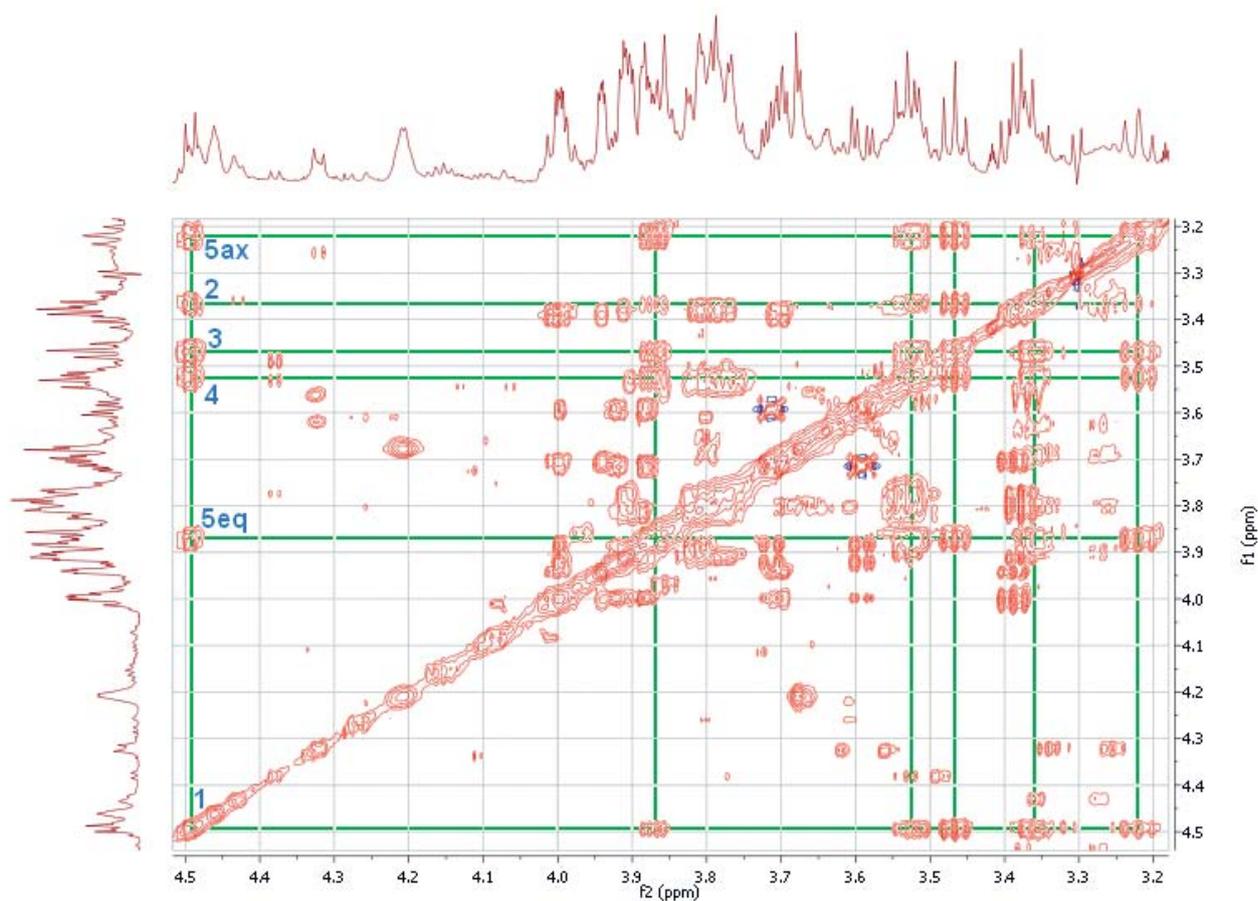


Fig. 6. A highlighted TOCSY network of a xylopyranose unit. The whole network is visible because of an efficient coherence transfer among vicinal protons thanks to their strong axial-axial interactions in the chair conformation. The indexing of 1–5 starts from the anomeric proton and ending with the CH₂ group incorporated inside the pyranose ring (5ax in the axial position and 5eq in the equatorial position).

spectrum. Mutual 3-bond ¹H—¹³C correlations were observed through O-glycosidic bonds. The most important correlations are depicted in the figure 7. The NMR identification was later on supported by a MS-MS fragmentation analysis (J. Marák, unpublished).

Table 1 summarizes the chemical shifts of protons and carbons in the molecule. The labeling is the same as in the figure 7.

Conclusion

Thanks to the applied NMR instrumentation (especially the cold probe) and the methods it was possible to identify the structure of the molecule in the sample even if the sample was low concentrated, impure and we had almost no prior knowledge before the identification happened.

Acknowledgement

The authors wish to thank the State Program 2003SP200280203 and VEGA 1/0972/12 for the support.

