Pure Appl. Chem., Vol. 84, No. 6, pp. 1469–1478, 2012.
http://dx.doi.org/10.1351/PAC-CON-11-09-12

Assignment of stereochemistry in open-chain steroidal saponins*

Victoria L. Challinor and James J. De Voss‡

School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane 4072, Australia

Abstract: The major phytochemical constituents of the widely used medicinal herb *Chamaelirium luteum* ("false unicorn") are open-chain cholesterol-derived steroidal saponins. These are unusual in lacking the extra ring(s) derived from the steroidal side chain that are characteristic of the more commonly isolated furostanol and spirostanol saponins. The structures of the major steroidal saponins of *C. luteum* were determined using a combination of multistage mass spectrometry (MS^n), 1D and 2D NMR experiments, and chemical degradation. The flexible nature of the steroidal side chain in these saponins necessitated determination of their absolute stereochemistry via total synthesis and X-ray crystallography.

Keywords: natural products; steroidal saponins; structure elucidation; total synthesis.

STEROIDAL SAPONINS ISOLATED FROM MEDICINAL HERBS

Chamaelirium luteum (false unicorn)

Chamaelirium luteum (L.) A. Gray, commonly known as "false unicorn" or "helonias", is a herbaceous member of the Melanthiaceae family endemic to the eastern states of the USA and Ontario, Canada [1]. The underground rhizomes of *C. luteum* are used in traditional medicine for the treatment of a range of female reproductive health problems [2], and are also present in a number of commercial botanical dietary supplements. *C. luteum* is listed as endangered or threatened in a number of states in the USA [1], in part owing to over-collection of plants from the wild for medicinal use (wildcrafting) [3].

The phytochemical profile of *C. luteum* has received surprisingly limited attention, given the wide use of this herb in traditional medicine and its commercial use; the trade volume of *C. luteum* plant material for the USA alone was recently estimated at 9800–14 300 kg per annum [4]. Early primary literature postulated the presence of saponins, with the isolation in 1878 of the "bitter principle" of *C. luteum*, named "chamaelirin" [5], and its decomposition products, glucose and "chamaeliretin" [6]. Subsequently, Marker et al. reported the isolation of small amounts (0.02 % dry weight of plant material) of the steroidal aglycone diosgenin from *C. carolinianum* (synonymous with *C. luteum*) [7]. Since then, no work had been reported on the phytochemical constituents of *C. luteum*, despite its inclusion in a clinical trial [8] and an animal model study [9].

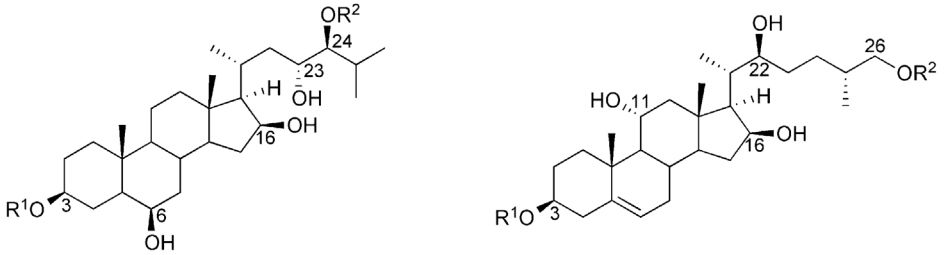
Steroidal saponins are common chemical constituents of medicinal herbs, and exhibit a wide range of biological activities, including cytotoxic, anti-inflammatory, hemolytic, antifungal, and antibacterial properties [10]. They consist of a steroidal (C_{27}) skeleton (aglycone), usually an oxidized cholesterol derivative, bearing multiple sugar residues at varying positions. The majority of steroidal

Pure Appl. Chem.* **84, 1297–1478 (2012). A collection of invited papers based on presentations at the 27th International Symposium on the Chemistry of Natural Products and the 7th International Conference on Biodiversity (ISCNP-27 & ICOB-7), Brisbane, Australia, 10–15 July 2011.
‡Corresponding author

saponins isolated from terrestrial plants fall into two structural subfamilies; the pentacyclic furostanol saponins, characterized by a hemiacteal moiety at C-22 and a glycosidic linkage at C-26, and the spirostanol saponins, which instead possess six rings formed via a bicyclic acetal at C-22. Our recent work on *C. luteum* [11,12] has led us to investigate a further, less common structural class of saponins, the open-chain cholesterol-derived steroidal glycosides. These lack the additional ring(s) derived from the C-17 side chain that are characteristic of furo- and spirostanol glycosides. Saponins of this type, while abundant in marine organisms such as starfish, sponges, and cnidarians [13–16], are relatively rare in terrestrial plants. With some exceptions (see, e.g., refs. [11,17–19]), open-chain steroidal glycosides are typically reported from plants as only minor constituents, along with saponins based on the more common furo- and spirostanol scaffolds. For example, the cholestene glycosides bethosides B and C were reported from the North American medicinal herb *Trillium erectum* L. (commonly known as "beth root") along with diosgenin- and pennogenin-derived saponins, and ecdysteroids [20].

