Available online at www.sciencedirect.com
ScienceDirect
Tetrahedron Letters
Tetrahedron Letters 48 (2007) 7323–7326

Synthesis of steroidal saponins bearing an aromatic E ring

Ziyu Wang, Ming Li, Xueting Liu and Biao Yu*

State Key Laboratory of Bio-organic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China

Received 18 July 2007; revised 7 August 2007; accepted 8 August 2007
Available online 10 August 2007

Abstract—A facile synthetic approach toward the steroidal saponins bearing an aromatic E ring was developed starting from the readily available spirostan saponin.

Steroidal saponins constitute an extremely diverse and abundant family of plant metabolites, with a broad range of biological activities.[1] The steroidal aglycones are biosynthetically derived from the 30-carbon oxidosqualene; removal of the three methyl groups of a lanostane precursor produces the 27-carbon steroidal skeleton.[2] Subsequent degradation may take place to form steroids with less than 27-carbons. In 2003, Nohara et al. disclosed from *Solanum aethiopicum* three steroidal saponins having novel 29/30-carbon skeletons.[3] These saponin compounds, named aethioside A–C, bearing an additional aromatized E ring are unprecedented. In 2005, an additional congener was identified from *Paris polyphylla*.[4] The origin of the additional carbons and the bio-formation of the aromatized E ring are yet to be known. Herein we report a facile approach toward the synthesis of the steroidal saponins bearing an aromatic E ring.

Recently, we have developed an effective approach for the conversion of the ready available spirostan saponins (e.g., **4**; Fig. 1) into their furostan counterparts,[5] where an oxidative ring open of the spirostan E and F rings was employed to provide a 16,22-di-one (e.g., **3**) as a versatile intermediate for further elaboration.[5,6] We envisioned the ready formation of the furan or

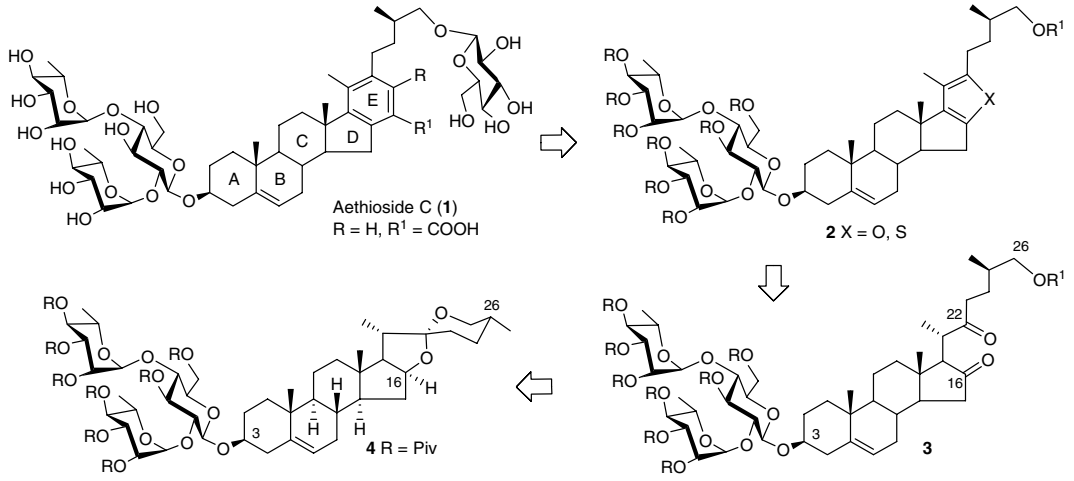
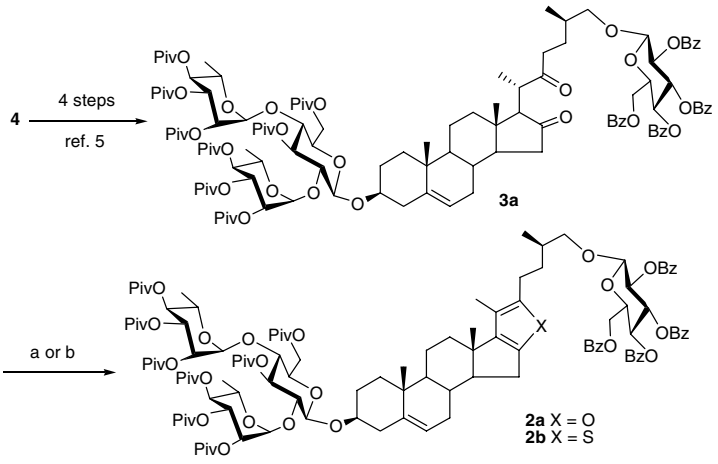


Figure 1. The natural steroidal saponin aethioside C (**1**) and its retrosynthetic analysis.

