Features in the NMR spectra of the aglycones of Agave spp. saponins. HMBC method for aglycone identification (HMAI)

SCHOLARONE™
Manuscripts

1   **Features in the NMR spectra of the aglycones of *Agave spp.* saponins.**

2   **HMBC method for aglycone identification (HMAI)**

3   **Short title:** HMBC Method for aglycone identification (HMAI) of *Agave spp.* saponins

4

5   Ana M. Simonet[1], Alexandra G. Durán[1], Andy J. Pérez[2], and Francisco A. Macías[1]

6

7

8   [1] Allelopathy Group, Department of Organic Chemistry, Institute of Biomolecules

9   (INBIO), Campus de Excelencia Internacional (ceiA3), School of Science, University of

10   Cadiz, C/ República Saharaui nº 7, 11510 Puerto Real, Cadiz, Spain.

11   [2] Departamento de Análisis Instrumental, Facultad de Farmacia, Universidad de

12   Concepción, Concepción 4191996, Chile.

13

14   **Correspondence**

15   Ana M. Simonet and Francisco A. Macías, Department of Organic Chemistry, School of

16   Science, University of Cadiz, C/ República Saharaui nº 7, 11510 Puerto Real, Cadiz,

17   Spain. Phone: +34 956012770; fax: +34 956016193; e-mail: ana.simonet@uca.es;

18   famacias@uca.es

19

23

25 **Abstract**

26

27 **Introduction:**

28 The analysis and detection of steroidal saponins is mainly performed using

29 chromatographic techniques coupled with Mass Spectrometry. However, Nuclear

30 Magnetic Resonance (NMR) spectroscopy is a potential tool that can be combined with

31 these techniques to obtain an unambiguous structural characterization.

32 **Objective:**

33 This work provides a review of the ^{13}C and ^{1}H NMR spectroscopic data of aglycones

34 from *Agave* saponins reported in the literature and also the development of an easy

35 identification method for these natural products.

36 **Methods:**

37 The database Scifinder was used for spectroscopic data collection in addition to data

38 obtained from the Cadiz Allelopathy research group. The keywords used were *Agave*,

39 spirostanic, furostanic, and saponin.

40 **Results:**

41 The shielding variations produced by functional groups on the aglycone core and the

42 structural features of the most representative aglycones from *Agave* species are

43 described. The effects are additive for up to four long-range connectivities. A method

44 for the identification of aglycones (HMAI) is proposed to classify aglycones from

45 *Agave* spp. through the use of ^{1}H NMR and HMBC experiments.

46 **Conclusions:**

47 The HMBC spectrum is representative of the structural features of aglycones from

48 *Agave* spp. The HMAI method allowed the identification of pure saponins or mixtures

49 thereof and this method can be used in combination with chromatographic techniques

50   coupled with Mass Spectrometry to provide a more thorough analysis of *Agave* samples

51   that contain aglycones.

52   **Short abstract**

53   ^{1}H and ^{13}C shielding variations produced by functional groups on the aglycone core and

54   the structural features of the most representative aglycones from *Agave* species are

55   described. The effects are additive for up to four long-range connectivities. A method

56   for the identification of aglycones (HMAI) is proposed to classify aglycones from

57   *Agave* spp. through the use of ^{1}H NMR and HMBC experiments. This method allowed

58   the identification of pure saponins or mixtures thereof and it could be used in

59   combination with chromatographic techniques coupled with Mass Spectrometry to

60   provide a more thorough analysis of *Agave* samples that contain aglycones.

61

62   **Keywords:** saponin, aglycone, *Agave*, HMBC, HMAI, ^{1}H NMR, ^{13}C NMR,

63   identification

65

1.   INTRODUCTION

1.1. Saponins. Definition and biological activities

Saponins are secondary metabolites that are glycosidic in character and have specific natural properties. Despite their well-known biological activities, the specific role and mechanism of action of saponins are not fully established. It has been suggested that saponins could play a significant defensive role against microorganisms, because various fungi produce saponin-detoxifying enzymes,[1] and against mammals or insect herbivores as an antifeedant.[2] Moreover, physiological effects associated with plant growth regulation and development have been reported as possible functions for saponins.[3]

These natural products, which are found in a vast variety of plant species and in some marine organisms, consist of a hydrophobic triterpene or sterol backbone and a hydrophilic carbohydrate chain formed by monosaccharide units, with the structural features linked together by a glycosidic bond (Figure 1). Based on the structures of the aglycone skeletons, saponins can be divided into two main groups, namely steroidal and triterpenoid saponins. Triterpenoid saponins are mostly found in dicotyledonous species whereas monocots mainly produce steroidal saponins.[2,4] Steroidal saponins have a hydrophobic nucleus or sapogenin constituted by 27 carbon atoms.

The amphipathic nature of saponins means that they can act as surfactants and in most cases stable they give soap-like foams in aqueous solutions and they have been used as natural soaps and detergents since ancient times.[5] The aforementioned properties allowed these compounds to rupture erythrocytes and cause irreversible damage to the membrane lipid bilayer. This hemolytic activity is one of the first effects reported and is

4

89   mode of action that is most widely accepted for the biological activities shown by

90   saponins.[6] However, in a study performed by Wang and co-workers the correlation

91   between hemolytic and cytotoxic activities of a collection of steroidal saponins was

92   evaluated. The results indicated that cytotoxic activity does not always relate with

93   hemolytic activity, thus suggesting that steroidal saponins execute the two activities in

94   different mechanisms.[7]

95   It has been demonstrated that the biological activities of saponins are dependent on their

96   structures. For example, a thorough study of 28 sapogenins and spirostane-type

97   saponins against pathogenic fungi showed that those saponins with less oxygenation in

98   the steroidal core and a sugar moiety of four or five monosaccharide units exhibited

99   significant activity.[8] The phytotoxicity of these secondary metabolites has also been

100   tested. The effects of 28 steroidal saponins on the standard target species *Lactuca sativa*

101   were evaluated. Strong root growth inhibition was noted for those saponins with four or

102   more sugar units in the saccharide chain and oxygenation, especially at the C-12

103   position of the aglycone skeleton.[9–11] Moreover, in other studies it has been

104   demonstrated that the activity is highly dependent on the monosaccharide features.[12]

105   Likewise, other structure-activity relationship studies (SARs) of steroidal saponins on

106   HL-60 (human promyelocytic leukemia) cells showed that cytotoxicity was dependent

107   on the aglycone backbone and also the sugar moieties and their sequences.[13] These

108   results indicated that the cytotoxic effects shown by these saponins could be due to non-

109   specific detergent effects with changes in membrane architecture. Nonetheless, the level

110   of damage is considerably different among the saponins tested, and two of them with

111   two sugar residues in the carbohydrate chain caused cell death through an apoptotic

112   process. Other saponins with these structural features have shown other mechanisms of

113   action.[14,15]

114   All of the findings outlined above support the idea that mechanisms other than

115   membrane damage are also involved.

116   The genus *Agave* is one of the main sources of steroidal saponins. The sugar moieties

117   present in saponins from this genus are *β*-D-glucopyranosyl, *β*-D-galactopyranosyl, *β*-

118   D-xylopyranosyl and *α*-L-rhamnopyranosyl units.[16] On the one hand, taking into

119   account the sapogenin backbone, saponins of *Agave* can be classified as spirostanol

120   glycosides and furostanol glycosides. On the other hand, monodesmosidic saponins are

121   those in which the sugar chain is present at only one position (generally at C-3) of the

122   sapogenin, while bidesmosidic saponins have two sugar units located at two different

123   points of the aglycone core. Most of the bidesmosidic furostanol saponins are

124   glycosylated at the C-3 and C-26 positions, with a *β*-D-glucopyranoside usually present

125   in the latter position.

126   **1.2. Drawbacks in the isolation and elucidation of saponins. Analysis of**

127   **mixtures.**

128   There are several methods to obtain saponin-rich extracts, including conventional

129   (maceration, Soxhlet and reflux extraction) and green (ultrasound-assisted, microwave-

130   assisted and accelerated solvent extraction) techniques.[17] Crude extracts are commonly

131   mixtures of saponins with a wide range of polarities and structural diversity, differing

132   even between plant organs. Steroidal saponins with up to seven sugar units have been

133   reported.[18] Thus, the isolation of saponins remains a challenge and requires the use of

134   various separation techniques and different adsorbents to achieve, in most cases, the

135   isolation of the major saponins.

136   *Agaves* have been widely used for their high carbohydrate content to obtain *Agave* sap,

137   sweeteners and, after fermentation, alcoholic beverages such as pulque, mescal or

138   tequila. Chromatographic techniques coupled with mass spectrometry have been used to

139   determine the saponins content and biochemical changes that occur during beverage

140   processing. Thus, for instance, this methodology has allowed the quantification and

141   identification of saponins in wild and cultivated populations used for the production of

142   mescal and pulque,[19] after the use of micropropagation to allow mass production of

143   *Agaves*,[20] under *in vitro* drought stress,[21] their variation during plant ripening stage,[22,23]

144   in concentrated *Agave* sap produced in different states of Mexico,[24] or changes in the

145   saponins profile by microorganisms after *Agave* sap fermentation.[25]

146   Mass spectrometry is a valuable tool that provides structural information on saponins,

147   including the fragmentation pattern of the sugar chain. This technique allows the

148   assignment and quantification of the saponins previously isolated from the species being

149   analyzed. However, structural assignment of saponins from different species cannot be

150   achieved, since there are isomers with identical masses and fragmentation patterns.

151   Given the influence of the structure on the biological activity of saponins, it would be

152   beneficial to complete these studies with nuclear magnetic resonance (NMR)

153   spectroscopic techniques. These experiments provide unambiguous information on the

154   position and stereochemistry of the functional groups present in the aglycone, as well as

155   the nature and connectivity of the different sugars on the carbohydrate chain.

156   **1.3. Evolution of the structural elucidation and assignment procedures**

157   The first studies that addressed the structural elucidation of secondary metabolites from

158   the *Agave* genus were focused on sapogenins.[26] Physical properties such as melting

159   point determination and chemical transformations, including elemental analysis and

160   oxidation and reduction (redox) reactions, were the most common techniques used for

161   the determination of known compounds. For a few years, the detection of certain

162   sapogenins was performed by measuring the ultraviolet absorption maximum[27] in the

163   typical region for an α,β-unsaturated ketone system. This approach allowed the

164   identification of 9-dehydrospirostan-12-ones.

165   The first structural elucidation of saponins was carried out by the total hydrolysis of the

166   isolated saponins followed by determination of melting point, IR, MS and

167   chromatographic mobility to identify the corresponding aglycone.[28] Specific techniques

168   were subsequently applied to identify the monosaccharides. Interglycosidic linkages

169   were determined by acid hydrolysis and permethylation followed by acid hydrolysis of

170   these units to obtain the different protosaponins. The identification of monosaccharides

171   was achieved by comparison with known samples. In some cases, these methods were

172   combined with mass spectrometry or FAB-MS to ascertain the sequence of sugars.[29]

173   The presence of furostane-type saponins was determined by thin layer chromatography

174   using Ehrlich's reagent for a preliminary examination and subsequent conversion to

175   their spirostanic derivatives through specific enzymes facilitated their further

176   elucidation.[30]

177   In the early years, NMR techniques were used for the identification of the aglycone

178   moiety after a hydrolysis reaction.[31] ^{13}C NMR experiments proved to be very useful for

179   the determination of less complex saponins. Thus, glycosylation shift rules began to be

180   applied to determine characteristic signal shifts (downfield or upfield) at the α- and β-

181   positions with respect to the –OH groups that were glycosylated.[32,33] Sugar chains with

182   up to six units were elucidated using the fragmentation patterns observed by mass

183   spectrometry and by comparison with ^{13}C NMR data reported in the literature.[34]

184   The characterization of these secondary metabolites is not an easy task but the advances

185   in NMR technology have provided a non-destructive way to achieve this

8

186 characterization. The emergence of two-dimensional NMR experiments allowed the

187 elucidation and signal assignment to be carried out in a more thorough and reliable

188 way.[35] One of the main advantages is that an unambiguous characterization of complex

189 sugar moieties can be achieved. Sugar linkage analysis and spatial arrangements can be

190 determined by HMBC and 1D or 2D NOESY/ROESY experiments.[36]

191 These two-dimensional NMR experiments have been used since around 2000 for the

192 determination of saponins and this technique requires a lower amount of pure

193 compound. Moreover, the development of higher field NMR spectrometers (at least 500

194 MHz) has made the complete assignments of the ^{1}H NMR spectra feasible, especially in

195 overlapping zones.[37]

196 As mentioned previously, these metabolites have potentially attractive biological

197 activities. Given that there is a high structure-activity correlation, including the

198 oxygenation pattern of the aglycone and nature of the sugar moiety, it is crucial to

199 achieve the most unambiguous determination possible.

200 Furthermore, because of the amphipathic nature of these compounds and the presence of

201 structurally related forms with very similar polarities, their separation can be tricky and

202 LC-MS and NMR techniques have been used to screen saponins by hyphenated

203 analytical platforms[38] or metabolomics.[39]

204 **2. SALIENT AND COMMON FEATURES IN THE NMR SPECTRA OF**

205 **SAPONIN AGLYCONES FROM *AGAVE* SPP.**

206 A systematic compilation of the ^{13}C NMR chemical shifts for steroidal-type saponins

207 dates from the 1980s.[40] Reviews covering the most characteristic ^{13}C signals to

208 determine the ring fusion and spirostane/furostane skeleton have been published.[35]

209   Likewise, taking into account the ^{1}H NMR chemical shift patterns of methylene C-26 of

210   a range of saponins, the stereochemistry of C-25 could be ascertained.[41–43]

211   Recently, a comprehensive review of the structural features of saponins from *Agave*

212   species was published by Sidana and co-workers.[16] The present review will focus on the

213   influence of the main structural characteristics of ^{1}H and ^{13}C chemical shifts and will

214   also provide an exhaustive overview of patterns and signals that could be key clues to

215   identify the aglycone. The significant signals due to either their sensitivity to

216   neighboring functional groups or their fast and feasible detection in one- and two-

217   dimensional spectra will also be discussed.

218   It is unusual to find the full assignments of the ^{1}H NMR data for steroidal saponins.

219   Nonetheless, the chemical shifts of saponins with the main aglycone structural

220   characteristics and the most common sugar chains have been selected (Tables 1 and 2).

221   All NMR data were obtained from samples in deuterated pyridine in order to avoid the

222   influence of the solvent on the chemical shift. This solvent is the most widely used for

223   this kind of compound and it is able to dissolve saponins with a variable range of

224   solubilities.[35]

225   [1]H and ^{13}C spectroscopic data were adjusted by comparing the chemical shifts described

226   at positions 22 and 27 using TMS as internal reference. Thus, for ^{1}H NMR values, the

227   setting is from 0 to –0.4 ppm and for ^{13}C NMR data from 0.4 to –0.3 ppm (see

228   supporting information).

229   The selected data show the consistency in the assignments as the observed error range is

230   ± 0.4 ppm for ^{13}C NMR and ± 0.1 ppm for ^{1}H NMR in methylenes and ± 0.04 ppm in

231   methyl and methine groups. The error range has been distinguished on the basis of the

232   types of hydrogens present in the structure because the methylene assignments are less

10

233　accurate. In this case, the determination was indirect through the use of two-dimensional

234　experiments and because the rings A and B methylenes are significatively influenced by

235　the nature of the sugar chain.

236　Axial and equatorial orientations in the methylene groups are not usually described and

237　therefore they are defined as 'a' and 'b' in increasing order of chemical shift. In the case

238　of assignment, the axial positions are generally more shielded than the equatorial ones.