Structure elucidation of major steroidal saponins from *C. luteum*

The major steroidal saponins of *C. luteum* all possess an open-chain cholesterol-derived aglycone. Chamaelirosides A and B (**1** and **2**) contain the (23*R*,24*S*)-chiograsterol B aglycone (**3**) [11], while heloside A (**4**) and a related minor constituent heloside B (**5**) are derived from the unique helogenin skeleton (**6**) [12] (Fig. 1). The structure and stereochemistry of these saponins was determined using a combination of multistage mass spectrometry (MS^n), 1D and 2D NMR experiments, chemical degradation, X-ray crystallography, and total synthesis.



1 Chamaeliroside A R^1 = β-D-Glc-(6)-β-D-Glc R^2 = H
2 Chamaeliroside B R^1 = β-D-Glc-(6)-β-D-Glc R^2 = β-D-Fuc
3 (23*R*,24*S*)-Chiograsterol B R^1, R^2 = H
4 Heloside A R^1, R^2 = β-D-Glc
5 Heloside B R^1 = β-D-Glc R^2 = H
6 Helogenin R^1, R^2 = H

Fig. 1 Major steroidal saponins isolated from *C. luteum* and their aglycones.

The process of structure elucidation of a steroidal saponin is exemplified by that used for heloside A (**4**), a bidesmodic glycoside containing a 3β,11α,16β,22,26-pentahydroxy-cholestene skeleton (**6**). Fragmentation by positive-ion electrospray ionization (ESI)-MS^n provided important information on the number and nature of the sugar units present in **4**. Along with the molecular formula of $C_{39}H_{66}O_{15}$ provided by high-resolution ESI-MS, the neutral loss of two 180 Da units from the quasi-molecular ion $[M + Na]^+$ at *m/z* 797 suggested the presence of two hexose monosaccharides, in addition to a C_{27} skeleton bearing five oxygens. The accompanying sugar units were identified as two β-glucose residues through examination of their vicinal coupling constants, which were obtained by selective irradiation of the sugar anomeric resonances in 1D total correlation spectroscopy (TOCSY) NMR experiments. The absolute configurations of the monosaccharides were determined by enantioselective GC analysis, in which the saponin **4** was subjected to acid-catalyzed methanolysis followed by per-trifluoroacetylation of the resultant methyl-glycosides, and comparison with authentic D- and L-sugar standards [21].

A helpful starting point for structure elucidation of the aglycone was the heteronuclear multiple-bond correlation (HMBC) correlations of the steroid methyl groups. Of all the HMBC correlations of the methyl groups, only those of H_3-18 (singlet) and H_3-21 (doublet, J = 7.0 Hz) intersect (at C-17), thus allowing rapid assignment of the four steroidal methyl signals and their neighboring groups (Fig. 2). Correlations observed between the sugar anomeric hydrogens and the steroid core carbons in the HMBC spectrum of **4** also established the sites of glycosidic linkages, revealing the attachment of the two β-D-glucose residues separately at positions 3 and 26.

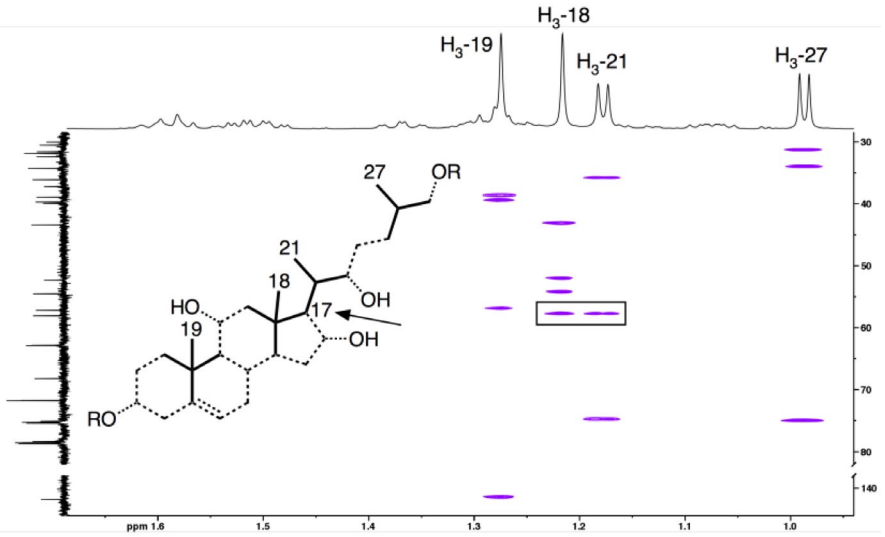


Fig. 2 Segment of the HMBC spectrum of **4**, showing correlation of both H_3-18 and H_3-21 with C-17 (R = β-D-glucose, C_5D_5N, 500 MHz).