References

- Avato P, Tava A (1995) *Phytochemistry* 40: 141–147.
- Chevallier A (1996) *The Encyclopedia of Medicinal Plants*. Dorling Kindersley, London.
- Dobelis IN (1989) *Magic and Medicine of Plants*. Reader's Digest, New York.
- Grieve M (1982) *A Modern Herbal*. Dover Publications, New York.
- Gudej J, Nazaruk J (2001) *Fitoterapia* 72: 839–840.
- Li W, Asada Y, Koike K, Nikaido T, Furuya T, Yoshikawa T (2005) *Tetrahedron* 61: 2921–2929.
- Mizutani K, Kasai R, Nakamura M, Tanaka O (1989) *Carbohydrate Research* 185: 27–38.
- Morikawa T, Li X, Nishida E, Ito Y, Matsuda H, Nakamura S, Muraoka O, Yoshikawa M (2008) *Journal of Natural Products* 71: 828–835.
- Schöpke T, Wray V, Rzazewska B, Hiller K (1991) *Phytochemistry* 30: 627–631.
- Sohretoglu D., Karakas FP, Stujber M, Turker AU, Calis I, Yalcin FN, Liptaj T (2012) *Biomedical Papers* 156: 69.
- Vincken JP, Heng L, de Groot A, Gruppen H (2007) *Phytochemistry* 68: 275–297.
- Yoshikawa M, Li X, Nishida E, Nakamura S, Matsuda H, Muraoka O, Morikawa T (2008) *Chemical & Pharmaceutical Bulletin* 56: 559–568.

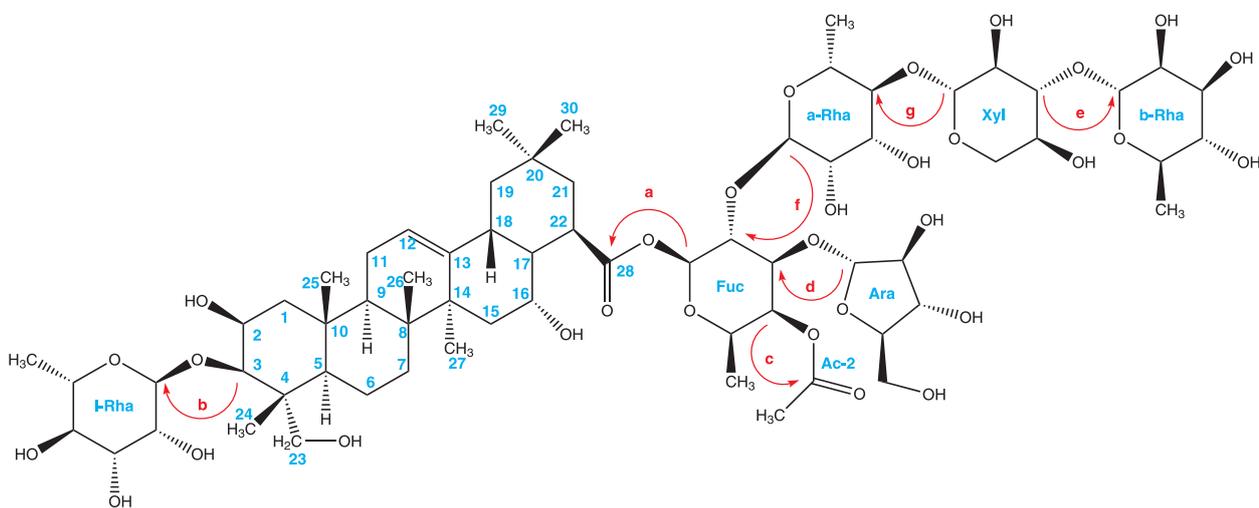
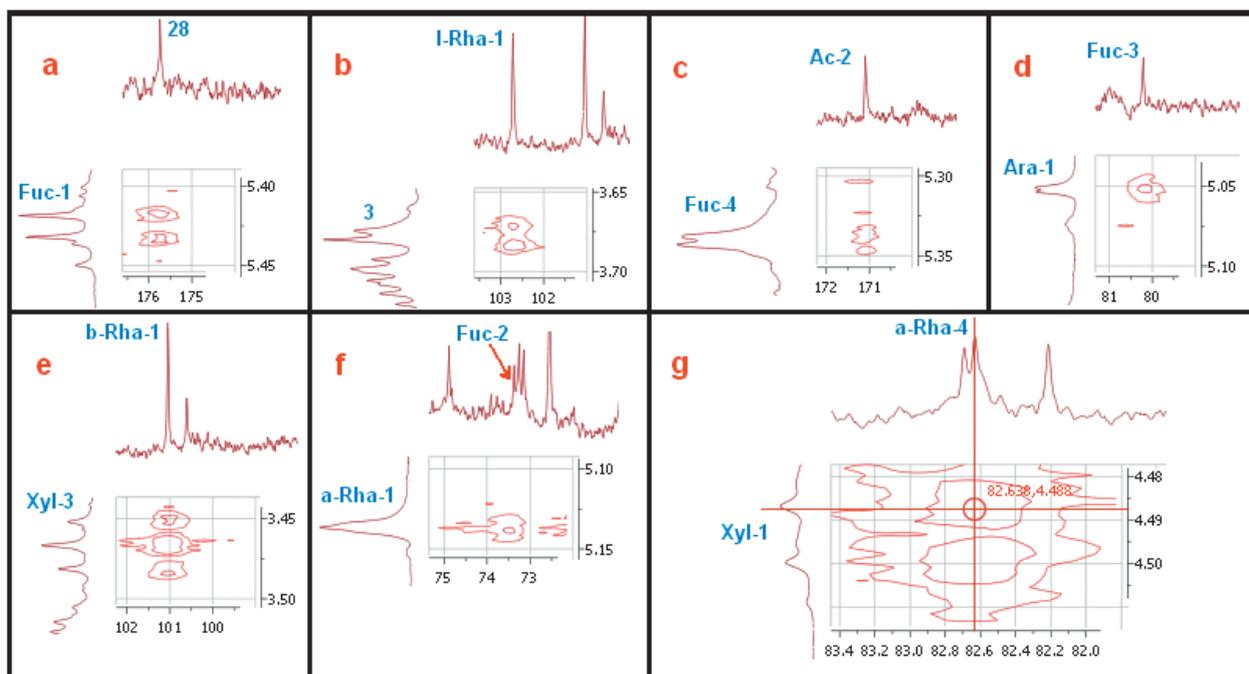