239　The spatial arrangement of ring A/B is described in the discussion section for those

240　cases where it has been reported in the bibliography and is of interest.

241　Steroidal-type saponins are the most widely represented structures belonging to *Agave*

242　species.[16] On the one hand, the structural features included are a spirostane or furostane

243　backbone and chiral centers at C-5 and C-25. On the other hand, the following

244　functional groups are also considered, double bonds at C-5, C-9(11) and C-25(27),

245　hydroxyl groups either in an α disposition at C-2, C-6 and C-23 or β disposition at C-2,

246　C-12 and C-24, and finally the presence or absence of a carbonyl group at C-12.

247　The stereochemistry of hydroxylated positions is the same in most cases. Epimers at

248　these positions are unusual in the *Agave* genus and the revision of some structures is

249　needed. For example, it was reported that the epimer with a hydroxyl group in a β

250　orientation at C-23 from *Agave fourcroydes*[44] had also been described from *A.*

251　*cantala*.[32] Nevertheless, it was confirmed that this saponin had the hydroxyl group in an

252　*α* orientation rather than *β*. Moreover, the epimer with a hydroxyl group in the β

253　orientation at C-6 was found in *A. cantala* and its structural elucidation showed that the

254　aglycone was chlorogenin. However, the hydroxyl group in this saponin is in an α

255　orientation. These facts confirm that some structures need to be reviewed.

256   Saponins that are representative of those with specific structural characteristics in the

257   aglycone core have been selected to discuss the most relevant and diagnostic NMR

258   spectroscopic features. The chemical shifts are within the error range determined as

259   valid for this review. Sugar moieties bonded at C-3 have direct influence at positions on

260   ring A to a greater or lesser extent. As a consequence, a greater range of error will be

261   allowed when describing the influence on the chemical shifts of that ring.

262   **2.1. Positions of ^{1}H and ^{13}C chemical shifts for ring F depending on the C-25**

263   **configuration**

264   Although the changes are not huge in the aglycone structure or backbone of steroidal

265   saponins reported from the *Agave* genus, some minor variations may considerably

266   hinder the correct elucidation of this structure. This is the case for the configuration at

267   C-25 of ring F, which could be defined not only as R or S but also as part of a double

268   bond with C-27 (Figure 2).

269   Configuration R is, however, the most common within this genus of plants. For this

270   reason, based on those saponins[45] (**1–3**) with tigogenin as the aglycone (**I**, Table 1), the

271   influence of the C-25 configuration changes over adjacent carbon positions could be

272   easily highlighted (Figure 2, Table 3). When C-25 has the S configuration, as in

273   compound **24**,[46] an upfield shift occurs for the carbon chemical shifts within ring F

274   (Figure 2), especially for C-25 and C-23, which are shifted upfield by 3.1 ppm and 5.2

275   ppm, respectively.

276   The ^{1}H NMR spectra also show chemical shift variations in ring F when C-25 is S. This

277   configuration leads to a downfield shift of around 0.37 ppm for the methyl group at C-

278   27, as well as greater separation of δ_H for pairs of geminal protons in methylenes at C-

279   23, 24, and 26, which is presumably caused by the axial orientation of methyl 27. Such

12

280   a separation has provided the basis for the well-known Agrawal's rule, which is

281   currently used to predict the configuration at C-25 with the $\Delta\delta_H$ of 2H-26.[35,43] It is worth

282   mentioning that despite a change in the C-25 configuration from R to S, the δ_H of H-25

283   remains unaltered. A slight influence has also been observed on the ^{1}H resonance for the

284   methyl at C-21 when C-25 is S, which usually appears 0.04 ppm upfield. While this

285   small variation could be within the error considered in this report, it has been regularly

286   observed when spectra were acquired under the same conditions.

287   Thirdly, when there is a double bond between C-25 and C-27 (**11**),[47] an obvious

288   downfield shift occurs for the now allylic protons at C-24 and C-26. The signals in [1]H

289   NMR spectra appear with chemical shift values higher than 2.2 ppm or 4 ppm for CH_2-

290   24 and 26, respectively, relative to the 25R derivative (**1**).

2.2. Substitution at C-23 and C-24 of the ring F

292   Methylene protons at positions C-23 and C-24 of ring F in spirostanic saponins from the

293   *Agave* genus are often substituted with hydroxyl groups. In both cases the hydroxyl

294   group is oriented equatorially, with the one at C-23 being a free hydroxyl group whereas

295   that at C-24 is usually glycosylated with a glucopyranose moiety (Figure 2).

296   Given that all of the saponins with such substitution patterns isolated to date from

297   *Agave* genus show a relative R configuration at C-25 (when C-24 is oxygenated the

298   absolute configuration of C-25 is S), a convenient model to compare and highlight

299   resonance changes is again tigogenin (**I**). When substituted, the δ_C values for the C-23

300   and C-24 now appear at 67.5 ppm (**58**) and 81.5 ppm (**62**), respectively,[48,49] in the [13]C

301   NMR spectra. It may be that both positions are substituted in a given compound and, in

302   this case, their δ_C are obviously shifted downfield, which is caused by a deshielding

303   effect of the neighboring electronegative oxygen (Table 3). In ^{1}H NMR spectra,

304   however, the methine proton chemical shifts are slightly modified when both C-23 and

305   C-24 are substituted. Similarly, the carbon chemical shift of C-22, a characteristic

306   spirostanic carbon, is usually shifted downfield by 2.4 ppm when only C-23 is

307   oxygenated, while this shift increases to 3.4 ppm when C-24 is also substituted. In the

308   situation where both C-23 and C-24 are oxygenated, the δ_C for C-26 (Table 3) is

309   affected in an additive way and this is clearly visible for compound **60**[50] when compared

310   with **58** and **62**.

311   A hydroxyl group at C-23 can also cause an upfield shift of the C-20 signal when

312   compared to tigogenin (**I**), which may be by 6.0 ppm or 7.3 ppm, as in compounds **58**

313   and **60**. The opposite effect is observed for δ_H of H-20 and H-17, the signals of which

314   are shifted downfield by 1.07 ppm and 0.08 ppm.

315   Despite the fact that the ^{1}H and ^{13}C resonance signals for methyl 21 are not greatly

316   affected by substitutions in ring F, the presence of a glycosidic linkage at C-24 causes

317   an upfield shift of its δ_H by 0.1 ppm. This effect, however, is not visible when

318   substitution occurs in both positions C-23 and C-24, with the hydroxyl group at C-23

319   being responsible in this case for the observed chemical shift variations for C-20 and C-

320   21. This double substitution, on the other hand, may also disturb the local magnetic

321   fields at C-25 and C-27 and therefore the carbon chemical shifts, with glycosylation at

322   C-24 being crucial. Instead, in the ^{1}H NMR spectrum the δ_H are additively shifted

323   downfield by 0.49 ppm (**60**) and 0.52 ppm (**60**) for methine 25 and methyl 27,

324   respectively.

2.3. Spirostane/Furostane C-22

326   The dioxygenated quaternary carbon at position 22 of aglycones could be found as a

327   ketal or a hemiketal structure. This carbon is the joint between rings E and F in

328   spirostanic saponins, while in furostanic compounds ring F is opened to yield a

329   hemiketal carbon at this position. Such a carbon could be found with its hydroxyl group

330   substituted by a methoxyl group, which has been reported to be a consequence of the

331   use of methanol during the purification process.[51] C-26 of furostanic saponins is in most

332   cases glycosylated with a glucopyranose.

333   Although the R configuration at C-25 is the most frequent in saponins isolated from the

334   *Agave* genus, saponins with an S configuration have also been reported.[16] On this basis,

335   we will now compare the most significant changes in chemical shifts from the ^{1}H and

336   ^{13}C NMR spectra between spirostanic (**I**) and furostanic (**V**) saponins, and between both

337   C-25 epimers R (**V**) and S (**VI**) of the latter. The methoxyl derivative is also included in

338   this discussion (**44**).[51]

339   Furostanic saponins present very different ^{1}H and ^{13}C NMR spectra than the spirostanic

340   compounds.[42] The ^{13}C chemical shift represents a significant difference when compared

341   to spirostanic saponins because the glycosylation causes a downfield shift by 8.6 ppm

342   for the C-26 signal (**43**) (Table 3).[51] Furthermore, the ^{13}C NMR signals of C-22 and C-

343   21 are shifted downfield by 1.4 ppm and 1.6 ppm, respectively, while the C-20 signal

344   moves upfield by 1.1 ppm.

345   Downfield shifts are generally observed in the ^{1}H NMR spectra of furostanic saponins

346   from the methine at C-17 to the methyl group at C-27. Indeed, the methyl groups at C-

347   21 and C-27 typically experience the most significant changes, with changes of 0.27

348   ppm and 0.29 ppm (**43**) (Table 3),[51] respectively. It is worth noting that opening of ring

349   F in furostanic saponins also induces a slight but consistent downfield shift of the ^{1}H

350   NMR signal for the methyl group at C-18 (by 0.05 ppm).

351   The C-25 epimers of furostanic saponins give rise to very similar ^{13}C NMR spectra.

352   However, in the ^{1}H NMR spectra there is a downfield shift of methyl-27 for 25R, and

353   the chemical shift separation between the geminal 2H-26 increases as described by

354   Agrawal's law.[41,42]

355   In the case where a methoxyl group is attached at C-22, a downfield shift by 2.0 ppm is

356   observed for this carbon signal (**44**), which is confirmed by the appearance of a new

357   signal at 47.2 ppm corresponding to the methoxyl group. The opposite effect is

358   observed for the signal of C-23, which is usually considerably shifted upfield by 6.5

359   ppm (**44**) (Table 3),[51] with a similar change also observed for methyl 27 (0.4 ppm),

360   albeit to a lesser extent. In the ^{1}H NMR spectrum a methoxyl signal is observed at 3.24

361   ppm and signals for H-16 to C-24 are mostly shifted upfield with respect to the

362   hydroxyl derivative (**43**), with changes of 0.08 and 0.14 ppm for the methyl signals H-

363   18 and H-21.

2.4. Oxygenation at C-12

365   Several saponins described from the *Agave* genus have a carbonyl group at C-12 (Table

366   1), while in some other cases this position has been found to be substituted with a β-

367   hydroxyl group instead. Changes induced by these functionalizations in the neighboring

368   nuclei are discussed below (Figure 3) and these may serve as diagnostic signals for

369   structure elucidation.

370   We have noted that a carbonyl group at C-12 of the aglycones may in some way affect

371   the chemical shifts for all other positions – except for the A and F rings, including when

372   the latter is opened as in furostanic saponins. As one would expect, the α-carbons with

373   respect to the carbonyl group are the most deshielded (**18**) (Table 4).[48] The carbon

374   located at a distance of two bonds, namely C-9, is also deshielded but to a lesser extent.

375   An exclusive feature of the *Agave* saponins with a carbonyl group at C-12 is a C-17

376   signal that is shielded by 8.6 ppm in the ^{13}C NMR spectra. This signal can be diagnostic

377   of the presence of a carbonyl group in the molecule (Table 4). Methyl groups at C-21,

378   C-19 and C-18 are also shielded and the signals move upfield by 1.0 ppm, 0.5 ppm, and

379   0.4 ppm, respectively.

380   The carbonyl group at C-12 also causes a deshielding effect in the ^{1}H NMR spectra.

381   This is the case for the signals of the methylene group at C-11, which are shifted

382   downfield by around 1.0 ppm (Table 3). The four methine protons H-8, H-9, H-14 and

383   H-17 are also shifted downfield (Figure 3) – especially the latter proton, which is

384   deshielded by 0.95 ppm (**18**) (Table 4).[48] Such an effect is also apparent for δ_H of the

385   methyl groups at C-18 and C-21, the signals of which move by 0.23 ppm and 0.19 ppm,

386   respectively (**18**).

387   In the furostanic saponins from the *Agave* genus that contain a carbonyl group at C-12

388   (**32**),[10] the ^{1}H and ^{13}C NMR signals of rings D and E seem to be affected by the sum of

389   the influence of the carbonyl and that of the opened ring F. The positions that show this

390   additive effect are highlighted in Table 4.

391   When a β-hydroxyl group is attached at C-12 instead of a carbonyl, only the

392   neighboring carbons (C-11 and C-13) experience a downfield shift of their δ_C and this is

393   by 10.6 ppm and 6.1 ppm, respectively (**47**) (Table 4).[10] The rest of the carbons in ring

394   C, however, are shielded (Figure 3). In addition, the δ_C of the closest methyl group (C-

395   18) is also shifted upfield, in this case by 5.2 ppm, while the only ring E signals that are

396   affected are those for C-20 (deshielded, 1.2 ppm) and C-21 (shielded, 0.5 ppm).

397   In the ^{1}H NMR spectra of those saponins with a β-hydroxyl group at C-12 (**47**) (Table

398   4),[10] all δ_H from H-9 to H-23a are deshielded to a greater or lesser extent (0.04 ppm to

17

399   0.44 ppm) (Figure 3). The significant signals of methyl groups at C-18 and C-21 are

400   shifted downfield by 0.26 ppm and 0.29 ppm, respectively.

2.5. The double bond between C-9 and C-11

402   In addition to the carbonyl group at C-12, an α,β-unsaturation between C-9 and C-11

403   has been found in some saponins from the *Agave* genus. Such conjugation causes a

404   shielding effect on the carbonyl carbon (Figure 3), the signal of which is often found 8.4

405   ppm upfield (**28**) (Table 4).[52]

406   Besides the downfield shift observed for several carbons as a consequence of the α,β-

407   unsaturated carbonyl group, such as for C-10 (+3.9 ppm), C-8 (+1.8 ppm) and C-13

408   (+10.8 ppm), the double bond also produces a distortion in the ring and this is

409   presumably the cause of shielding of C-14 by 3.4 ppm compared to those saponins

410   without the unsaturation (**18**, **32**). The opposite effect is found for the methyl group at

411   C-19, the signal of which is downfield shifted by 8.1 ppm (**28**), while the other methyl

412   at C-18 is shifted upfield by 1.3 ppm (Figure 3).

413   In the ^{1}H NMR spectra (**29**) (Table 4)[53], the most affected signals are those for H-8, H-

414   19, and H-14, when compared to the hecogenin (**VII**) derivative (**18**), and these are

415   deshielded by 0.97 ppm, 0.17 ppm and 0.71 ppm, respectively. A similar influence is

416   observed for the rest of the rings in the aglycone.

2.6. Free hydroxyl group at C-6 and its glycosylation

418   Position 6 of aglycones can also be oxygenated with a hydroxyl group, which is

419   oriented equatorially and could also be glycosylated[16]. In the case of the free hydroxyl

420   group, α positions (one bond) are deshielded in both the ^{1}H and ^{13}C NMR spectra

421   (Figure 4). The chemical shifts of signals in ^{1}H NMR spectra for H-4ax, H-8, H-9 and

422   H-19 are significantly affected and are deshielded by 0.15 ppm, 0.19 ppm, 0.13 ppm

423   and 0.06 ppm, respectively (**45**) (Table 5),[44] while δ_C for the methyl C-19 is deshielded

424   by 1.2 ppm (**45**) in the ^{13}C NMR spectrum. However, positions beyond ring B are not

425   affected.

426   In cases where the hydroxyl at C-6 is glycosylated, a strong deshielding effect for the

427   signal of C-6 (by 11.6 ppm, **46**) with regards to **45** is observed in the ^{13}C NMR spectra

428   (Table 5).[10] The signals for adjacent carbons are shifted slightly upfield when compared

429   to those in saponins with a free hydroxyl group at C-6 (**45**). In the ^{1}H NMR spectra

430   those signals belonging to protons with an equatorial orientation, such as the

431   glycosylation, are the most affected. In this way, signals for C-4$_{eq}$ and C-7$_{eq}$ are

432   deshielded and can be found at 3.39 ppm and 2.57 ppm (**46**). Glycosylation does not

433   affect the signals for axial protons or the C-19 methyl group when compared with **45**.