The stereochemistry of the ring junctions and substituents in a steroidal aglycone may be determined using a 2D rotating frame Overhauser effect spectroscopy (ROESY) experiment. Correlations observed in the ROESY spectrum of **4** revealed the *trans* fusion of the steroid B/C and C/D rings, the β orientation of the substituents at C-3, C-16, and C-17, and the α orientation of the C-11 hydroxy group, while the configuration of C-20 was indicated by a cross-peak between H-20 and H-18 (Fig. 3). The flexible nature of C-20 to C-27 in helosides A (**4**) and B (**5**), as well as in chamaelirosides A (**1**) and B (**2**), meant that the stereochemistry along this side chain was unable to be determined spectroscopically, in contrast to typical furo- and spirostanol saponins [22,23].

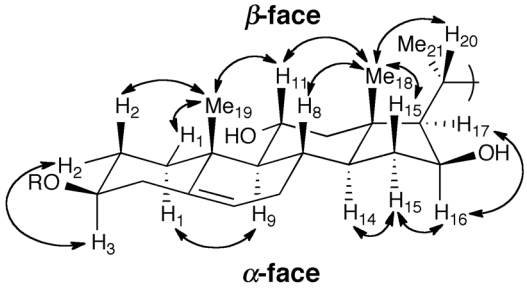
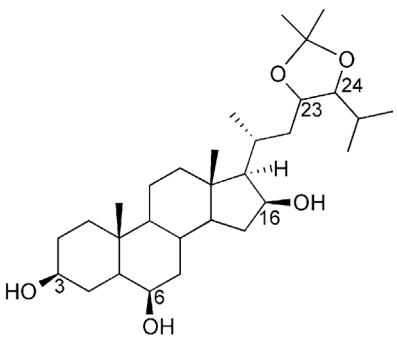


Fig. 3 Key ROESY correlations observed in heloside A (**4**).

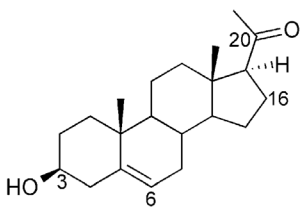
STEREOCHEMISTRY IN OPEN-CHAIN STEROIDAL SAPONINS

Chamaelirosides A and B

Chamaelirosides A (**1**) and B (**2**) contain an open-chain steroidal aglycone (**3**) that possesses the same planar structure as chiograsterol B, a minor metabolite isolated from *Chionographis japonica* in 1968 [24]. The absolute configuration of C-23 and C-24 in the side chain of chiograsterol B had not been defined, although the diol was suggested to have an *erythro* configuration, based upon the H-23/H-24 coupling constant in the corresponding acetonide derivative **7** [24]. Since only very limited ^{1}H NMR spectroscopic data were available in the literature for comparison [24], we were unable to draw any conclusion about the relative stereochemistry of **1** and **2**. We therefore synthesized simple *erythro* and *threo* acetonide models, and comparison of their ^{1}H NMR spectra with the acetonide **7** derived from the aglycone of **1** and **2** revealed that the *erythro* configuration was present in these saponins [11]. We thus set out to synthesize both *erythro* aglycone diastereomers, (23*R*,24*S*)-**3** and (23*S*,24*R*)-**3**, to confirm the structures and determine the absolute stereochemistry of the major saponins of *C. luteum*.