Keywords: Steroidal saponin; Diels–Alder reaction; Thiophene; Furan; Aethioside.
* Corresponding author. Tel.: +21 64163300x1407; fax: +21 64166128; e-mail: byu@mail.sioc.ac.cn

0040-4039/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2007.08.028

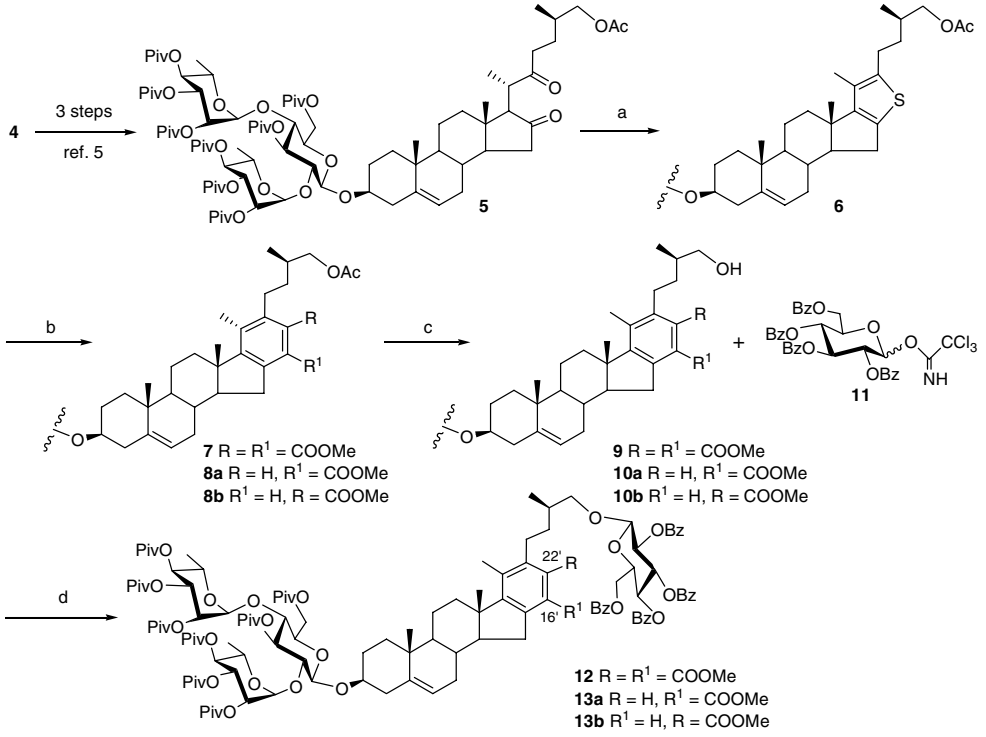


Scheme 1. Reagents and conditions: (a) *p*-TsOH (cat.), Ac_2O, reflux, 2 h, 65%; (b) Lawesson's reagent, 4 Å MS, toluene, reflux, overnight, 78%.

thiophene derivatives (e.g., **2**) from the 16,22-di-one to obtain the novel steroidal saponin derivatives with a heterocyclic E ring. Moreover, an oxidative Diels–Alder cycloaddition of the furan/thiophene E ring with an acetylene dienophile would generate a benzene E ring, thus providing an access to the unusual aethioside-type saponins.

We first tried the furan/thiophene ring formation on the 16,22-di-one-3,26-bisglycoside **3a**, a precursor to the furostan saponins (Scheme 1).[5] Thus, reflux of 16,22-di-one **3a** in Ac_2O in the presence of a catalytic amount of *p*-toluenesulfonic acid for two hours afforded the desired furan derivative **2a** as a major product in 65% isolated yield.[7] However, this furan derivative decomposed quickly at storage or under conditions for the subsequent Diels–Alder reaction. Alternatively, dione **3a** was treated with Lawesson's reagent in the presence of 4 Å MS in toluene;[8] the desired thiophene derivative **2b** was obtained in a good 78% yield. Nevertheless, under the oxidative Diels–Alder cycloaddition conditions (*m*-CPBA, $BF_3{\cdot}Et_2O$, dimethyl acetylenedicarboxylate, or methyl propiolate), which have been well documented by Mataka et al.,[9] the thiophene derivative **2b** proceeded into a complex mixture.

The bulky 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl residue at the 26-OH that is proximal to the thiophene ring in **2b** might play a detrimental role in the above reaction. Thus, we attempted the transformation from



Scheme 2. Reagents and conditions: (a) Lawesson's reagent, 4 Å MS, toluene, reflux, overnight, 90%; (b) dimethyl acetylenedicarboxylate or methyl propiolate, $BF_3{\cdot}Et_2O$, *m*-CPBA, 4 Å MS, toluene, −70 to 0 °C, 72% (for **7**, overnight); 56% (for **8a/b**, a week); (c) NaOMe, $MeOH/CH_2Cl_2$ (v:v 4:1), rt, overnight, 72% (for **9**); 92% (for **10a/b**); (d) TMSOTf (0.3 equiv), 4 Å MS, CH_2Cl_2, −40 °C to rt, overnight, 82% (for **12**); 79% (for **13a/b**).