434   A long-range effect that is observed is the shielding of the signal for H-16 by 0.13 ppm

435   (**46**) (Table 5), which is added to the influence of H-23 for saponin **58** (Table 3).

436   **2.7. The bridgehead methine at C-5: α- or β-spirostanes and the double bond**

437   **between C-5 and C-6**

438   Together with C-25, the other stereocenter that may vary its configuration within the

439   aglycone backbone of saponins reported from the *Agave* genus is C-5. This carbon is a

440   bridgehead at the junction between rings A and B, and in fact it has only two possible

441   configurations, i.e., *cis* when H-5 is β-oriented (equatorial) or *trans* when it is α-

442   oriented (axial) (Figure 4). In a similar way to C-25, some saponins have also been

443   reported to have a double bond between C-5 and C-6. As saponins described from

444   *Agave* plants mostly have H-5α (**1**) (Table 5),[45] we discuss below the comparison

445 (Figure 4) with those that contain H-5β (**4**)[54] along with those that are unsaturated in C-

446 5 (**9**).[55]

447 The ^{13}C NMR signals for rings A and B of those saponins with a *cis* junction between

448 these two rings are commonly shielded. Such an effect is especially apparent for the

449 methines at C-5 and C-9 and methylenes at C-1 and C-7, which are shifted upfield by

450 7.5 ppm, 13.9 ppm, 6.1 ppm and 5.6 ppm, respectively (Table 5).[54] Agrawal described

451 the C-5, C-7 and C-9 effects[35] by comparison of the aglycones 5α-spirostan-3β-ol and

452 5β-spirostan-3α-ol.[56] It is necessary to note that for saponins of the *Agave* genus the H-

453 5β saponins found are 5β-spirostan-3α-ol and C-1 effects are therefore also observed.

454 The opposite effect is observed for the methyl group at C-19, the signal of which is

455 strongly deshielded (by 11.7 ppm downfield, **4**).

456 In the ^{1}H NMR spectra the observed effect is rather a deshielding and the proton signals

457 of H-3 and H-19 are particularly noteworthy (by 0.40 ppm and 0.22 ppm) (**4**). The *cis*

458 junction of rings A and B also affects the signals of methines oriented towards the α-

459 face beyond these two rings, such as C-14 (+0.07 ppm), C-16 (+0.05 ppm) and C-17

460 (+0.06 ppm).

461 Finally, the presence of a double bond between C-5 and C-6 produces a strong

462 downfield shift in positions C-4 and C-19 (Figure 4), although the δ_C for the allylic

463 carbon C-7 is not affected. Significant shielding is experienced by methines at C-8 (−3.4

464 ppm) and C-9 (−3.8 ppm) in the ^{13}C NMR spectrum of saponin **9**, which has diosgenin

465 (**III**) as the aglycone. The opposite effect is observed in the ^{1}H NMR spectra, where all

466 signals from rings A and B are deshielded, including the methyl C-19 (0.23 ppm). As is

467 common in these systems, the most affected signals are those in the axial position for

468   allylic methylenes C-4 and C-7 and for C-9 (signals move by 1.08 ppm, 0.69 ppm and

469   0.38 ppm) (Table 5).

470   For the H-5β and Δ^5 saponins from the *Agave* genus with a carbonyl (**22**)[54] and (**26**)[10] or

471   hydroxyl group (**48**)[47] at C-12, the ^{1}H and ^{13}C NMR signals for positions 8 to 18 are

472   affected by the sum of its influence. The positions that show this additive effect are

473   highlighted in Table 5.

2.8. Hydroxyl group at C-2

475   A hydroxyl group can frequently be found attached at C-2 and this is always in an

476   equatorial disposition regardless of whether the saponins are H-5α, H-5β or Δ^5 (Figure

477   5).

478   In this case, we will compare the H-2 hydroxylated derivatives (**35**),[45] (**38**),[54] (**40**)[55] of

479   saponins with aglycones that have the same H-5 configuration (**I**, **II** and **III**) (Table 6).

480   As observed in the ^{13}C NMR spectra, the positions adjacent to the hydroxyl group are

481   deshielded, as one would expect, and this is more pronounced for C-1. Positions further

482   away are slightly affected to different extents, perhaps due to spatial relationships

483   between the new hydroxyl group and positions in rings A and B, particularly for the *cis*

484   series (**38**)[54] (Figure 5).

485   The most significant difference in the ^{13}C NMR spectra is observed for the signal of C-

486   19, which in the H-5β case (**38**)[54] it is unaffected, while a downfield shift by 1.2 ppm or

487   1.0 ppm is observed for H-5α (**35**) or Δ^5 (**40**)[45,55] In the case of C-5 (Table 6), the

488   observed effect is characteristic for each kind of structure and a hydroxyl group at C-2

489   does not affect the chemical shift of C-5 when the ring junction A/B is *trans*.[45]

490   However, this signal is shielded by 0.5 and 1.0 ppm, respectively, when the ring fusion

21

491 is *cis* (**38**)[54] or there is a double bond at C-5 (**40**).[55] Shielding is found for the signal of

492 C-4 for aglycones with H-5α and Δ^5 (0.6 ppm and 1.6 ppm), but a downfield shift by 1.3

493 ppm is observed when H-5 is β (**71**).

494 The ^{1}H NMR signals for ring A are sensitive to the nature of the sugar chain attached at

495 C-3, as described in the next section. These signals usually have a large variability, but

496 deshielding is the most common effect (Figure 5) except for the H-3 signal, which in the

497 H-5α (**35**)[45] and Δ^5 derivatives[55] is shifted upfield by between 0.04 ppm and 0.11 ppm

498 (Table 6). For H-5β aglycones,[54] on the other hand, this signal is shifted downfield by

499 0.1–0.14 ppm. The influence extends into ring B to a lesser extent and the signal for H-9

500 is sensitive to the presence of a hydroxyl group at C-2 (H-5α and Δ^5 +0.09 ppm and

501 +0.07 ppm; H-5β −0.08 ppm).

502 For the 2-hydroxylated saponins described from the *Agave* genus with a carbonyl (**50**),

503 (**53**), (**55**)[11] or hydroxyl (**61**)[11] group at C-12, the ^{1}H and ^{13}C NMR signals of positions 8

504 to 18 are affected by the sum of these two factors. The positions that show such additive

505 effects are highlighted in Table 5.

2.9. The main glycosylation at C-3

507 Although unusual, it is possible to find *Agave* saponins with a free hydroxyl group at C-

508 3. However, such saponins are frequently glycosylated at C-6 and C-24.[16] In any case,

509 the absence of glycosylation affects the chemical shifts for carbons C-2, C-3, C-4 and

510 C-5.[35]

511 The influence of the sugar chain attached at C-3 on the signals for rings A and B is

512 analyzed below by considering a selection of sugar chains containing glucose,

513 galactose, and also chains that include up to six sugar moieties (Table 2).

22

514   The sugar chains found to date in *Agave* plants are divided into two groups, namely

515   those with one or two sugar units (short chain) and those with more than three (long

516   chain).[16] Short chain saponins with only one sugar unit usually contain glucose or

517   galactose attached at C-3, while these two monosaccharaides can also be found in a

518   disaccharide chain with different connections. In most cases for the second group, a

519   glucopyranosyl-(1-4)-galactopyranosyloxy unit is attached to position C-3, which in

520   addition may be further branched with other units of glucose, xylose and also

521   rhamnoses.

522   With the aim of studying the influence of the sugar chain on the chemical shifts for

523   rings A and B, we have selected a series of saponins that have the same aglycone

524   moiety and different sugars chains.

525   Saponins with a *cis* junction between rings A and B (H-5β series) usually contain a

526   chain with up to three sugar units and a wide variety of inter-connection types. In these

527   cases, it is quite difficult to identify trends since the δ_C values for these rings may

528   change by ±0.6 ppm, which is close to the range of error. A consistent downfield shift

529   by 0.4 ppm for C-3 in different aglycones is observed when the sugar sequence is

530   formed by a glucose unit 1-4 connected to a galactose (S2C) (**5**, **15**, **22**),[54] or when a

531   glucose is 1-2 connected to the glucose unit of S2C (S3A) (**4**, **16**),[54] when compared

532   with glucopyranosyl derivatives (**11**, **20**, **25**, **30**) (Figure 6).[46,47]

533   The variability observed in the ^{1}H NMR spectra is also within the range of error, which

534   is usually higher for methylenes at C-1, C-6 and C-7. The signal for proton H-3 is also

535   affected (upfield shift by 0.06 ppm) by the previously mentioned sugar chains when

536   compared with glucopyranosyl derivatives (Figure 6). When the sugar chain is longer in

23

537   terms of the number of units, the protons farthest from H-3 are also affected, as

538   observed by comparison between sugar chains S2C and S3A (Figure 6).

539   Saponins with a *trans* fusion (H-5α series) of rings A and B often contain four or more

540   sugar units (Table 2). The ^{1}H and ^{13}C NMR spectra of these materials show little

541   variation and this can be considered to be within the margin of error.

542   **2.10. Concluding structural remarks**

543   The structural features and functional groups present in saponins from *Agave* spp. have

544   a strong influence on ^{1}H and ^{13}C NMR chemical shifts at distances of up to four-bonds.

545   Additionally, in those cases in which there are significant structural changes through

546   space, these effects can be observed as long-range correlations.

547   Thus, rings C and F are influenced by functionalization at C-22, C-23 and C-12, while

548   shielding of rings A and B is altered by functional groups at positions C-2, C-5 and C-6.

549   Moreover, these effects are additive, especially for the combination of oxygenations at

550   C-12 and other structural characteristics. A hydroxyl group at C-2 on the three saponin

551   backbones defined by the nature of H-5 (H-5α, H-5β or Δ^5) has its peculiarities, since

552   different spatial arrangements are adopted. It is therefore useful to confirm the structural

553   elucidation and complete assignments performed by NMR spectroscopic techniques in

554   order to understand these regularities.

555   The most common errors in the assignment of signals in the aglycone core are due to

556   the complexity of certain regions in the spectra. For example, oxygenated positions,

557   which give signals between 3 and 5 ppm in the ^{1}H NMR spectra and 90 and 60 ppm in

558   ^{13}C, are overlapped with chemical shifts of the sugar moiety. The affected signals of the

559   aglycone skeleton are those at the C-2, C-3, C-6, C-12, C-23 and C-24 positions.[57] In

560   contrast, methylene signals of the aglycone backbone usually appear upfield (below 3

561   ppm and 50 ppm in ^{1}H and ^{13}C NMR spectra, respectively). HETCOR or HSQC

562   experiments are used for the structural elucidation of these positions, although it should

563   be noted that these signals are overlapped, which can result in errors in the assignment[58]

564   or the determination of only one of the methylene signals in the ^{1}H NMR spectrum.

565   TOCSY and HSQC-TOCSY experiments[55] are used in order to achieve an unambiguous

566   structure elucidation and signal assignment. These techniques allow the assignment of

567   overlapping signals because they are in a spin system with sufficient deshielded signals

568   to distinguish their correlations easily.

569   In some cases, a knowledge of patterns allows us to propose the revision of assignments

570   or structures described in the bibliography. The NMR data for some described structures

571   are consistent with those previously reported for ^{13}C NMR spectra, but the ^{1}H NMR

572   chemical shifts differ significantly, even in the case of the methyl signals.[59] A

573   significant example is the saponin described from *Hosta plantaginea*,[60] which has an

574   aglycone skeleton that is included in this paper. The authors proposed an *α,β*-

575   unsaturated ketone on ring C as a functional group of the aglycone **XI**. Although the [13]C

576   NMR data are in accordance with the signals reported previously,[52] the key ^{1}H NMR

577   chemical shifts (for instance H-11 at 5.87 ppm or H-21 at 1.33 ppm) are not consistent

578   with the proposed structure. In fact, the chemical shift of the methyl group at H-21 is

579   affected by the conjugation of the carbonyl group at C-12. This methyl signal shows

580   values between 1.32 and 1.34 ppm when it is not in conjugation or in the range from

581   1.41 to 1.39 ppm when it is conjugated (Table 7).[10,52,53]

582   The steroidal saponin described from *Agave attenuata*[61] has sarsapogenin (25S-5β-

583   spirostanol) as the aglycone. The spectroscopic data for the C-19 position are in

25

584   agreement with an H-5β disposition, but the rest of the ^{13}C NMR chemical shifts of

585   rings A and B are very diverse and they do not have the characteristic shielding data for

586   H-5β in relation to H-5α-spirostanes (see section 2.7). Moreover, the structural

587   elucidation was carried out by comparison of ^{13}C NMR data of sapogenins.[40] Therefore,

588   a further confirmation of the chemical structure through the use of two-dimensional

589   NMR experiments is suggested in this case.

590   **3. IDENTIFICATION OF SAPONINS USING ^{1}H NMR AND HMBC SPECTRA**

591   The comprehensive study of the effects that functional groups and structural

592   characteristics have on the aglycone signals in the ^{1}H and ^{13}C NMR spectra of *Agave*

593   saponins indicated that signals are affected up to four bonds away, while changes in

594   signals that are further away lie within error. Full assignment of signals in the ^{13}C NMR

595   spectrum has been performed for the last 35 years and these data have been reported for

596   all saponins. Regarding ^{1}H NMR spectra, the signals for methyl groups have also been

597   widely described. Assignment of the rest of the signals has been possible with the aid of

598   high resolution instruments, including the use of multidimensional and selective

599   excitation experiments. However, the signals due to methylene groups have a larger

600   range of error.

601   Most of the functionalization found in saponins from *Agave* plants are up to four bonds

602   away from protons of methyl groups. For this reason, we propose the use of ^{1}H NMR

603   and HMBC spectra only for aglycone identification when the saponin is pure or is

604   present in a non-complex mixture.

605   The signals of methyl groups in ^{1}H NMR spectra are easily visible due to their high

606   intensity (three equivalent protons). In the case of *Agave* saponins two methyl singlets

607   (CH_3-18 and -19) and two doublets (CH_3-21 and -27) are observed, except for those

26

608   derivatives with a double bond between C-25 and C-27, which are easily recognizable

609   by the presence of two broad singlets (4.76 and 4.79 ppm).

610   The HMBC spectrum allows correlations to be observed through distances of two and

611   three bonds from methyl groups. The main advantage of using this technique is that one

612   can use as references the signals with a less ambiguous assignment since these are the

613   best described in the literature. In addition, smaller amounts of sample are required for

614   acquisition when compared to a good quality ^{13}C NMR spectrum. Furthermore, with

615   predictive techniques such as NUS[62] it is possible to find the optimal time for the

616   spectrum acquisition that guarantees its use for the study and even quantification of

617   saponins.

618   The method described below for the identification of saponins from *Agave* plants could

619   also be applied for other genera with saponins with the same structural features. The

620   method begins with the methyl groups (singlets and doublets) as reference signals,

621   without the need to assign each signal, from which correlations can be observed in the

622   HMBC spectrum that help to distinguish specific structural features.

3.1. HMBC correlations of methyl doublets

624   The methyl doublets (CH$_3$-21 and -27) typically present correlations with signals from

625   26 ppm up to 113 ppm (Table 7). This latter signal belongs to C-22 (109-113 ppm),

626   three bonds away from methyl-21, and may be used for the rapid distinction of methyl-

627   21. On the other side of the range can be found C-24, with its resonance below 30 ppm

628   for those saponins without functionalization in ring F (spirostanic or furostanic). For the

629   doublet of methyl-27, the 25R-spirostanic saponins present a more shielded H-27 signal

630   than the rest, with a δ_H of less than 0.80 ppm. That signal appears close to 1 ppm in

631   furostanic and 25S-spirostanic saponins and between them it is easy to recognize a

27

632 correlation with C-26 at δ_C 75.3 or δ_C 65.2, respectively. Given that C-23 is four bonds

633 away from C-21 and C-27, the presence of a hydroxyl group at C-23 can be determined

634 based on a correlation between H-27 and C-24 (38.9 ppm). In order to determine the

635 presence of a glucopyranosyloxy group at C-24, the three-bond correlation with this

636 methyl doublet is useful. Methyl-27 shows a three-bond correlation with a carbon at

637 81.5 ppm or at 87.9 ppm when C-23 is also hydroxylated.