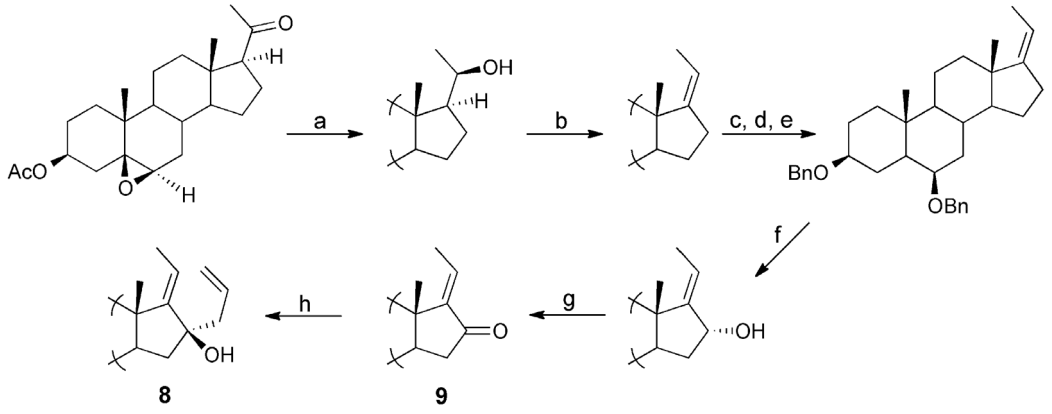
7



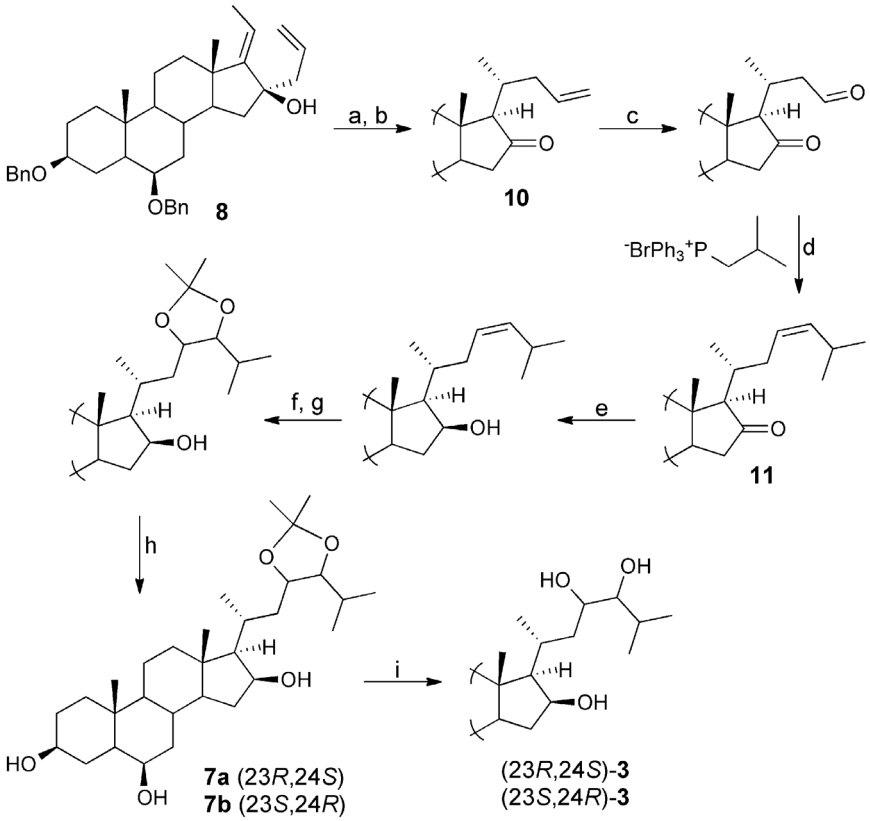
Pregnenolone

The commercially available steroid pregnenolone was identified as a suitable starting material, requiring β-hydroxylation at C-6 and C-16 and introduction of the side chain at C-20 to access the aglycone **3**. The key step in the synthesis was the oxy-anion accelerated Cope rearrangement of the tertiary allylic alcohol **8**, which was readily prepared as a single diastereomer from the β-epoxide of pregnenolone acetate (Scheme 1). We had investigated a number of other possibilities for installing the side chain, for example, conjugate addition to **9**, but none were successful.

The key Cope rearrangment of **8** proceeded in excellent yield to give keto-olefin **10** as a single diastereomer with the required (17*R*,20*R*) stereochemistry, in line with literature precedent [25] (Scheme 2). Subsequent conversion of **10** to the keto-aldehyde and Wittig condensation to form *Z*-alkene **11** completed the backbone of the steroidal side chain. Attempted enantioselective dihydroxylation of **11** under Sharpless conditions failed, but dihydroxylation of **11** under Upjohn conditions, followed by direct conversion to the diastereomeric *erythro* acetonides and removal of the benzyl protecting groups, afforded **7a** and **7b**. Crystallization of the major acetonide triol **7a** yielded crystals suitable for X-ray analysis, and the structure obtained revealed the (23*R*,24*S*) absolute configuration of the side chain. Comparison of the ^{1}H and ^{13}C NMR spectra of diastereomeric acetonides **7a** and **7b**, with the acetonide **7** derived from the aglycone of **1** and **2**, showed that the naturally occurring stereochemistry in these saponins is (23*R*,24*S*). Finally, acetal hydrolysis of **7a** and **7b** generated the diastereomeric pentahydroxy sterols (23*R*,24*S*)-**3** and (23*R*,24*S*)-**3**. A comparison of their physical properties with those reported previously for chiograsterol B [24] suggested that it too probably contains the (23*R*,24*S*) stereochemistry.



Scheme 1 *Reagents and conditions*: (a) $NaBH_4$, $MeOH/CH_2Cl_2$, 0 °C, 68 %. (b) $POCl_3$, pyridine, 0–25 °C, 71 %. (c) $NaCNBH_3$, $BF_3{\cdot}Et_2O$, THF, 60 °C, 87 %. (d) NaOH, MeOH, 25 °C, 85 %. (e) BnI, NaH, THF/DMF, 25 °C, 2 days, 72 %. (f) SeO_2, *t*BuOOH, CH_2Cl_2, 0 °C, 66 %. (g) Dess–Martin periodinane, CH_2Cl_2, 25 °C, 78 %. (h) Allyl MgBr (3 equiv), $AlCl_3$ (3 equiv), THF, –78 °C, 89 %.