16,22-di-one-26-*O*-acetate **5**, which was readily prepared from the spirostan dioscin pivalate **4** in three steps and 68% yield (Scheme 2).[5] Treatment of 16,22-di-one **5** with Lawesson's reagent provided the thiophene derivative **6** in an excellent 90% yield. Thiophene **6** was then applied to the Mataka's cycloaddition conditions, where the thiophene ring was expected to be selectively oxidized at low temperature with *m*-CPBA in the presence of $BF_3{\cdot}Et_2O$ to yield the thiophene *S*-monoxide; subsequent addition to a dienophile and extrusion of the resulting O═S-bridge in the [4+2]-cycloadduct under the oxidative surroundings should provide the desired benzene derivative.[9] Indeed, when dimethyl acetylenedicarboxylate was used as an dienophile and toluene as the solvent (overnight), the corresponding benzene derivative **7** was isolated in a good 72% yield. However, under similar conditions, the reaction with methyl propiolate was found much sluggish, providing a pair of the inseparable regio-isomers (**8a/b**, 3:7) in 56% yield over a week of reaction. Selective removal of the 26-*O*-acetyl group in **7** and **8a/b** was achieved with NaOMe in MeOH/CH_2Cl_2, providing **9** and **10a/b**. Glycosylation of the 26-ols (**9** and **10a/b**) with 2,3,4,6-tetra-*O*-benzoyl-D-glucopyranosyl trichloroacetimidate **11** under the promotion of TMSOTf afforded the desired bisglycosides **12** and **13a/b** in about 80% yield.[10] Final removal of the ester groups on **12** and **13a/b**, and the previous compounds **2b**, **6**, **7**, and **8a/b** as well, with LiOH in a mixture solvent of MeOH/THF/H_2O (to compromise the solubility), afforded novel steroidal saponins with a benzene or thiophene E ring.

In summary, we have shown the ready conversion of a spirostan saponin into the novel saponin derivatives with an aromatized E ring, including the furan, thiophene, and benzene ring. To achieve the effective synthesis of the unusual aethioside-type saponins, the regioselectivity in the Diels–Alder cycloaddition with the E-thiophene-ring requires improvement. In addition, diversity-oriented synthesis based on the present chemistry and the screening of the bioactivities of these compounds are our undergoing projects.

Acknowledgments

Financial support from the Chinese Academy of Sciences (KGCX2-SW-213 and KSCX2-YW-R-23) and the National Natural Science Foundation of China (20321202) are gratefully acknowledged. X.L. also thanks the China Postdoctoral Science Foundation for a grant.