638 The doublet signal for the methyl group at C-21 is more deshielded than that of C-27,

639 usually by between 1.0 ppm and 1.6 ppm, and it is very easy to distinguish due to its

640 HMBC correlation with the signal of C-22 (109–113 ppm).

641 The structural features that are most easily recognizable by using an HMBC spectrum of

642 a saponin is the kind of oxygenation at C-12 (hydroxyl or carbonyl group), furostanic or

643 spirostanic nature or a combination of these two. The presence of a carbonyl group at C-

644 12 causes a strong shielding at C-17 and this may be observed in correlations from

645 methyl-21, which is crucial for discriminating between the two possible

646 functionalizations at this position. Methyl-21 also serves to distinguish a methoxylated

647 furostanic saponin, since it has a common three bond correlation with the methoxyl

648 group (3.24 ppm) on C-22.

649 The methyl-21 signal appears at around 1.4 ppm in the ^{1}H NMR spectrum of saponins

650 with a hydroxyl group or a,β-unsaturated carbonyl group at C-12. Based on its HMBC

651 correlation with C-17 (hydroxyl group **47**: 63.0 ppm; a,β-unsaturated carbonyl group

652 **28**: 54.5 ppm) these can be distinguished from each other (Table 7).

653 The carbon signal for C-20 has a value of around 35 ppm only when C-23 is

654 oxygenated. This may be distinguished from those saponins that also contain

655 glucosylation at C-24 on the basis of deshielding experienced by C-22 (up to 112.7

28

656   ppm). The only two types of saponins that have a signal of C-22 at 111.6 ppm are those

657   with a hydroxyl group at C-23 or C-24. Only in the latter case, however, does the signal

658   of C-20 have a chemical shift of 42.1 ppm.

659   **3.2. HMBC correlations of methyl singlets**

660   With the previous assignment of the C-17 chemical shift, its three bond HMBC

661   correlation with a three-proton singlet allows the assignment of methyl C-18. Now,

662   from this methyl it is possible to identify the kind of functionalization at C-12. Thus, the

663   HMBC correlation with a ^{13}C signal at 79.3 ppm (**47**), 212.8 ppm (**18**, **32**) or 204.4 (**28**)

664   can confirm the presence of a hydroxyl, a carbonyl, or an α,β-unsaturated carbonyl

665   group, respectively (Table 8).

666   Spirostanic saponins with a hydroxyl group at C-23 give a ^{1}H NMR signal for methyl

667   C-18 at 0.96 ppm (**58**). The correlation with the signal of C-17 at around 62 ppm,

668   together with the absence of correlations with signals above 200 ppm, can confirm such

669   a structural feature. It is worth mentioning that the glucopyranosyl group at C-24 causes

670   strong shielding of the methyl-18 signal (**62**: 0.71 ppm) when other functionalization is

671   not present in ring F. On the other hand, when functionalization is not present on rings

672   C or F, the C-18 signal is observed in the range from 0.76 ppm to 0.85 ppm and it

673   differs by 1.0 ppm for the carbon signal of C-17 of spirostanic and furostanic saponins

674   (**40**: 62.9 ppm; **43**: 63.9 ppm) (Table 8). The rest of the correlations that have not

675   previously been mentioned, i.e., those with C-13 and C-14, can be grouped into two

676   ranges between 39.0 ppm and 42 ppm, and between 54 ppm and 57 ppm.

677   As described before, the signal of the angular methyl-19 can be influenced by

678   functionalization in rings A–C, but also by the nature of the sugar chain attached at C-3.

679   As described above for methyl-18, several signals between 35 ppm and 46 ppm and

29

680   between 53 ppm and 55 ppm (C-1, C-5, C-9 and C-10) show HMBC correlations with

681   methyl-19. The chemical shifts of these signals are significantly influenced by the

682   presence of functional groups in rings A and B. However, they are rarely observed away

683   from these ranges and they are therefore not relevant for structure elucidation.

684   The double bond between C-5 and C-6 or between C-9 and C-11 that is usually present

685   in saponins from *Agave* is three bonds away from methyl-19. In this way, the HMBC

686   correlations of this methyl singlet with carbon signals at 141.1 ppm (C-5, **9**), or with

687   that at 171.3 ppm (C-9, **28**), can be used for their assignment.

688   A *cis* junction between rings A and B (H-5β) gives rise to a strong deshielding of the

689   methyl-19 signal and in the ^{1}H NMR spectrum this signal is in the range between 0.79

690   and 0.84 ppm. In the case of a *trans* junction (H-5α) this methyl-19 signal is at higher

691   field, except for those compounds with a double bond between C-9 and C-11.

692   The C-6 position is a four-bond ^{1}H-^{1}H correlation from the methyl group at C-19. When

693   this position is oxygenated, the HMBC spectrum shows a three-bond correlation

694   between methyl-19 and C-5 between 50 and 53 ppm (Table 8). The hydroxylation at C-

695   6 can also significantly deshield the signal of H-4$_{eq}$, which may be found between 3.36

696   ppm and 3.39 ppm when a sugar chain is also attached at C-3, or at 3.23 ppm if a free

697   hydroxyl is present at C-6. This effect is easily detectable since the spectroscopic region

698   where this signal appears is not frequently hindered by overlap with other proton

699   signals. In this way, a doublet with a coupling constant of 12 Hz is easily visible and is

700   diagnostic of the aforementioned structural feature.

701   The position C-2 is also a four-bond correlation from methyl C-19 and its presence

702   causes appreciable deshielding of the resonance for the C-1 position. Its correlation is

703   found between 35 and 46 ppm in the HMBC spectrum and the overlap of several signals

30

704 can also be observed (Table 8). This deshielding makes it easier to distinguish between

705 saponins that are hydroxylated (**38**, 40.5 ppm)[54] and non-hydroxylated (**4**, 30.8 ppm)[54]

706 at the C-2 position for H-5β saponins (Table 8). Furthermore, a downfield shift for C-5

707 of 1 ppm can be readily observed for saponins with a double bond between C-5 and C-

708 6, Δ^5 (**9**: 141.1 ppm[55]; **40**, **43**, **44**, **62**: 140.1 ppm[49,51,55]) (Table 8). On the other hand, the

709 methyl group C-19 is influenced by the C-2 hydroxyl group, as observed in the [1]H

710 NMR spectrum. Nevertheless, the downfield shift produced for the C-19 singlet is close

711 to the error range and it is also very sensitive to the sugar chain chemical shifts. Most of

712 the H-5α saponins found in *Agave* spp. have a sugar chain with three or more units,

713 which include a galactose bonded at the C-3 position (Table 2). The anomeric proton in

714 the ^{1}H NMR spectrum is not overlapped with other signals in the range from 4.81 ppm

715 to 4.88 ppm when there is no functional group at the C-2 position (**1-3**, **18**, **47**).[45,48]

716 However, when there is a hydroxyl group this value *is* between 4.89 and 4.92 ppm (**35**,

717 **50**, **51**, **55**, **61**).[10,45] All of these considerations – together with the corresponding

718 HMBC correlations reported for methyl at C-19 (Table 8) – can indicate the presence or

719 absence of a hydroxyl group at C-2 on the 5α-spirostanol core.

720 **3.3 Flowchart for the HMBC Method for aglycone identification (HMAI).**

721 Two flowcharts (Figures 7 and 8) are proposed as a tool for the identification of

722 aglycones of saponins from the *Agave* species (Table 1) by ^{1}H NMR and HMBC

723 experiments. Priority has been given to ^{13}C NMR signals for the detection of structural

724 features due to their lower sensitivity to the solvent and the nature of the sugar moiety

725 linked at C-3 when compared to methyl resonances in ^{1}H NMR spectrum.

726 Both ranges of chemical shifts and absolute values are provided in the flowchart. In this

727 last case, values within the error range established for this work for ^{1}H and ^{13}C NMR

31

728   signals should be considered. Spectra should be referenced to deuterated pyridine (7.55

729   ppm and 135.6 ppm for ^{1}H and ^{13}C NMR, respectively).

730   Prior assignment is not needed and only proton signals (for three equivalent protons)

731   should be distinguished between 1.6 ppm and 0.5 ppm. These signals are readily

732   recognizable and correspond to secondary methyl groups at C-21 and C-27, which are

733   doublets and singlets for angular methyl groups C-18 and C-19. The flowchart should

734   be started with the methyl doublets that provide information on rings C–F. If only one

735   doublet is observed it is verified that the aglycone contains a double bond at C25(27),

736   while if two doublet methyl signals are observed it is the doublet that is more deshielded

737   that is analyzed first (usually C-27).

738   Secondly, methyl groups that give rise to singlets should be investigated and the most

739   deshielded position will be applied in the flowchart (Figure 8) to assign each methyl

740   group. In the case of the methyl at C-18, signals from HMBC experiments mainly

741   provide structural information about ring C and this should be identical to the data

742   obtained in the study of the methyl group doublets.

743   In some cases, the flowchart indicates that HMBC values for a specific methyl should

744   be revised. In this situation, taking into consideration the structural features of the

745   remaining methyls, Tables 7 and 8 should be used.

746   In a case where the HMBC signals are different to those indicated in the flowchart, the

747   saponin should have other structural characteristics that will require elucidation.

748   A representative example for the application of a saponin in the flowchart is described

749   below (Figure 9). More examples with structural features included and excluded from

750   this study are provided in the supporting information.

751   A ^{1}H NMR spectrum contains two doublet signals for three equivalent protons at 1.16

752   ppm and 0.72 ppm. The most deshielded doublet has HMBC correlations with the

753   signals at 31.9, 38.9 and 66.0 ppm. As the signal did not show correlations with values

754   lower than 30 ppm (D1-Figure 7) or in the range 109–113 (D5-Figure 7), 88 ppm (D11)

755   and 81 ppm (D12), it is verified that it has a correlation at approximately 39 ppm (D13),

756   specifically at 38.9 ppm. This indicates that the methyl analyzed is found at C-27 of a

757   25R-spirostanic-type saponin and it possesses a hydroxyl group in an α disposition at C-

758   23. The second methyl doublet at 1.16 ppm in the ^{1}H NMR spectrum has HMBC

759   correlations with 35.9, 62.6 and 111.7 ppm and these correlations do not fit premise D1.

760   Nonetheless, a correlation is observed between 109 ppm and 113 ppm (D5-Figure 7),

761   the characteristic chemical shift of the C-22 position in the ^{13}C NMR spectrum. This

762   may confirm that this doublet corresponds with the C-21 position. Moreover,

763   correlations at 54–55 ppm (D6) and approximately 109 ppm (D9) are not observed.

764   Thus, by following the indications established in the flowchart, spectroscopic data for

765   C-21 should be compared with those listed in Table 7 (D10). Comparison of the

766   shielding data confirms that this compound is a 25R-spirostanic-type saponin with a

767   hydroxyl group in an α disposition on C-23 (**58**).

768   The application of the flowchart to the methyl singlets is described in more detail below

769   (Figure 8). The most downfield shift at 0.96 ppm in the ^{1}H NMR spectrum shows

770   HMBC correlations with signals at 41.0 ppm, 56.5 ppm and 62.6 ppm. Since it does not

771   have any correlation with values higher than 70 ppm, premises from S1 to S8 are

772   discarded (Figure 9). On the other hand, a correlation at 62.6 ppm fits with premises S9

773   and S11 (Figure 8) and therefore this singlet is due to the methyl group at C-18. The

774   HMBC connectivity data are consistent with those reported for a 25R-spirostanic-type

775   saponin with a hydroxyl group in an α orientation at C-23 (**58**, Table 8). It is worth

776   highlighting that the two connectivities that should appear at 40.5 ppm (C-12) and 41.4

777   ppm (C-13) are overlapped and they are observed as a single signal at 41.0 ppm. The

778   second methyl singlet (0.74 ppm) corresponds with the C-19 position on the aglycone

779   skeleton and it shows multiple-bond ^{1}H-^{13}C connectivities at 37.2 ppm, 51.4 ppm and

780   53.9 ppm. As in the case of the methyl group at C-18, correlations with values higher

781   than 70 ppm are not observed, although correlations with values higher than 50 ppm are

782   found (S9-Figure 8). Moreover, there are no correlations with chemical shifts greater

783   than 60 ppm (S11) but there are between 50 ppm and 52.5 ppm (S13). Besides, the

784   presence of a doublet signal at 3.23 ppm and a coupling constant of 12 Hz in the ^{1}H

785   NMR spectrum (S14) confirms an aglycone with a glucopyranosyloxy group linked to

786   C-6 and a free hydroxyl group at the C-3 position. These structural features led us to

787   identify this aglycone as 25R-spirostane-3β,6α,23α-triol (**XXVI**, Table 1).

788   The example explained below is a mixture of saponins with the same sugar moiety

789   bonded at C-3 on the aglycone core. Singlet and doublet signals can be observed in the

790   ^{1}H NMR spectrum. Some of these signals are overlapped because the differences

791   between these are far from the methyl groups under study. Doublet signals are observed

792   between 1.38 ppm and 1.32 ppm along with overlapping regions at around 0.66 ppm,

793   which could be doublets (Table 9). HMBC correlations (29.2, 30.5 and 66.9 ppm, Table

794   9) are consistent with the C-27 position (D1-Figure 7). As the H-27 resonance is below

795   0.80 ppm (D2-Figure 7), 25R-spirostane saponins can be proposed. Other downfield

796   doublet signals have similar correlations. The signals found at 109.3 ppm or 109.5 ppm

797   are consistent with them being due to C-21 (D5-Figure 7). Furthermore, the second

798   correlation at 54.5 ppm or 54.3 ppm (D6-Figure 7) and signals at 109.5 ppm (D7) led us

799   finally to D8. In this way, it can be determined that a saponin with chemical shift at 1.38

800   ppm in ^{1}H NMR spectrum has an α,β-unsaturated carbonyl group at the C-12 position,

801   while a saponin with proton shielding at 1.32 ppm contains a single carbonyl group.

802   Several singlet signals are found in the ^{1}H NMR spectrum. The singlets corresponding

803   to the C-18 position of each saponin can be distinguished by considering correlations of

804   C-17, since they are related with the doublet methyl signals at C-21 assigned previously.

805   A correlation at 54.5 ppm connects the most deshielded doublet (1.38 ppm) with a

806   singlet at 0.98 ppm. One of the HMBC correlations for this singlet that appears at 204.3

807   ppm is found within the range established in S3 (204 and 205 ppm, Figure 8). This fact

808   indicates that there is an α,β-unsaturated carbonyl group at the C-12 position and it

809   supports the characteristic identified through the correlations of the methyl doublet. The

810   other two singlet signals, which are very close to each other (1.05 ppm and 1.04 ppm),

811   share the correlation at 54.2 ppm as overlapped doublets found at 1.32 ppm. There is

812   also a correlation at 212.7 ppm within the range 212–214 ppm (S1-Figure 8). The

813   proton chemical shift observed at 1.05 ppm indicates that the spirostanic saponins have

814   a carbonyl group at C-12 (S2-Figure 8), a situation that was already deduced through

815   the analysis of the corresponding methyl doublet.