Scheme 2 *Reagents and conditions*: (a) KH, $(MeOCH_2)_2$, reflux, 1 h. (b) H_2O, –10 °C, 91 %. (c) O_3, CH_2Cl_2, –78 °C, $P(Ph)_3$, 55 %. (d) BuLi, Et_2O, –78 °C, 42 %. (e) DIBAL-H, Et_2O, –78 °C, 72 %. (f) OsO_4, *N*-methylmorpholine-*N*-oxide (NMO), THF/*t*BuOH/H_2O, 25 °C. (g) Acetone, TsOH, 25 °C, (23*R*,24*S*): 53 %, (23*S*,24*R*): 25 %. (h) H_2, Pd/C, THF, 25 °C, **7a**: 78 %, **7b**: 72 %. (i) Dowex 50-X8, MeOH, 25 °C, 2 days, (23*R*,24*S*)-**3**: 75 %, (23*S*,24*R*)-**3**: 85 %.

Revision of the absolute configurations of bethosides B and C

The steroidal saponins helosides A and B (**4** and **5**) isolated from *C. luteum* also contain an open-chain steroidal aglycone [12]. To determine the absolute configurations of C-22 and C-25 in the side chain of these saponins, their corresponding aglycone helogenin (**6**) was isolated via semipreparative reverse-phase high-performance liquid chromatography (RPHPLC) performed on the acid hydrolysate of a crude *C. luteum* extract. Crystallization of **6** from a methanol/water solution pleasingly yielded white needles suitable for X-ray crystallography, and the crystal structure obtained both confirmed the constitution of helogenin, and revealed the absolute configuration of the side chain as (22*S*,25*R*) [12]. While helogenin had not been isolated previously, the closely related 3β,16β,22,26-tetrahydroxy-cholest-5-ene skeleton (lacking C-11 hydroxylation) had been reported as the aglycone in saponins isolated from *Allium tuberosum* [26], *Solanum lyratum* [27] and *S. anguivi* [28], and *Trillium erectum* [20]. The ^{1}H and ^{13}C NMR data for the side chain in the helosides **4** and **5** [12] were in surprisingly good agreement ($\Delta\delta_H \leq 0.04$ ppm, $\Delta\delta_C \leq 0.3$ ppm in C_5D_5N for positions 20 to 27) with those of the structurally analogous saponins bethosides B and C isolated from *T. erectum* (**12** and **13**, Fig. 4), which we had reported to possess the epimeric 22*R* configuration [20].

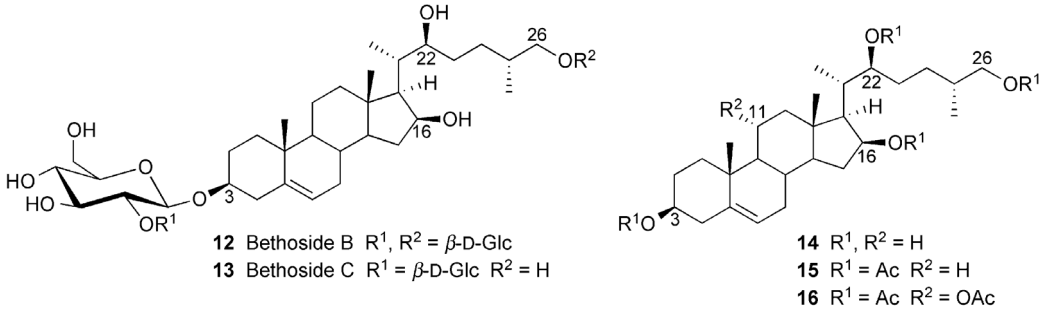


Fig. 4 Saponins **12** and **13** from *T. erectum* and their derivatives **14** and **15**, and the helogenin penta-acetate **16**.

Our assignment of the (22*R*,25*R*) stereochemistry in **12** and **13** was based upon earlier work, in which the absolute configuration of C-22 in the 3β,16β,22,26-tetrahydroxy-cholest-5-ene aglycone (**14**) was determined via Horeau's method [27]. A comparison of the ^{1}H NMR data for the aglycone tetra-acetate **15** derived from bethosides B and C, with the limited chemical shifts reported for synthetic tetra-acetates (22*S*)-**15** and (22*R*)-**15** [27], led to our assignment of (22*R*,25*R*) stereochemistry in **12** and **13** [20]. To investigate further the apparently conflicting stereochemical assignment of bethosides **12** and **13**, we chose to isolate their aglycone (**14**) directly from the acid hydrolysate of a crude *T. erectum* extract. Unexpectedly, it proved difficult to identify bethosides B and C in different *T. erectum* extracts; it was clear that these saponins were absent, or present as only minor constituents in three of the four samples of this herb obtained commercially (Fig. 5). The difficulty encountered in locating the bethoside saponins underlines the problem of variability in the composition of different herbal extracts and the importance of phytochemically profiling medicinal plants.