Fig. 7. Important HMBC correlations used for the structure determination.

Tab. 1. The proton and carbon chemical shifts of the identified molecule.

Position	¹ H	¹³ C	Position	¹ H	¹³ C
1	1.20 dd, 2.8; 13.5 2.03 m	43.7	l-Rha-1	4.86, br.s	102.7
2	4.21 m	70.4	l-Rha-2	3.91 m	70.9
3	3.68 d, 3.5	81.1	l-Rha-3	3.78 dd, 9.5; 2.7	70.7
4	-	42.0	l-Rha-4	3.38 t, 9.5	72.5
5	1.33 d, 10.6	46.6	l-Rha-5	3.81 m	68.9
6	1.44 m 1.50 m	17.6	l-Rha-6	1.24 d, 6.1	16.5
7	1.41 d, 11.0 1.64 dd, 8.4; 11.0	32.2	Xyl-1	4.49 d, 7.6	105.5
8	-	39.5	Xyl-2	3.36 br.t, 9.0; 7.6	74.9
9	1.64 dd, 6.0; 11.0	47.0	Xyl-3	3.47 br.t, 9.0; 9.4	82.7
10	-	36.3	Xyl-4	3.53 br.dt, 9.2; 10.0; 5.3	68.4
11	1.93 m 2.00 m	23.2	Xyl-5ax	3.22 br.t, 11.6; 10.0	65.7
12	5.34 d, 2.6	122.1	Xyl-5eq	3.87 dd, 11.6; 5.3	
13	-	143.2	a-Rha-1	5.14 d, 1.1	100.6
14	-	41.6	a-Rha-2	3.90 m	70.3
15	1.45 dd, 15.0; 2.7 1.69 dd, 15.0; 3.7	35.0	a-Rha-3	3.80 dd, 9.3; 2.8	70.9
16	4.46 m	73.2	a-Rha-4	3.53 br.t, 10.1; 9.3	82.6
17	-	48.9	a-Rha-5	3.77 m	67.8
18	2.94 dd, 13.8; 3.4	40.9	a-Rha-6	1.32 d; 6.1	17.0
19	1.06 dd, 12.0; 3.4 2.30 br.t, 13.8; 12.0	46.6	Ara-1	5.05 d, 1.6	110.2
20	-	29.9	Ara-2	4.00 dd, 3.8; 1.6	82.2
21	1.17 m 1.92 m	35.1	Ara-3	3.88 dd, 6.2; 3.8	76.7
22	1.77 m 1.91 m	30.5	Ara-4	3.92 m	83.9
23	3.27 d, 11.1 3.33 d, 11.1	64.3	Ara-5_1	3.59 dd, 12.2; 4.1	60.8
24	0.91 s	13.4	Ara-5_2	3.71 dd, 12.2; 3.1	
25	1.33 s	17.0	b-Rha-1	5.14 d, 1.0	101.1
26	0.80 s	16.5	b-Rha-2	3.94 dd, 3.0; 1.0	70.8
27	1.39 s	25.7	b-Rha-3	3.70 dd, 9.3; 3.0	70.7
28	-	175.8	b-Rha-4	3.39 br.t, 9.9; 9.3	72.5
29	0.88 s	31.9	b-Rha-5	4.00 m	68.5
30	0.95 s	23.6	b-Rha-6	1.24 d; 6.5	15.1
Fuc-1	5.43 d, 8.0	93.4	Ac-1	2.14 s	19.4
Fuc-2	3.81 br.t, 9.7; 8.0	73.4	Ac-2	-	171.1
Fuc-3	3.89 dd, 9.7; 3.4	80.2			
Fuc-4	5.34 d, 3.4	73.3			
Fuc-5	3.85 m	69.7			
Fuc-6	1.06 d, 6.4	15.1			