816   Overall, it can be stated that the three saponins are 25R-spirostane type with a carbonyl

817   group at C-12 and that one of them is α,β-unsaturated.

818   The three remaining singlets in the HMBC spectrum correspond to methyl groups at C-

819   19 (Table 9). The most deshielded methyl group, at 0.80 ppm, shows a correlation with

820   171.3 ppm, which is in accordance with decision S5. For the third time, it can be

821   concluded that this saponin has an α,β-unsaturated carbonyl group in ring C and,

822   furthermore, there is no functionalization in rings A or B and the saponin is H-5α. The

823   other two methyl singlets, which belong to aglycones with the same functionalization in

824   rings C-F, are separated by 0.7 ppm and this suggests that the difference between them

825   is on ring A or B. HMBC correlations for both saponins are very similar and they are

826   around 55 ppm (S9-Figure 8). Since there are no correlations between 50 ppm and 53

827   ppm (S11), it can be determined that the A/B ring junction is *trans* for both compounds.

828   The anomeric proton H-1Gal signal appears below 4.88 ppm and this is related to

829   saponins without a hydroxyl group at C-2 (S18-Figure 8), although other signals appear

830   up to 4.88 so there could be some overlapped signal of H-1Gal of a C-2 hydroxylated

831   derivative. Thus, on comparing the correlations of the two methyl signals at 0.72 ppm

832   and 0.64 ppm with those previously reported for compounds **35** and **1** (with and without

833   a hydroxyl group at C-2, respectively) (Table 8), it can be deduced that the most

834   deshielded signal fits with the presence of a hydroxyl group at C-2, while the most

835   shielded resonance is due to the absence of further functionalizations in rings A and B.

836   The HMAI method discussed here afforded the assignment of each methyl group for a

837   mixture of three saponins (Table 9) and identified these aglycones as **VII**, **XI** and **XXII**

838   (Table 1). It is worth mentioning that methyl doublet and singlet signals are separated

839   from each other in ranges within the pattern rules established in section 2. For instance,

840   the methyl at C-19 is particularly affected. This methyl is influenced by the presence of

841   a hydroxyl group at C-2, which produces a shielding variation of 0.7 ppm (Table 6). In

842   our study this change was 0.8 ppm for aglycones **VII** and **XXII**. Moreover, a

843   deshielding of 0.17 ppm is observed if there is an α,β-unsaturated carbonyl group (Table

844   4). In the aforementioned example, this shielding variation is 0.16 ppm for aglycones

845   **VII** and **XI**. Additionally, long-range correlations due to the presence of a hydroxyl

846   group at C-2 are also observed with the C-18 and C-21 signals as a slight shielding

847   (0.01) which – although within the error range of this study – allows us to distinguish

848   the two signals (Table 9).

849   **4. CONCLUSIONS**

850   The pattern rules noted in ^{1}H and ^{13}C NMR spectra for the most representative

851   aglycones from the *Agave* species indicate that functionalization and structural

852   modifications produce a shielding variation over a long-range connectivity of up to four

853   bonds. These effects are usually within the measurement error range at higher multiple-

854   bond correlation. Besides, the effects are additive and the most influenced signals can

855   be used to identify structural features. Aglycones can be analyzed through three-bond

856   ^{1}H-^{13}C correlations observed in HMBC experiments for the methyl groups. These

857   signals are particularly intense and this fact can be used to reduce the acquisition time of

858   the experiment. Predictive techniques (NUS) can also be used to obtain the desired

859   correlations in a short time. The chemical shifts for these methyl groups and ^{13}C NMR

860   data are usually reported in the bibliography and a spectroscopic data source should be

861   available to identify other types of saponin.

862   Saponins with a combination of structural requirements, including those evaluated in

863   this study (Table 1), can be identified through the method for aglycone identification

864   (HMAI). This method includes a flowchart that facilitates the identification of

865   aglycones. Besides, the method identifies saponins and proposes structural elucidation

866   for those that do not fit the premises described. This method has been tested with

867   HMBC spectra of different saponins and with data reported in the bibliography. ^{13}C

868   NMR signals are mainly used because they are less influenced by long-range effects or

869   small variations in deuterated solvents.

870   The application of the HMAI method to a mixture of three saponins allowed the

871   identification of each aglycone, and it was ascertained that the signals have the same

872   HMBC correlations as pure saponins within the error range established. Moreover, the

873   shielding range of the methyl groups in the mixture of saponins fits the patterns

874   observed in the ^{1}H NMR spectra and the results are therefore reinforced.

875 Overall, the HMBC spectrum of methyl signals is proposed as a starting point for the

876 identification of aglycones from *Agave* saponins. These studies, combined with HPLC-

877 MS techniques, can be used for quality control or to monitor products that contain

878 *Agave* saponins.

879

880

881

882

883

884 **ACKNOWLEDGMENTS**

889 **ORCID**

890 Ana M. Simonet   0000-0002-6516-1783

891 Alexandra G. Durán   0000-0002-9799-0850

892 Andy J. Pérez   0000-0001-6717-2040

893 Francisco A. Macías   0000-0001-8862-2864

894

REFERENCES

1   Osbourn A. Saponins and plant defence — a soap story. *Trends Plant Sci* 1996;1:4–9.

2   Hussain M, Debnath B, Qasim M, et al. Role of Saponins in Plant Defense Against Specialist Herbivores. *Molecules* 2019;24:2067.

3   Faizal A, Geelen D. Saponins and their role in biological processes in plants. *Phytochem Rev* 2013;12:877–893.

4   Yang Y, Laval S, Yu B. Chemical Synthesis of Saponins. In: Horton D, editor. Advances in Carbohydrate Chemistry and Biochemistry, First edit. London: Elsevier Inc., 2014. p. 137–226.

5   Oleszek W, Bialy Z. Chromatographic determination of plant saponins—An update (2002–2005). *J Chromatogr A* 2006;1112:78–91.

6   Sparg SG, Light ME, van Staden J. Biological activities and distribution of plant saponins. *J Ethnopharmacol* 2004;94:219–243.

7   Wang Y, Zhang Y, Zhu Z, et al. Exploration of the correlation between the structure, hemolytic activity, and cytotoxicity of steroid saponins. *Bioorg Med Chem* 2007;15:2528–2532.

8   Yang C-R, Zhang Y, Jacob MR, et al. Antifungal Activity of C-27 Steroidal Saponins. *Antimicrob Agents Chemother* 2006;50:1710–1714.

9   Calle JM, Pérez AJ, Simonet AM, et al. Steroidal Saponins from *Furcraea hexapetala* Leaves and Their Phytotoxic Activity. *J Nat Prod* 2016;79:2903–

39

917   2911.

918   10   Pérez AJ, Calle JM, Simonet AM, et al. Bioactive steroidal saponins from *Agave*

919   *offoyana* flowers. *Phytochemistry* 2013;95:298–307.

920   11   Pérez AJ, Simonet AM, Calle JM, et al. Phytotoxic steroidal saponins from

921   *Agave offoyana* leaves. *Phytochemistry* 2014;105:92–100.

922   12   Teshima Y, Ikeda T, Imada K, et al. Identification and Biological Activity of

923   Antifungal Saponins from Shallot (*Allium cepa* L. Aggregatum Group). *J Agric*

924   *Food Chem* 2013;61:7440–7445.

925   13   Mimaki Y, Yokosura A, Kuroda M, et al. Cytotoxic Activities and Structure-

926   Cytotoxic Relationships of Steroidal Saponins. *Biol Pharm Bull* 2001;24:1286–

927   1289.

928   14   Fujino T, Yokosuka A, Higurashi H, et al. AU-1 from Agavaceae plants causes

929   transient increase in p21/Cip1 expression in renal adenocarcinoma ACHN cells

930   in an miR-34-dependent manner. *J Nat Med* 2017;71:36–43.

931   15   Fujino T, Yokosuka A, Ichikawa H, et al. AU-1 from Agavaceae plants

932   downregulates the expression of glycolytic enzyme phosphoglycerate mutase. *J*

933   *Nat Med* 2018;72:342–346.

934   16   Sidana J, Singh B, Sharma OP. Saponins of *Agave*: Chemistry and bioactivity.

935   *Phytochemistry* 2016;130:22–46.

936   17   Cheok CY, Salman HAK, Sulaiman R. Extraction and quantification of saponins:

937   A review. *Food Res Int* 2014;59:16–40.

938   18   Krzyzanowska J, Kowalczyk M, Oleszek W. Analysis of Plant Saponins. In:

939   Encyclopedia of Analytical Chemistry. Chichester, UK: John Wiley & Sons, Ltd,

940   2014. p. 1–21.

941   19   Urbina CJF, Casas A, Martínez-Díaz Y, et al. Domestication and saponins

942   contents in a gradient of management intensity of agaves: *Agave cupreata*, *A.*

943   *inaequidens* and *A. hookeri* in central Mexico. *Genet Resour Crop Evol*

944   2018;65:1133–1146.

945   20   Puente-Garza CA, García-Lara S, Gutiérrez-Uribe JA. Enhancement of saponins

946   and flavonols by micropropagation of *Agave salmiana*. *Ind Crops Prod*

947   2017;105:225–230.

948   21   Puente-Garza CA, Meza-Miranda C, Ochoa-Martínez D, et al. Effect of in vitro

949   drought stress on phenolic acids, flavonols, saponins, and antioxidant activity in

950   *Agave salmiana*. *Plant Physiol Biochem* 2017;115:400–407.

951   22   Leal-Díaz AM, Santos-Zea L, Martínez-Escobedo HC, et al. Effect of *Agave*

952   *americana* and *Agave salmiana* Ripeness on Saponin Content from Aguamiel

953   (Agave Sap). *J Agric Food Chem* 2015;63:3924–3930.

954   23   Puente-Garza CA, Espinosa-Leal CA, García-Lara S. Steroidal Saponin and

955   Flavonol Content and Antioxidant Activity during Sporophyte Development of

956   Maguey (*Agave salmiana*). *Plant Foods Hum Nutr* 2018;73:287–294.

957   24   Santos-Zea L, Rosas-Pérez AM, Leal-Díaz AM, et al. Variability in Saponin

958   Content, Cancer Antiproliferative Activity and Physicochemical Properties of

959   Concentrated Agave Sap. *J Food Sci* 2016;81:H2069–H2075.

960 25 Figueroa L, Santos-Zea L, Escalante A, et al. Mass Spectrometry-Based

961   Metabolomics of Agave Sap (*Agave salmiana*) after Its Inoculation with

962   Microorganisms Isolated from Agave Sap Concentrate Selected to Enhance

963   Anticancer Activity. *Sustainability* 2017;9:2095.

964 26 Marker RE, Wagner RB, Ulshafer PR, et al. Sterols. CLVII. Sapogenins. LXIX.

965   1 Isolation and Structures of Thirteen New Steroidal Sapogenins. New Sources

966   for Known Sapogenins. *J Am Chem Soc* 1943;65:1199–1209.

967 27 Wagner RB, Forker RF, Spitzer PF. The Δ9-12-Keto Steroidal Sapogenins. *J Am*

968   *Chem Soc* 1951;73:2494–2497.

969 28 Wilkomirski B, Bobeyko VA, Kintia PK. New steroidal saponins of *Agave*

970   *americana*. *Phytochemistry* 1975;14:2657–2659.

971 29 Sati OP, Rana U, Chaukiyal DC, et al. A New Spirostanol Glycoside from *Agave*

972   *cantala*. *J Nat Prod* 1987;50:263–265.

973 30 Kiyosawa S, Hutoh M, Komori T, et al. Detection of Proto-type Compounds of

974   Diosgenin-and Other Spirostanol-Glycosides. *Chem Pharm Bull (Tokyo)*

975   1968;16:1162–1164.

976 31 Sharma SC, Sati OP. A spirostanol glycoside from *Agave cantala*.

977   *Phytochemistry* 1982;21:1820–1821.

978 32 Sati OP, Pant G, Miyahara K, et al. Cantalasaponin-1, A Novel Spirostanol

979   Bisdesmoside from *Agave cantala*. *J Nat Prod* 1985;48:395–399.

980 33 Yi D, Yan-Yong C, De-Zu W, et al. Steroidal saponins from a cultivated form of

981   *Agave sisalana*. *Phytochemistry* 1989;28:2787–2791.

982   34   Uniyal GC, Agrawal PK, Sati OP, et al. A spirostane hexaglycoside from *Agave*

983   *cantala* fruits. *Phytochemistry* 1991;30:4187–4189.

984   35   Agrawal PK, Jain DC, Pathak AK. NMR spectroscopy of steroidal sapogenins

985   and steroidal saponins: An update. *Magn Reson Chem* 1995;33:923–953.

986   36   Plock A, Beyer G, Hiller K, et al. Application of MS and NMR to the structure

987   elucidation of complex sugar moieties of natural products: exemplified by the

988   steroidal saponin from *Yucca filamentosa* L. *Phytochemistry* 2001;57:489–496.

989   37   Yokosuka A, Mimaki Y, Sashida Y. Four New 3,5-Cyclosteroidal Saponins from

990   *Dracaena surculosa. Chem Pharm Bull (Tokyo)* 2002;50:992–995.

991   38   Khakimov B, Tseng L, Godejohann M, et al. Screening for Triterpenoid Saponins

992   in Plants Using Hyphenated Analytical Platforms. *Molecules* 2016;21:1614.

993   39   Kim HK, Saifullah, Khan S, et al. Metabolic classification of South American

994   Ilex species by NMR-based metabolomics. *Phytochemistry* 2010;71:773–784.

995   40   P. K. Agrawal, D.C. Jain, R. K. Gupta, et al. Carbon-13 NMR Spectoscopy of

996   Steridal Sapogenins and Steridal Saponins. *Phytochemistry* 1985;24:2479–2496.

997   41   Agrawal PK. Assigning stereodiversity of the 27-Me group of furostane-type

998   steroidal saponins via NMR chemical shifts. *Steroids* 2005;70:715–724.

999   42   Agrawal PK. Dependence of1H NMR chemical shifts of geminal protons of

1000   glycosyloxy methylene (H2-26) on the orientation of the 27-methyl group of

1001   furostane-type steroidal saponins. *Magn Reson Chem* 2004;42:990–993.

1002 43   Agrawal PK, Burkholz T, Jacob C. Revisit to 25R/25S Stereochemical Analysis

1003   of Spirostane-type Steroidal Sapogenins and Steroidal Saponins via ^{1}H NMR

1004   Chemical Shift Data. *Nat Prod Commun* 2012;7:709–711.

1005 44 Ohtsuki T, Koyano T, Kowithayakorn T, et al. New chlorogenin hexasaccharide

1006   isolated from *Agave fourcroydes* with cytotoxic and cell cycle inhibitory

1007   activities. *Bioorg Med Chem* 2004;12:3841–3845.

1008 45 Calle JM. Estudios de fitotoxicidad y relación estructura-actividad (SAR) de

1009   saponinas esteroidales. Ph. D. Thesis, University of Cadiz, 2015.

1010 46 Qu L, Ruan J, Wu S, et al. Separation and Bioactive Assay of 25R/S-Spirostanol

1011   Saponin Diastereomers from *Yucca schidigera* Roezl (Mojave) Stems. *Molecules*

1012   2018;23:2562.

1013 47 Qu L, Wang J, Ruan J, et al. Spirostane-Type Saponins Obtained from *Yucca*

1014   *schidigera*. *Molecules* 2018;23:167.