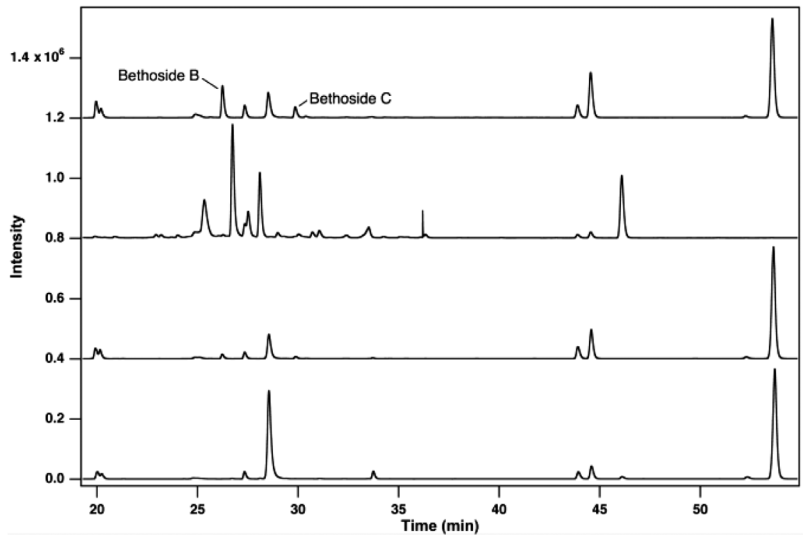
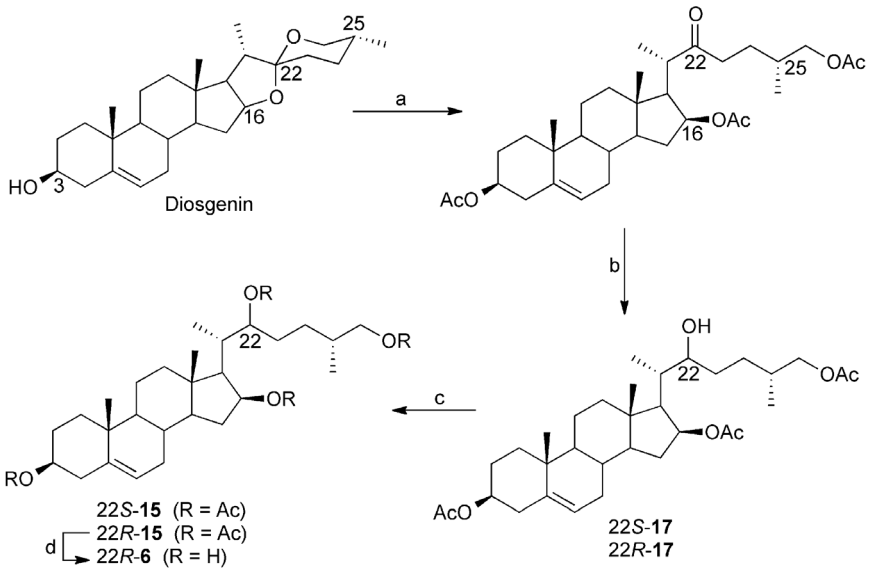


Fig. 5 RPHPLC-evaporative light-scattering detection (ELSD) traces of crude *T. erectum* extracts obtained from four different commercial samples (8–100 % aqueous CH_3CN over 90 min). Bethosides B (**12**) and C (**13**) appear to be present in only one of the four samples.

The close agreement of the ^1H and ^{13}C NMR data for the side chain in **14** and the helogenin aglycone **6** ($\Delta\delta_H \leq 0.04$ ppm, $\Delta\delta_C \leq 0.2$ ppm in C_5D_5N for positions 20 to 27), as well as for their per-acetylated derivatives **15** and **16** ($\Delta\delta_H \leq 0.02$ ppm, $\Delta\delta_C \leq 0.2$ ppm in $CDCl_3$), strongly suggested that 22*S* is the naturally occurring stereochemistry in both the heloside and bethoside saponins [29]. Crystallization of the bethoside aglycone **14** from a methanol/water solution afforded single crystals suitable for X-ray crystallography, and the structure obtained revealed that it indeed possessed the (22*S*,25*R*) absolute configuration of the side chain as seen in the helosides. This work has led to the revision of **12** and **13** to (22*S*)-bethoside B and (22*S*)-bethoside C, respectively [29].