1015 48 Pérez AJ. Estudio fitoquímico de especies nativas de Cuba pertenecientes a la

1016   familia Agavaceae y evaluación de sus actividades biológicas. Ph. D. Thesis,

1017   University of Cadiz, 2011.

1018 49 Mimaki Y, Watanabe K, Sakagami H, et al. Steroidal Glycosides from the

1019   Leaves of *Cestrum nocturnum*. *J Nat Prod* 2002;65:1863–1868.

1020 50 Jin J-M, Zhang Y-J, Yang C-R. Four New Steroid Constituents from the Waste

1021   Residue of Fibre Separation from *Agave americana* Leaves. *Chem Pharm Bull*

1022   *(Tokyo)* 2004;52:654–658.

1023 51 Macías FA, Guerra JO, Simonet AM, et al. Characterization of three saponins

1024   from a fraction using [1]D DOSY as a solvent signal suppression tool.

44

1025   Agabrittonosides E-F. Furostane Saponins from *Agave brittoniana* Trel. spp.

1026   *Brachypus*. *Magn Reson Chem* 2010;48: 350-355.

1027   52   Mimaki Y, Kanmoto T, Kuroda M, et al. Steroidal Saponins from the

1028   Underground Parts of Hosta longipes and Their Inhibitory Activity on Tumor

1029   Promoter-Induced Phospholipid Metabolism. *Chem Pharm Bull (Tokyo)*

1030   1995;43:1190–1196.

1031   53   Benito J. Aislamiento biodirigido de saponinas de la especie *Agave americana*.

1032   Master thesis, University of Cadiz, 2017.

1033   54   Yokosuka A, Jitsuno M, Yui S, et al. Steroidal Glycosides from *Agave utahensis*

1034   and Their Cytotoxic Activity. *J Nat Prod* 2009;72:1399–1404.

1035   55   Macías FA, Guerra JO, Simonet AM, et al. Characterization of the fraction

1036   components using 1D TOCSY and 1D ROESY experiments. Four new spirostane

1037   saponins from *Agave brittoniana* Trel. spp. *Brachypus*. *Magn Reson Chem*

1038   2007;45:615–620.

1039   56   Lajis NH, Abdullah ASH, Salim SJS, et al. Epi-Sarsasapogenin and epi-

1040   smilagenin: two sapogenins isolated from the rumen content of sheep intoxicated

1041   by *Brachiaria decumbens*. *Steroids* 1993;58:387–389.

1042   57   Eskander J, Lavaud C, Harakat D. Steroidal saponins from the leaves of *Agave*

1043   *macroacantha*. *Fitoterapia* 2010;81:371–374.

1044   58   Simmons-Boyce JL, Tinto WF, McLean S, et al. Saponins from *Furcraea selloa*

1045   var. *marginata*. *Fitoterapia* 2004;75:634–638.

1046   59   da Silva BP, Parente JP. A New Bioactive Steroidal Saponin from *Agave shrevei*.

1047   *Zeitschrift für Naturforsch C* 2005;60:57–62.

1048   60   Wang M, Xu Z, Peng Y, et al. Two New Steroidal Saponins with Antifungal

1049   Activity from *Hosta plantaginea* Rhizomes. *Chem Nat Compd* 2016;52:1047–

1050   1051.

1051   61   Mendes TP, De Medeiros Silva G, Da Silva BP, et al. A new steroidal saponin

1052   from *Agave attenuata*. *Nat Prod Res* 2004;18:183–188.

1053   62   Xiao H-H, Lv J, Mok D, et al. NMR Applications for Botanical Mixtures: The

1054   Use of HSQC Data to Determine Lignan Content in Sambucus williamsii. *J Nat*

1055   *Prod* 2019;82:1733–1740.

1056   63   Mina S, Melek FR, Abdel-khalik SM, et al. Pharmacological Activities of *Agave*

1057   *seemanniana* and Isolation of Three Steroidal Saponins. *European J Med Plants*

1058   2014;4:271–283.

1059   64   Zou P, Fu J, Yu H, et al. The NMR studies on two new furostanol saponins from

1060   *Agave sisalana* leaves. *Magn Reson Chem* 2006;44:1090–1095.

1061   65   Yokosuka A, Mimaki Y, Kuroda M, et al. A new steroidal saponin from the

1062   leaves of *Agave americana*. *Planta Med* 2000;66:393–396.

1063

1064

1065

1066   **Figure legends**

1067   **Figure 1** Representative structure of a steroidal saponin.

1068   **Figure 2** ^{1}H and ^{13}C NMR signals influenced by functionalization at ring F of *Agave*

1069   saponin aglycones.

1070   **Figure 3** ^{1}H and ^{13}C NMR signals influenced by functionalization at ring C of *Agave*

1071   saponin aglycones.

1072   **Figure 4** ^{1}H and ^{13}C NMR signals influenced by functionalization at ring B of *Agave*

1073   saponin aglycones.

1074   **Figure 5** ^{1}H and ^{13}C NMR signals influenced by functionalization at ring A of *Agave*

1075   saponin aglycones.

1076   **Figure 6** ^{1}H and ^{13}C NMR signals influenced by sugar chains at ring A of *Agave*

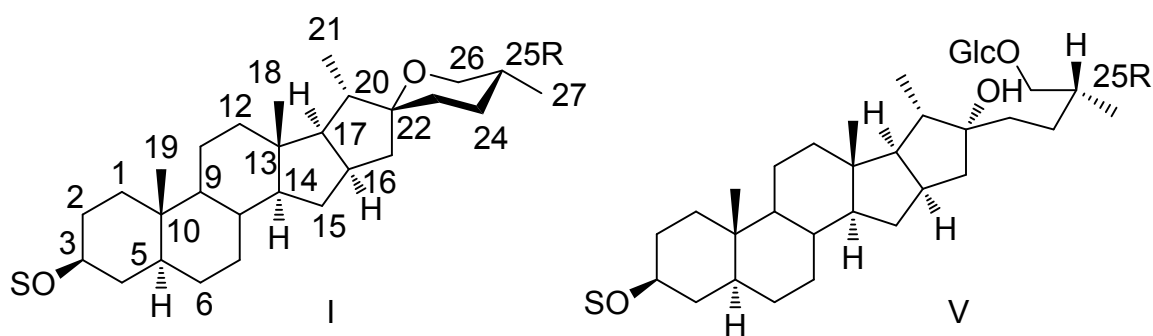1077   saponin aglycones.

1078   **Figure 7** Flowchart for the HMBC method for aglycone identification (HMAI) of

1079   saponins from *Agave* species for doublet signals.

1080   **Figure 8** Flowchart for the HMBC method for aglycone identification (HMAI) of

1081   saponins from *Agave* species for singlet signals.

1082   **Figure 9** ^{1}H NMR and HMBC data with representation of HMAI decisions applied.

1084

1085 **Table 1** Saponin aglycones described in this review.

1086
1087

	C-2	C-5	C-6	C-9	C-12	C-23	C-24	C-22	C-25	Saponins
I		α						SP	R	**1-3**[45]
II		β						SP	R	**4-8**[45,54]
III		DB						SP	R	**9-10**[11,55]
IV		β						SP	DB	**11-12**[47]
V		α						F	R	**13-14**[63]
VI		β						F	S	**15-17**[54]
VII		α			CO			SP	R	**18-19**[48,52]
VIII		β			CO			SP	R	**20-23**[46,54]
IX		β			CO			SP	S	**24-25**[46]
X		DB			CO			SP	R	**26-27**[11]
XI		α		DB	CO			SP	R	**28-29**[52,53]
XII		β			CO			SP	DB	**30-31**[47]
XIII		DB			CO			F	R	**32-33**[11]
XIV		DB			CO			F	S	**34**[64]
XV	OHα	α						SP	R	**35-37**[45,52]
XVI	OHβ	β						SP	R	**38-39**[54]
XVII	OHα	DB						SP	R	**40-42**[11,55]
XVIII	OHα	DB						F	R	**43-44**[51]
XIX		α	OHα					SP	R	**45-46**[11,44]
XX		α			OHβ			SP	R	**47**[48]
XXI		β			OHβ			SP	DB	**48-49**[47]
XXII	OHα	α			CO			SP	R	**50-51**[10,11]
XXIII	OHα	DB			CO			SP	R	**52-54**[10,11]
XXIV	OHα	α		DB	CO			SP	R	**55-56**[10,52]
XXV		α	OHα		CO			SP	R	**57**[65]
XXVI		α	OHα			OHα		SP	R	**58-59**[48]
XXVII		α				OHα	OHβ	SP	R*	**60**[50]
XXVIII	OHα	DB				OHβ		SP	R	**61**[10]
XXIX	OHα	DB					OHβ	SP	R*	**62**[49]
XXX		α	OHα			OHα	OHβ	SP	R*	**63**[50]

1088
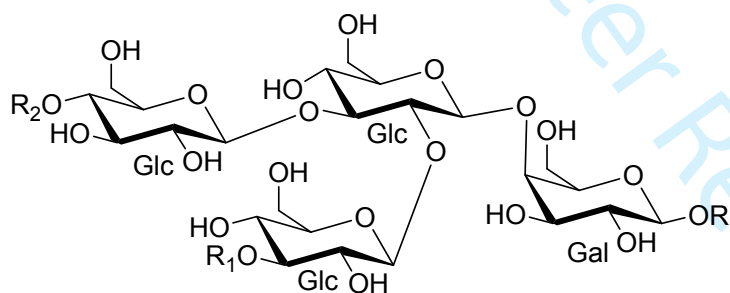
1089 OH: hydroxyl; DB: double bond; CO: carbonyl; SP: spirostanic; F: furostanic; R/S/α/β: chiral center
1090 configuration. *R is the relative configuration.

Table 2 Sugar chains of saponins described in this review.
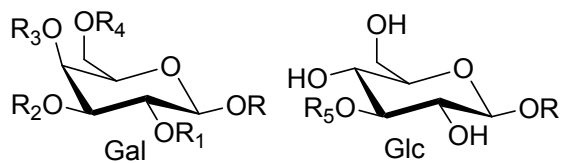
Name	R$_1$	R$_2$	R$_3$	R$_4$	R$_5$
S2A	Glc	H	H	H	
S2B	H	Glc	H	H	
S2C	H	H	Glc	H	
S2D	H	H	H	Glc	
S2E					Glc
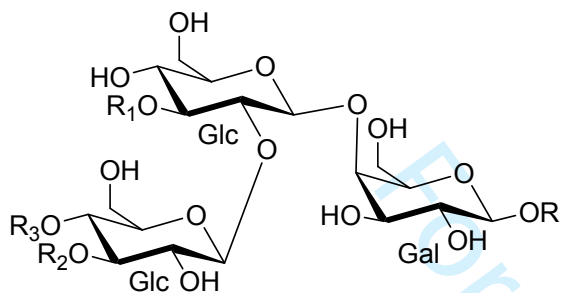
Name	R$_1$	R$_2$	R$_3$
S3A	H	H	H
S4A	Xyl	H	H
S5A	Xyl	Xyl	H
S5B	Xyl	Rha	H
S5C	Xyl	Glu	H
S5D	Xyl	H	Rha

Name	R$_1$	R$_2$
S4B	H	H
S5E	H	Rha
S6A	Glu	Rha

1091

Table 3 Influence of structure and functionalization of *Agave* saponin aglycons on ^{1}H and ^{13}C NMR chemical shifts. F Ring.

	I	IX	IV	XXVI	XXIX	XXVII	XVIII		VI	XVIII
Functionalization of F ring.										
Aglycon	I	IX	IV	XXVI	XXIX	XXVII	XVIII		VI	XVIII
C-22	SP	SP	SP	SP	SP	SP	F		F	F-OM
C-23				OHα		OHα				
C-24					OHβ	OHβ				
C-25	R	S	DB	R	R*	R*	R		S	R
^{13}C NMR chemical shifts (ppm. bold numbers) or differences.										
Saponin	**1**	**24**	**11**	**58**	**62**	**60**	**43**	**43**	**34**	**44**
16	**81.0**	-1.1[b]		0.7	0.5	1.0[a]		**81.2**		
17	**62.9**	-8.7[b]		-0.3	-0.6	-0.9[a]	1.0	**63.9**		
20	**41.9**	1.2[b]		-6.0		-7.3	-1.1	**40.8**		
21	**14.9**	-1.1[b]					1.6	**16.5**		
22	**109.3**	0.5		2.4	2.3	3.4	1.4	**110.7**		2
23	**31.6**	-5.2	1.7	67.5	9.2	**71.4**	5.7	**37.3**		-6.5
24	**29.1**	-2.9		9.8	81.5	87.9	-0.6	**28.5**		
25	**30.6**	-3.1	144.4	1.2	7.6	7.3	3.8	**34.4**		
26	**66.7**	-1.5	-1.7	-0.7	-1.6	-2.6[a]	8.6	**75.3**		
27	**17.2**	-0.9	108.7		-3.7	-4.0[a]		**17.5**		-0.4
^{1}H NMR chemical shifts (ppm. bold numbers) or differences.										
Saponin	**1**	**24**	**11**	**58**	**62**	**60**	**43**	**43**	**34**	**44**
16	**4.53**	-0.04[b]	0.05				0.38	**4.91**		-0.50
17	**1.77**	1.00[b]	0.06	0.08		0.08[a]	0.13	**1.9**		-0.20
18	**0.8**	0.26[b]		0.16	-0.09	0.21	0.05	**0.85**		-0.08
20	**1.93**	-0.07[b]		1.07		1.08[a]	0.27	**2.2**		
21	**1.12**	0.23[b]	-0.04	0.04	-0.10	0.03	0.18	**1.3**		-0.14
23a	**1.62**	-0.24	0.14	**3.82**	0.31	**3.84**	0.35	**1.97**		-0.20
23b	**1.67**	0.13	0.09		0.96		0.34	**2.01**		
24a	**1.54**	-0.23	0.68	0.21	**4.01**	**3.96**	0.11	**1.65**		0.13
24b	**1.54**	0.56	1.15	0.54			0.47	**2.01**		-0.23
25	**1.56**			0.24	0.30	0.49[a]	0.34	**1.9**		
26a	**3.48**	-0.13	0.53	-0.04	0.06	0.12	0.11	**3.59**	-0.11	
26b	**3.56**	0.46	0.89	-0.04	0.05	0.04	0.36	**3.92**	0.15	
27a	**0.67**	0.38	**4.76**	0.05	0.45	0.52[a]	0.29	**0.96**	0.05	
27b			**4.79**							

1094

OH: hydroxyl; DB: double bond; SP: spirostanic; F: furostanic; R/S/α/β: chiral center configuration. *R is the relative configuration.

Difference values are obtained from data for ring F. C-25R. spirostanic or furostanic aglycone. Error data: ^{13}C NMR: ±0.4 ppm; ^{1}H NMR: methylene signals ±0.1 ppm. others ±0.04 ppm.

[a] These differences are in agreement with additive effects of C-23 and C-24 hydroxyl groups and [b] additive effects of C-12 carbonyl group and C-25S.

1101

1102

Table 4 Influence of structure and functionalization of *Agave* saponin aglycones on ^{1}H and ^{13}C NMR chemical shifts. C Ring.

	Functionalization				
Aglycone	**I**	**VII**	**XIII**	**XX**	**XI**
C-5	α	α	DB	α	α
C-11					DB
C-12		CO	CO	OHα	CO
C-22	SP	SP	F	SP	SP
C-25	R	R	R	R	R
^{13}C NMR chemical shifts (ppm, bold numbers) or differences.					
Saponin	**1**	**18**	**32**	**47**	**28**
8	**35.1**	-0.8	-4.2^b	-0.7	1.8
9	**54.1**	1.4	-1.7^b	-0.5	**171.30**
10	**35.6**	0.7	2.0^b		3.9
11	**21.1**	16.9	16.5	10.6	**120.00**
12	**40**	**212.8**	**212.9**	79.3	**204.30**
13	**40.5**	14.9	14.9	6.1	10.8
14	**56.1**			-0.9	-3.4
15	**32**	-0.6	-0.1		
16	**81**	-1.3	-1.2		-0.8
17	**62.9**	-8.6	-8.0^a		-8.4
18	**16.5**	-0.4	-0.4	-5.2	-1.3
19	**12.2**	-0.5	6.7^b		6.1
20	**41.9**	0.7	-0.6^a	1.2	1.1
21	**14.9**	-1.0	0.4^a	-0.5	-1.2
^{1}H NMR chemical shifts (ppm, bold numbers) or differences.					
Saponin	**1**	**18**	**32**	**47**	**29**
8	**1.39**	0.31	0.42^b		0.97
9	**0.47**	0.38	0.81^b	0.15	
11a	**1.16**	1.02	1.12^b	0.31	**5.76**
11b	**1.36**	0.97	1.16^b	0.44	
12a	**1.01**			3.48	
12b	**1.63**				
14	**0.99**	0.34	0.41^b	0.04	0.71
15a	**1.38**	0.16	0.22	0.17	0.27
15b	**2**	0.06	0.06	0.07	0.15
16	**4.53**	-0.08	0.32^a	0.07	-0.04
17	**1.77**	0.95	1.16^a	0.38	0.85
18	**0.8**	0.23	0.33^a	0.26	0.19
19	**0.62**		0.29^b		0.17
20	**1.93**	-0.05	0.26^a	0.25	0.05
21	**1.12**	0.19	0.41^a	0.29	0.27

OH: hydroxyl; DB: double bond; CO: carbonyl; SP: spirostanic; F: furostanic; R/S/α/β: chiral center configuration.

Difference values are obtained from data for 25R-spirostanic aglycone. Error data: ^{13}C NMR: ±0.4 ppm; ^{1}H NMR: methylene signals ±0.1 ppm, others ±0.04 ppm.

^{a} These differences are in agreement with additive effects of C-12 carbonyl group and furostanic structure and ^{b} additive effects of C-12 carbonyl group and C-5 double bond.

Table 5 Influence of structure and functionalization of *Agave* saponin aglycones on ^{1}H and ^{13}C NMR chemical shifts. B Ring.

Aglycone	I	XIX	XIX	XI	II	VIII	XXI	III	X
Functionalization									
C-5	α	α	α	α	β	β	β	DB	DB
C-6		OHα	OGlcα						
C-9				DB					
C-12					CO		OHβ		CO
^{13}C NMR chemical shifts (ppm, bold numbers) or differences.									
Saponin	**1**	**45**	**46**	**28**	**4**	**22**	**48**	**9**	**26**
1	**36.9**			-1.9	-6.1	-6.5	-5.9	0.6	0.1
2	**29.8**				-2.8	-3.1	-2.9	0.4	0.1
3	**78.1**			-1.4	-3.4	-3.8	-3.8		
4	**34.7**	-5.3	-6.1		-4	-4.1	-4.2	4.6	4.4
5	**44.4**	7.7	6.5	-1.9	-7.5	-7.9	-7.6	**141.1**	**140.9**
6	**28.8**	68.3	79.9	-0.9	-1.8	-2.1	-1.7	**121.7**	**121.5**
7	**32.3**	10.3	9.1		-5.6	-6	-5.6		-0.5
8	**35.1**	-0.9	-1	1.8	0.4	-0.4[a]	-0.4[b]	-3.4	-4.2[a]
9	**54.1**			171.30	-13.9	-12.2[a]	-14.7[b]	-3.8	-1.8[a]
10	**35.6**	0.8	1.1	3.9	-0.4	0[a]	-0.3[b]	1.5	2[a]
14	**56.1**			-3.4			-0.8[b]	0.6	
19	**12.2**	1.2	1.2	6.1	11.7	10.8[a]	11.6[b]	7.3	6.7[a]
^{1}H NMR chemical shifts (ppm, bold numbers) or differences.									
Saponin	**1**	**45**	**46**	**29**	**4**	**22**	**48**	**9**	**26**
1a	**0.75**			0.40	0.69	0.47	0.68	0.18	0.10
1b	**1.47**				0.32	0.15	0.27	0.18	
2a	**1.60**				-0.18	-0.25	-0.10	0.08	
2b	**2.02**				-0.13	-0.23	-0.15	0.07	
3	**3.89**				0.40	0.35	0.43		-0.08
4a	**1.32**	0.15	0.12		0.37	0.34	0.40	1.08	1.06
4b	**1.77**		1.62			-0.07		0.87	0.88
5	**0.86**	0.31	0.37	0.17	1.19	1.12	1.15		
6a	**1.04**	**3.55**	**3.59**	0.12				**5.27**	**5.25**
6b	**1.08**			0.08	0.63	0.64	0.65		
7a	**0.75**	0.39		0.12	0.18	0.17	0.20	0.69	0.69
7b	**1.48**	0.70	1.09	0.23	-0.25	-0.18	-0.17	0.33	0.36
8	**1.39**	0.19	0.13	0.97	0.07	0.41[a]	0.13	0.09	0.42[a]
9	**0.47**	0.13			0.80	1.24[a]	0.99[b]	0.38	0.81[a]
11a	**1.16**					1.02[a]	0.31[b]	0.19	1.10[a]
11b	**1.36**				-0.07	0.98[a]	0.39[b]	0.08	1.14[a]
12a	**1.01**				0.06		**3.51**		
12b	**1.63**								
14	**0.99**			0.71	0.07	0.44[a]	0.14[b]		0.41[a]
15a	**1.38**			0.24		0.19[a]	0.20[b]		0.20[a]
15b	**2.00**			0.15			0.08[b]		0.08[a]
16	**4.53**		-0.13	-0.02	0.05	-0.0[a]	0.14[b]		-0.07[a]
17	**1.77**			0.85	0.06	1.02[a]	0.43[b]		1.01[a]
18	**0.80**			0.19		0.26[a]	0.26[b]		0.27[a]
19	**0.62**	0.06	0.06	0.17	0.22	0.18[a]	0.22	0.23	0.28

OH: hydroxyl; DB: double bond; CO: carbonyl; SP: spirostanic; F: furostanic; R/S/α/β: chiral center configuration.
Difference values are obtained from data for 25R-spirostanic aglycone. Error data: ^{13}C NMR: ±0.4 ppm; ^{1}H NMR: methylenes signals ±0.1 ppm, others ±0.04 ppm.
[a] These differences are in agreement with additive effects of C-12 carbonyl group and H-5β or Δ[5]-spirostanol and [b] additive effects of C-12 hydroxyl group H-5β-spirostanol.

1120

1121 **Table 6** Influence of structure and functionalization of *Agave* saponin aglycones on ^{1}H and ^{13}C NMR

1122 chemical shifts. A Ring.

|	**Functionalization**									
Sugar moiety	S6A	S3A	S5A	S5A	S3A	S3A	S5A	S5A	S5A	S4A
Aglycon	**I**	**XV**	**XXII**	**XXIV**	**II**	**XVI**	**III**	**XVII**	**XXIII**	**XXVIII**
C-2		OHα	OHα	OHα		OHβ		OHα	OHα	OHα
C-5	α	α	α	α	β	β	DB	DB	DB	
C-9				DB						
C-12			CO	CO					CO	OHβ
^{13}C NMR chemical shifts (ppm, bold numbers) or differences.										
Saponin	**1**	**35**	**50**	**55**	**4**	**38**	**9**	**40**	**53**	**61**
1	**36.9**	8.7	8.2	6.6[a]	**30.8**	9.7	**37.5**	8.3	7.8	8.3
2	**29.8**	70.3	70.2	70.3	**27.0**	79.7	**30.2**	70.1	69.8	**70.0**
3	**78.1**	6.5	5.9	5.4[a]	**74.7**	5.0	**78.3**	6.3	6.0	6.3
4	**34.7**	-0.6	-0.8	-1.0[a]	**30.7**	1.3	**39.3**	-1.6	-1.8	-1.7
5	**44.4**			-1.9[a]	**36.9**	-0.5	**141.1**	-1.0	-1.2	-1.0
6	**28.8**	-0.8	-1.0	-1.6[a]	**27.0**	-0.8	**121.7**			
7	**32.3**		-0.7		**26.7**		**32.3**		-0.6	
8	**35.1**	-0.6	-1.4[a]	1.1[a]	**35.5**		**31.7**	-0.6	-1.3[a]	-1.3[b]
9	**54.1**	0.2	1.3[a]	**170.5**	**40.2**	1.2	**50.3**		1.8[a]	-0.3[b]
10	**35.6**	1.2	2.5[a]	5.0[a]	**35.2**	1.7	**37.1**	0.9	1.3[a]	1.0[b]
19	**12.2**	1.2	0.7[a]	7.2[a]	**23.9**	-0.1	**19.5**	1.0	0.4[a]	1.0[b]
^{1}H NMR chemical shifts (ppm, bold numbers) or differences.										
Saponin	**1**	**35**	**50**	**55**	**4**	**38**	**9**	**40**	**53**	**61**
1a	**0.75**	0.37	0.34	0.80[a]	**1.44**	0.32	**0.93**	0.34	0.29	0.35
1b	**1.47**	0.70	0.53	0.74[a]	**1.79**	0.11	**1.65**	0.64	0.49	0.64
2a	**1.60**	**3.95**	**3.89**	**4.00**	**1.42**	**3.85**	**1.68**	**4.06**	**3.99**	**4.00**
2b	**2.02**				**1.89**		**2.09**			
3	**3.89**	-0.05	-0.07	-0.06	**4.29**	0.14	**3.87**	-0.06	-0.09	-0.08
4a	**1.32**	0.15	0.12	0.17	**1.69**	0.14	**2.40**	0.13	0.11	0.11
4b	**1.77**	0.05	0.07	0.14	**1.80**	0.11	**2.64**	0.04	0.07	0.05
5	**0.86**	0.11	0.07	0.29[a]	**2.05**	-0.03				
6a	**1.04**	-0.05		0.19[a]	**1.03**	0.03	**5.27**			
6b	**1.08**	0.02	0.04	0.08[a]	**1.71**	-0.06				
7a	**0.75**			0.12[a]	**0.93**		**1.44**			
7b	**1.48**			0.25[a]	**1.23**		**1.81**			
8	**1.39**		0.31[a]	0.97[a]	**1.46**	-0.04	**1.48**		0.30[a]	
9	**0.47**	0.09	0.49[a]		**1.27**	-0.08	**0.85**	0.07	0.51[a]	0.25[b]
11a	**1.16**	0.02	1.16[a]		**1.16**		**1.35**		1.02[a]	0.32[b]
11b	**1.36**	0.08	1.02[a]		**1.29**	0.15	**1.44**		1.08[a]	0.49[b]
19	**0.62**	0.07	0.08	0.24[a]	**0.84**	0.03	**0.85**	0.06	0.10	0.10

1123

1124 OH: hydroxyl; DB: double bond; CO: carbonyl; α/β: chiral center configuration.

1125 Difference values are obtained from data for aglycones I, II and III. Error data: ^{13}C NMR: ±0.4 ppm; ^{1}H NMR:

1126 methylene signals ±0.1 ppm, others ±0.04 ppm.

1127 [a] These differences are in agreement with additive effects of C-12 carbonyl group and spirostan-2,3-diols and [b] additive

1128 effects of C-12 hydroxyl group and spirostan-2,3-diols H-5β-spirostanol.

1129 **Table 7** HMBC correlations with doublet signals of methyl groups C-21 and C-27.
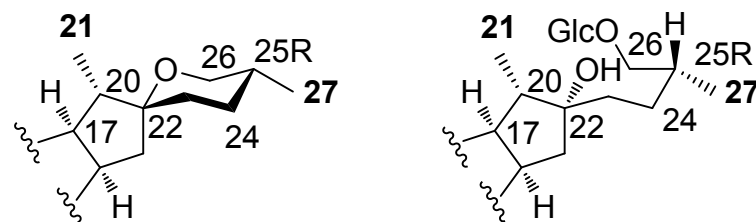
1130

| Structural Features | | | | | | HMBC signals | | | | | | | | |
C-9	C-12	C-23	C-24	C-22	C-25	**H-27**	C-24	C-25	C-26	**H-21**	C-17	C-20	C-22	Data
				SP	R	0.67	29.1	30.6	66.7	1.12	62.9	41.9	109.3	**1**
				SP	DB	-	-	-	-	1.08	63.2	41.9	109.4	**11**
				F	R	0.96	28.5	34.4	75.3	1.30	63.9	40.8	110.7	**43**
				FM	R	0.98	28.2	34.2	75.2	1.16	64.1	40.5	112.7	**44**
		OHα		SP	R	0.72	38.9	31.8	66.0	1.16	62.6	35.9	111.7	**58**
			OGlcβ	SP	R*	1.12	81.5	38.2	65.1	1.02	62.3	42.1	111.6	**62**
		OHα	OGlcβ	SP	R*	1.19	87.9	37.9	64.1	1.15	62.0	34.6	112.7	**60**
	CO			SP	S	1.05	26.2	27.5	65.2	1.35	54.2	43.1	109.8	**24**
	CO			SP	R	0.64	29.2	30.5	66.9	1.31	54.3	42.6	109.3	**18**
	CO			F	R	0.96	28.4	34.3	75.3	1.53	54.9	41.3	110.9	**32**
	CO			F	S	1.01	28.3	34.4	75.3	1.51	54.8	41.3	110.8	**34**
DB	CO			SP	R	0.67	29.2	30.5	67.0	1.38	54.5	43.0	109.5	**28**
	OHβ			SP	R	0.67	29.4	30.7	66.9	1.41	63.0	43.1	109.6	**47**

1131

1132 OH: hydroxyl; DB: double bond; CO: carbonyl; SP: spirostanic; F: furostanic; R/S/α/β: chiral center configuration.

1133 * R is the relative configuration; S is the absolute configuration because a glucopyranosyloxy moiety is at C-24.

Table 8 HMBC correlations with doublet signals of methyl groups C-18 and C-19.
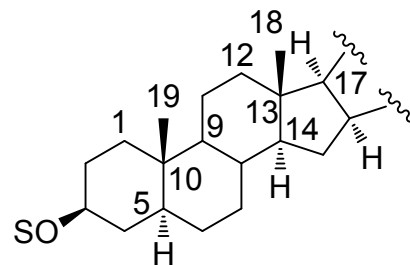
Structural Features								HMBC signals										
C-2	C-5	C-6	C-9	C-12	C-22	C-23	C-24	**H-18**	C-12	C-13	C-14	C-17	**H-19**	C-1	C-5	C-9	C-10	Data
	α				SP			0.80	40.0	40.5	56.1	62.9	0.62	36.9	44.4	54.1	35.6	**1**
	α				SP	OHα	OGlcβ	1.01	40.7	41.4	56.6	62	0.75	37.5	45.6	54.6	35.9	**60**
	α			CO	SP			1.03	212.8	55.4	55.9	54.3	0.61	36.6	44.4	55.5	36.3	**18**
	α		DE	CO	SP			0.98	204.3	51.3	52.7	54.5	0.79	35.0	42.5	171.3	39.5	**28**
OH α	α		DE	CO	SP			0.97	204.3	51.4	52.7	54.6	0.86	43.5	42.5	170.5	40.6	**55**
	α			OHβ	SP			1.06	79.3	46.6	55.2	63.0	0.64	37.2	44.7	53.6	35.9	**47**
OH α	α				SP			0.78	40.0	40.6	56.3	63.0	0.69	45.6	44.6	54.3	36.8	**35**
	β				SP			0.79	40.3	40.9	56.5	63.1	0.84	30.8	36.9	40.2	35.2	**4**
OH β	β				SP			0.77	40.2	40.8	56.3	63.1	0.87	40.5	36.4	41.4	36.9	**38**
	DB				SP			0.80	39.9	40.5	56.7	62.9	0.85	37.5	141.1	50.3	37.1	**9**
	DB			CO	F			1.13	212.9	55.4	56.0	54.9	0.91	37.0	140.9	52.4	37.6	**32**
OH α	DB				SP			0.78	39.8	40.5	56.5	62.9	0.91	45.8	140.1	50.2	38.0	**40**
OH α	DB				SP		OGlcβ	0.71	39.7	40.4	56.5	62.3	0.91	45.7	140.1	50.1	37.9	**62**
OH α	DB				F			0.85	39.9	40.8	56.5	63.9	0.92	45.8	140.1	50.3	38.0	**43**
OH α	DB				FM			0.77	39.6	40.8	56.4	64.1	0.91	45.7	140.1	50.2	37.9	**44**
	α	OHα			SP			0.81	40.0	40.7	56.2	62.9	0.68	37.6	52.1	54.0	36.4	**45**
	α	OGlcα			SP			0.76	40.0	40.8	56.4	63.0	0.68	37.5	50.9	53.8	36.7	**46**
	α	OGlcα			SP		OHα	0.96	22.7	41.4	56.5	62.6	0.74	37.8	51.3	54.0	36.8	**58**

OH: hydroxyl; DB: double bond; CO: carbonyl; SP: spirostanic; F: furostanic; α/β: chiral center configuration.