Though it was clear that the 22*S* absolute configuration is present in bethosides B and C (**12** and **13**) as well as helosides A and B (**4** and **5**), it was unclear whether the C-22 epimers of their aglycone could be distinguished by NMR spectroscopy. We therefore synthesized (22*S*)-**15** and (22*R*)-**15**, starting from the readily available steroidal aglycone diosgenin, which already contained the required 25*R* configuration. Spiroketal ring opening of the steroid E and F rings in diosgenin, as reported by Fernandez-Herrera et al. [30,31], was followed by reduction of the C-22 ketone to generate epimeric alcohols (22*S*)-**17** and (22*R*)-**17**, which were in turn acetylated to yield the desired (22*S*)-**15** and (22*R*)-**15** (Scheme 3).



Scheme 3 *Reagents and conditions*: (a) Fernandez-Herrera et al. [30,31] (b) $NaBH_4$, $CeCl_3 \cdot 7H_2O$, MeOH, CH_2Cl_2, 25 °C, starting material: 55 %; semi-preparative NP-HPLC 22*S*-**17**: 3 %, 22*R*-**17**: 10 %. (c) Ac_2O, pyridine, 25 °C, 22*S*-**15**: 53 %, 22*R*-**15**: 66 %. (d) LiOH, MeOH, THF, H_2O, 25 °C, 39 %.

The ^1H and ^{13}C NMR data for (22*S*)-**15** were identical to those of **15** derived from the naturally occurring aglycone of bethosides B and C, and distinct from those of (22*R*)-**15** (Fig. 6). Due to discrepancies between the physical properties of our (22*S*)-**15**, (22*R*)-**15** and the compounds reported in the earlier literature [27], it is unclear where the misassignment first arose. As a final step, (22*R*)-**15** was

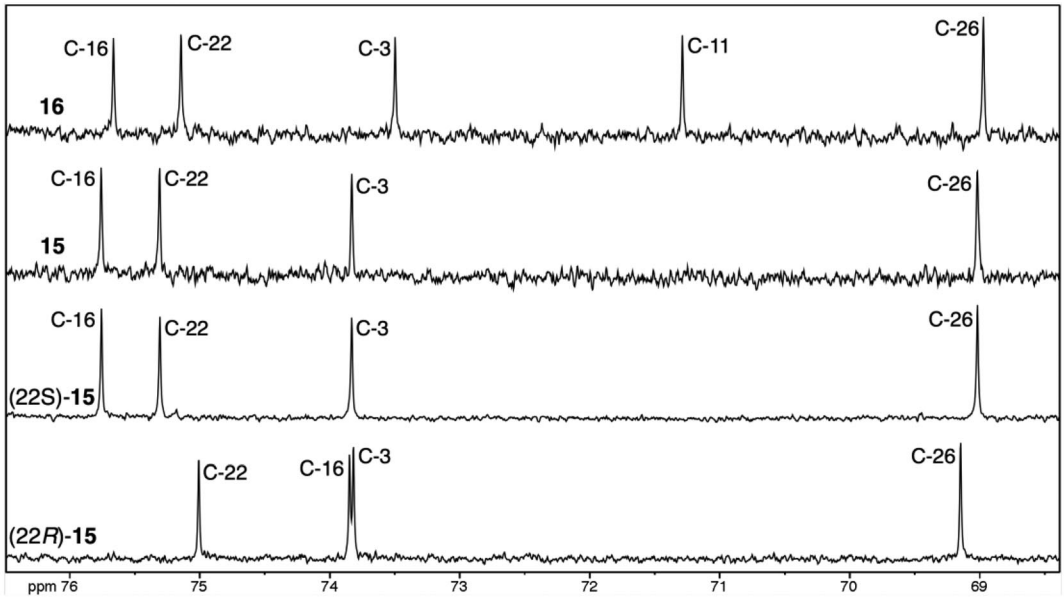


Fig. 6 Segment of the ^{13}C NMR spectra of naturally derived **15** and **16**, and synthetic (22*S*)-**15** and (22*R*)-**15** ($CDCl_3$, 188 MHz).

converted into the unnatural tetra-ol (22*R*)-**6**, whose ^{1}H and ^{13}C NMR chemical shifts differed from those of the natural aglycone, (22*S*)-**6**, at positions 16, 21, and 22 [29]. It was clear from comparison of the ^{1}H and ^{13}C NMR spectra of (22*S*)-**15** and (22*R*)-**15**, and (22*S*)-**6** and (22*R*)-**6**, that C-22 epimers of the bethoside aglycone can be distinguished by NMR spectroscopy.

CONCLUSIONS

The major phytochemical constituents of the widely used North American herb *C. luteum* are open-chain cholesterol-derived steroidal glycosides. The determination of absolute stereochemistry within the flexible steroidal side chain of these saponins proved challenging and required the use of X-ray crystallography and total synthesis. As yet, no obvious trends exist that would allow direct assignment of stereochemistry by NMR spectroscopy.

ACKNOWLEDGMENTS

The authors thank L. Lambert for assistance with the acquisition of NMR spectra (Centre for Advanced Imaging, The University of Queensland) and Dr R. Lehmann (Integria Healthcare) for assistance and support. The authors are grateful for funding from the United States Pharmacopeia (USP Fellowship 2008-2009 to V.L.C.) and an Australian Postgraduate Award to V.L.C.