2000

2001   **Table 9** HMBC correlations observed for saponin mixture.

2002

¹H NMR signal	HMBC correlations				Methyl assignment	Aglycone assignment
1.38 d	42.9	54.5	109.5		C-21	**XI**
1.33 d	42.6	54.3	109.3		C-21	**VII**
1.32 d	overlapped	overlapped	overlapped		C-21	**XXII**
1.05 s	54.2	55.5	212.7		C-18	**VII**
1.04 s	overlapped	overlapped	overlapped		C-18	**XXII**
0.98 s	51.3	52.6	54.5	204.3	C-18	**XI**
0.80 s	35.0	39.5	42.5	171.3	C-19	**XI**
0.72 s	37.2	44.7	55.3		C-19	**XXII**
0.66 d	29.2	30.5	66.9		C-27	**VII; XI; XXII**
0.64 s	36.4	44.4	55.5		C-19	**VII**

2003

56
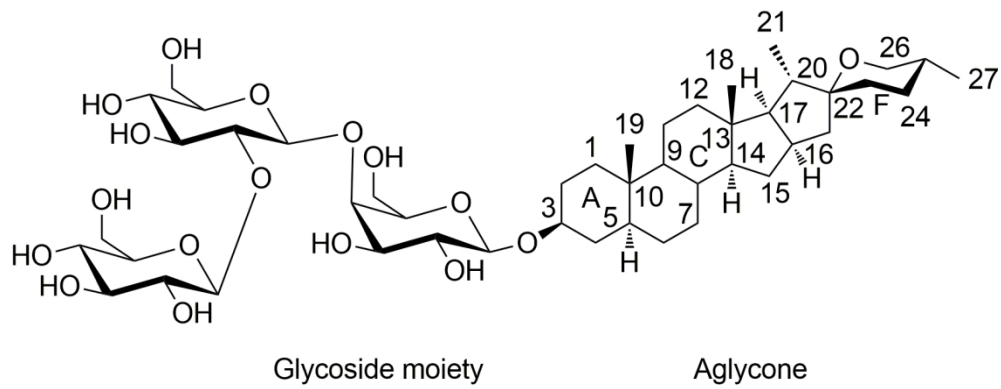
Figure 1 Representative structure of a steroidal saponin.

152x59mm (300 x 300 DPI)
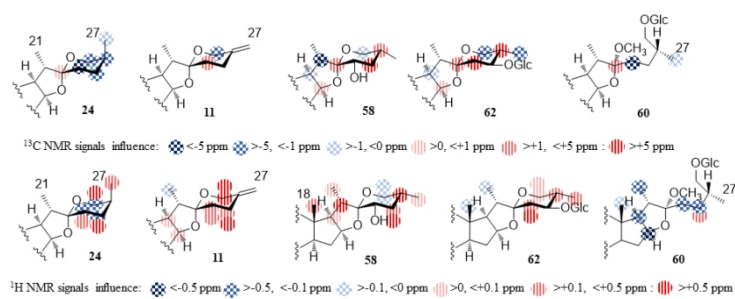
Figure 2 1H and 13C NMR signals influenced by functionalization at ring F of Agave saponin aglycones.

338x190mm (96 x 96 DPI)
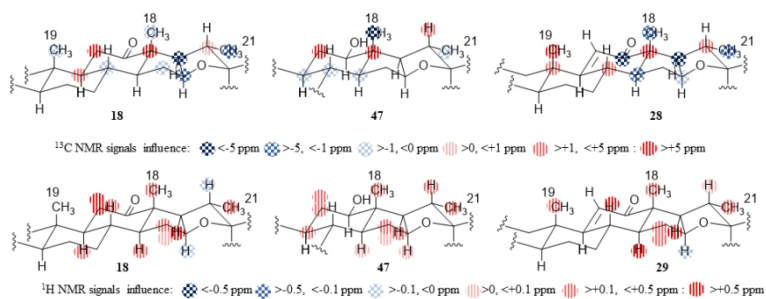
Figure 3 1H and 13C NMR signals influenced by functionalization at ring C of Agave saponin aglycones.
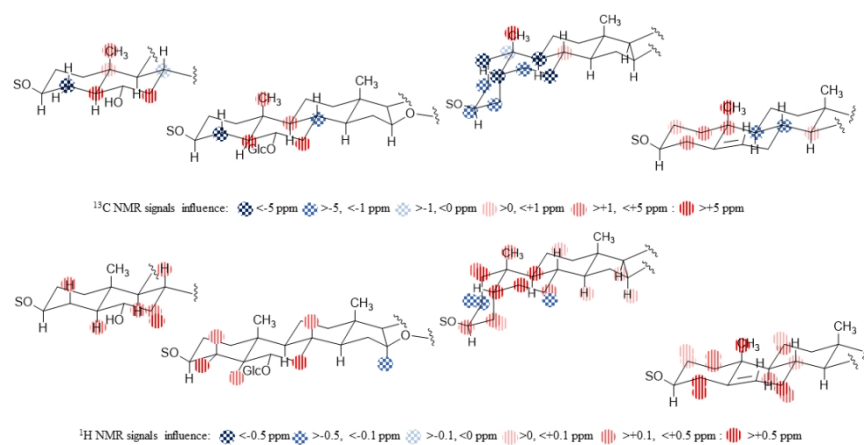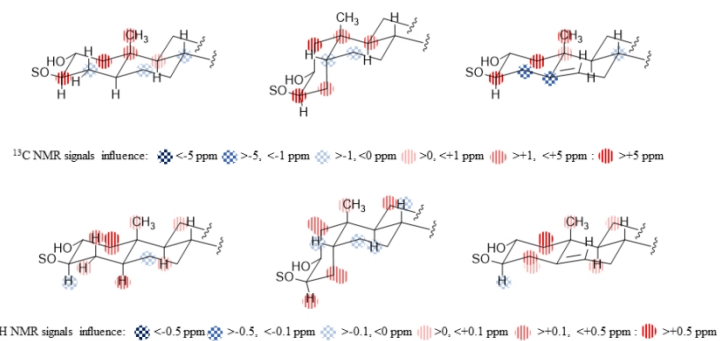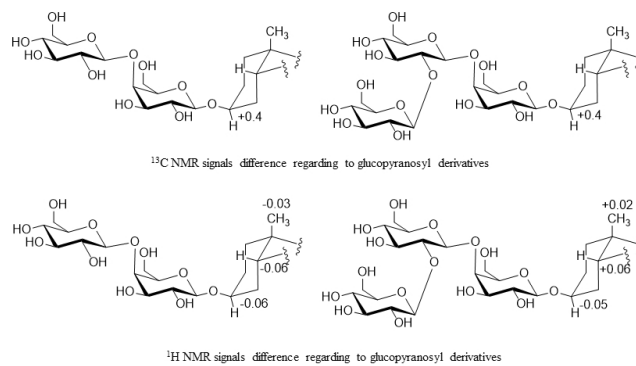
338x190mm (96 x 96 DPI)

Figure 4 1H and 13C NMR signals influenced by functionalization at ring B of Agave saponin aglycones.

338x190mm (96 x 96 DPI)

Figure 5 1H and 13C NMR signals influenced by functionalization at ring A of Agave saponin aglycones.

338x190mm (96 x 96 DPI)

Figure 6 1H and 13C NMR signals influenced by sugar chains at ring A of Agave saponin aglycones.
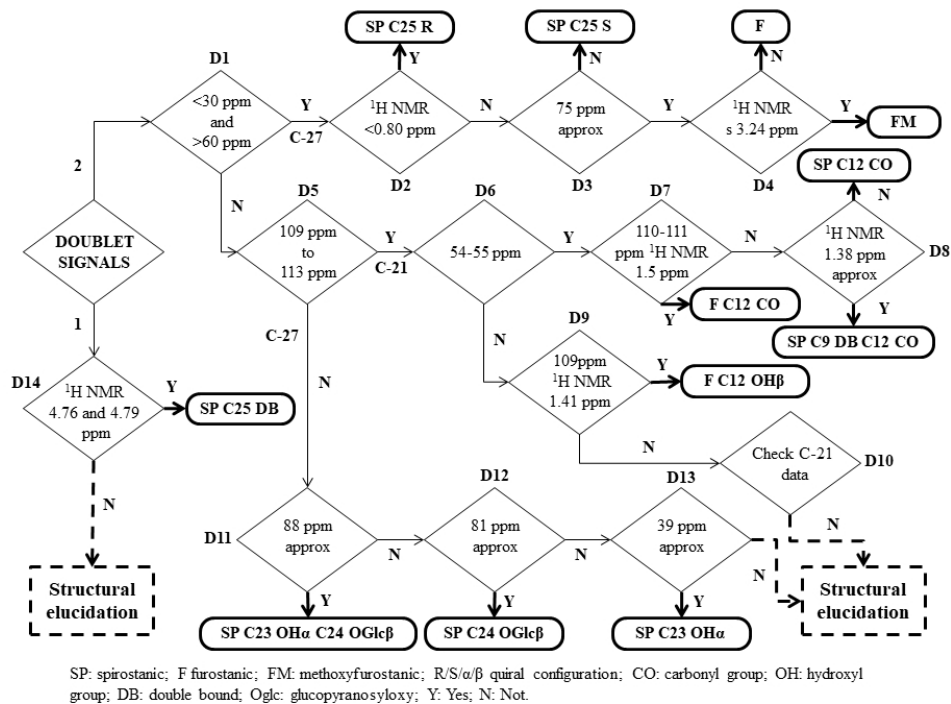
338x190mm (96 x 96 DPI)

Phytochemical Analysis

Figure 7 Flowchart for the HMBC method for aglycone identification (HMAI) of saponins from Agave species for doublet signals.
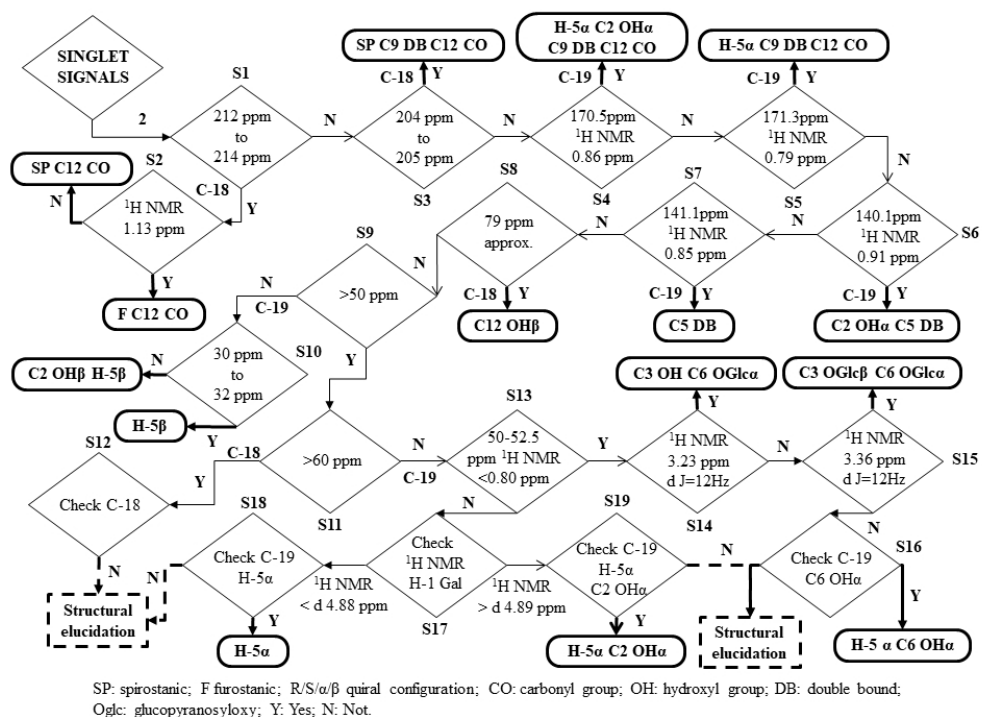
254x190mm (96 x 96 DPI)

Figure 8 Flowchart for the HMBC method for aglycone identification (HMAI) of saponins from Agave species for singlet signals.
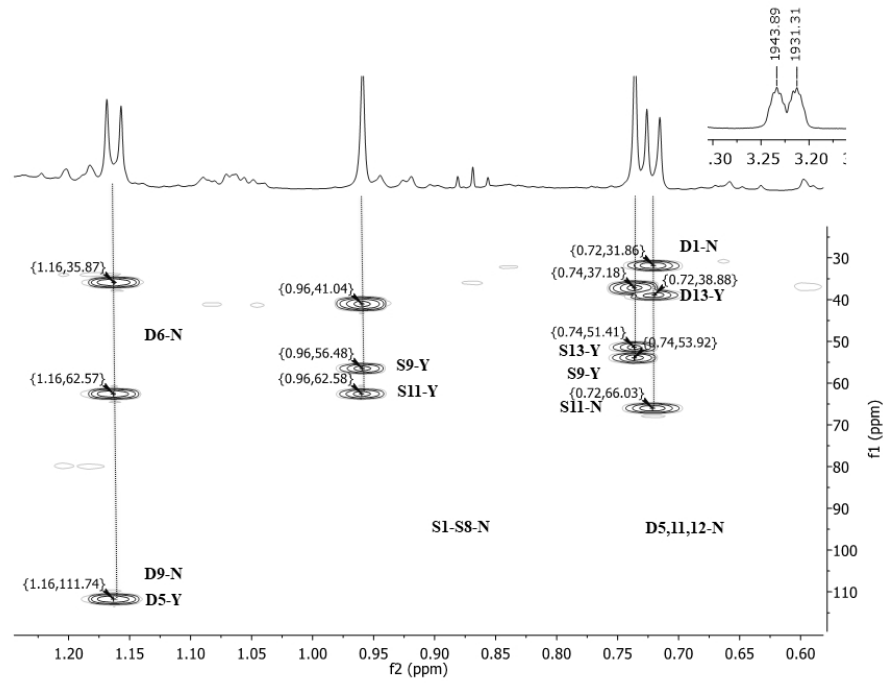
254x190mm (96 x 96 DPI)

Figure 9 1H NMR and HMBC data with a representation of HMAI decisions applied.

254x190mm (96 x 96 DPI)