REFERENCES

1. United States Department of Agriculture Natural Resources Conservation Service. *The PLANTS Database*, National Plant Data Center, <http://plants.usda.gov>, accessed 26 July 2010.
2. A. Chevallier. *The Encyclopedia of Medincal Plants*, Dorling Kindersley, London (1996).
3. M. McKenzie, A. Kirakosyan, P. B. Kaufman. In *Recent Advances in Plant Biotechnology*, A. Kirakosyan, P. B. Kaufman (Eds.), pp. 363, Springer (2009).
4. E. P. Burkhart, M. G. Jacobson. *Agrofor. Syst.* **76**, 437 (2009).
5. F. V. Greene. *Am. J. Pharm.* **50**, 250 (1878).
6. F. V. Greene. *Am. J. Pharm.* **50**, 465 (1878).
7. R. E. Marker, R. B. Wagner, P. R. Ulshafer. *J. Am. Chem. Soc.* **64**, 1283 (1942).
8. K. M. Newton, S. D. Reed, L. Grothaus, K. Ehrlich, J. Guiltinan, E. Ludman, A. Z. LaCroix. *Maturitas* **52**, 134 (2005).
9. S. S. W. Ng, W. D. Figg. *Anticancer Res.* **23**, 3585 (2003).
10. S. G. Sparg, M. E. Light, J. van Staden. *J. Ethnopharmacol.* **94**, 219 (2004).
11. N. J. Matovic, J. M. U. Stuthe, V. L. Challinor, P. V. Bernhardt, R. P. Lehmann, W. Kitching, J. J. De Voss. *Chem.—Eur. J.* **17**, 7578 (2011).
12. V. L. Challinor, J. M. U. Stuthe, P. V. Bernhardt, R. P. Lehmann, W. Kitching, J. J. De Voss. *J. Nat. Prod.* **74**, 1557 (2011).
13. L. Minale, R. Riccio, F. Zollo. *Prog. Chem. Org. Nat. Prod.* **62**, 75 (1993).
14. V. A. Stonik. *Russ. Chem. Rev.* **70**, 673 (2001).
15. V. A. Stonik, N. V. Ivanchina, A. A. Kicha. *Nat. Prod. Commun.* **3**, 1587 (2008).
16. N. V. Ivanchina, A. A. Kicha, V. A. Stonik. *Steroids* **76**, 425 (2011).
17. M. Kuroda, Y. Mimaki, Y. Sashida. *Phytochemistry* **52**, 445 (1999).
18. H. Yoshimitsu, M. Nishida, T. Nohara. *Chem. Pharm. Bull.* **48**, 556 (2000).
19. A. I. Hamed, W. Oleszek, A. Stochmal, C. Pizza, S. Piacente. *Phytochemistry* **65**, 2935 (2004).
20. P. Y. Hayes, R. Lehmann, K. Penman, W. Kitching, J. J. De Voss. *Phytochemistry* **70**, 105 (2009).
21. W. A. Konig, I. Benecke, H. Bretting. *Angew. Chem., Int. Ed. Engl.* **20**, 693 (1981).
22. P. K. Agrawal, P. Bunsawansong, G. A. Morris. *Phytochemistry* **47**, 255 (1998).

23. P. K. Agrawal. *Steroids* **70**, 715 (2005).
24. K. Takeda, H. Minato, A. Shimaoka. *Chem. Commun. (London)* 104 (1968).
25. M. Koreeda, Y. Tanaka, A. Schwartz. *J. Org. Chem.* **45**, 1172 (1980).
26. S. M. Sang, S. L. Mao, A. N. Lao, Z. L. Chen, C. T. Ho. *Food Chem.* **83**, 499 (2003).
27. S. Yahara, M. Ohtsuka, K. Nakano, T. Nohara. *Chem. Pharm. Bull.* **37**, 1802 (1989).
28. X. H. Zhu, H. Tsumagari, T. Honbu, T. Ikeda, M. Ono, T. Nohara. *Tetrahedron Lett.* **42**, 8043 (2001).
29. V. L. Challinor, P. Y. Hayes, P. V. Bernhardt, W. Kitching, R. P. Lehmann, J. J. De Voss. *J. Org. Chem.* **76**, 7275 (2011).
30. M. A. Fernandez-Herrera, J. Sandoval-Ramirez, S. Meza-Reyes, S. Montiel-Smith. *J. Mex. Chem. Soc.* **53**, 126 (2009).
31. M. A. Fernandez-Herrera, H. Lopez-Munoz, J. M. V. Hernandez-Vazquez, M. Lopez-Davila, M. L. Escobar-Sanchez, L. Sanchez-Sanchez, B. M. Pinto, J. Sandoval-Ramirez. *Bioorg. Med. Chem.* **18**, 2474 (2010).