References and notes

1. Hostettmann, K.; Marston, A. *Saponins*; Cambridge University Press: Cambridge, UK, 1995.
2. Vincken, J.-P.; Heng, L.; Groot, A. D.; Gruppen, H. *Phytochemistry* **2007**, *68*, 275–297.
3. Tagawa, C.; Okawa, M.; Ikeda, T.; Yoshida, T.; Nohara, T. *Tetrahedron Lett.* **2003**, *44*, 4839–4841.
4. Huang, Y.; Wang, Q.; Ye, W.-C.; Cui, L. J. *Chin. J. Nat. Med.* **2005**, *3*, 138–140.
5. Li, M.; Yu, B. *Org. Lett.* **2006**, *8*, 2679–2682.
6. (a) Bovicelli, P.; Lupattelli, P.; Fracassi, D.; Mincione, E. *Tetrahedron Lett.* **1994**, *35*, 935–938; (b) Yu, B.; Liao, J.; Zhang, J.; Hui, Y. *Tetrahedron Lett.* **2001**, *42*, 77–79; (c) Cheng, M. S.; Wang, Q. L.; Tian, Q.; Song, H. Y.; Liu, Y. X.; Li, Q.; Xu, X.; Miao, H. D.; Yao, X. S.; Yang, Z. *J. Org. Chem.* **2003**, *68*, 3658–3662.
7. Chaosuancharoen, N.; Kongkathip, N.; Kongkathip, B. *Synth. Commun.* **2004**, *34*, 961–983.
8. Jesberger, M.; Davis, T. P.; Barner, L. *Synthesis* **2003**, 1929–1958.
9. (a) Li, Y.-Q.; Thiemann, T.; Mimura, K.; Sawada, T.; Mataka, S.; Tashiro, M. *Eur. J. Org. Chem.* **1998**, 1841–1850; (b) Thiemann, T.; Li, Y.-Q.; Thiemann, C.; Sawada, T.; Ohira, D.; Tashiro, M.; Mataka, S. *Heterocycles* **2000**, *52*, 1215–1230.
10. Selected data for the new compounds. Compound **2a**: $[\alpha]_D^{25}$ −19 (*c* 0.7, $CHCl_3$); ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.00–7.72 (m, 8H), 7.52–7.17 (m, 12H), 5.83 (t, 1H, $J = 9.6$ Hz), 5.61 (t, 1H, $J = 9.9$ Hz), 5.47 (t, 1H, $J = 8.4$ Hz), 5.37–5.20 (m, 3H), 5.19–5.11 (m, 2H), 5.09–4.97 (m, 2H), 4.93 (s, 1H), 4.81 (d, 1H, $J = 1.5$ Hz), 4.76 (d, 1H, $J = 8.1$ Hz), 4.70 (s, 1H), 4.61–4.50 (m, 2H), 4.49–4.30 (m, 3H), 4.25–4.15 (m, 1H), 4.14–4.03 (m, 1H), 3.92–3.80 (m, 1H), 3.76–3.63 (m, 3H), 3.63–3.46 (m, 2H), 3.40–3.30 (m, 1H), 2.45–2.28 (m, 5H), 2.28–2.10 (m, 2H), 2.00–1.80 (m, 5H), 1.73 (s, 3H), 0.97 (s, 3H), 0.82 (s, 3H), 0.74 (d, 3H, $J = 6.3$ Hz); ^{13}C NMR (75 MHz, $CDCl_3$): δ 177.8, 177.1 (2C), 177.0, 176.8 (2C), 176.3, 176.2, 166.1, 165.8, 165.2, 165.1, 154.8, 153.1, 140.2, 136.6, 133.4, 133.2, 133.0 (2C), 129.8, 129.7, 129.7, 129.7, 129.6, 129.4, 128.9, 128.8, 128.4, 128.3, 128.2, 121.9, 111.4, 101.3, 99.3, 97.6, 97.1, 78.9, 76.1, 75.9, 75.0, 73.0, 72.1, 72.0, 71.9, 70.8, 70.4, 69.9, 69.8, 69.3, 69.1, 68.6, 68.1, 66.8, 63.3, 62.9, 60.5, 50.8, 49.4, 45.8, 41.0, 38.8, 38.8, 38.7, 38.6, 38.5, 37.1, 37.0, 35.4, 32.7, 32.3, 31.7, 30.6, 30.2, 29.8, 29.7, 29.5, 27.2, 27.1 (2C), 27.0 (3C), 26.6, 23.9, 20.4, 19.3, 18.0, 17.6, 17.2, 16.6, 8.7; MALDI-HRMS *m*/*z*: $[M+Na]^+$ calcd for $C_{119}H_{160}O_{33}Na$, 2140.0734; found: 2140.0829. Compound **2b**: $[\alpha]_D^{25}$ −16 (*c* 1.0, $CHCl_3$); ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.98–7.73 (m, 8H), 7.50–7.16 (m, 12H), 5.83 (t, 1H, $J = 9.0$ Hz), 5.61 (t, 1H, $J = 9.6$ Hz), 5.47 (t, 1H, $J = 9.3$ Hz), 5.37–5.20 (m, 3H), 5.18–5.11 (m, 2H), 5.08–4.97 (m, 2H), 4.92 (s, 1H), 4.80 (s, 1H), 4.76 (d, 1H, $J = 7.8$ Hz), 4.70 (s, 1H), 4.61–4.50 (m, 2H), 4.48–4.30 (m, 3H), 4.23–4.14 (m, 1H), 4.12–4.04 (m, 1H), 3.92–3.80 (m, 1H), 3.76–3.62 (m, 3H), 3.62–3.46 (m, 2H), 3.38–3.28 (m, 1H), 2.62–1.94 (m, 9H), 1.90 (s, 3H), 0.96 (s, 3H), 0.84 (s, 3H), 0.75 (d, 3H, $J = 6.6$ Hz); ^{13}C NMR (75 MHz, $CDCl_3$): δ 177.7, 177.1, 176.8, 176.2 (2C), 166.1, 165.8, 165.2, 165.0, 155.7, 140.2, 139.7, 135.3, 133.4, 133.1, 133.0, 129.8, 129.7, 129.4, 128.8, 128.3, 128.2, 126.8, 122.0, 101.3, 99.2, 97.6, 97.1, 78.9, 77.3, 76.1, 75.9, 74.9, 73.0, 72.1, 72.0, 71.9, 70.8, 70.4, 69.9, 69.8, 69.3, 69.1, 68.6, 68.1, 66.8, 63.2, 62.9, 61.5, 50.7, 44.1, 38.8, 38.8, 38.7, 38.5, 37.1, 36.9, 35.1, 32.8, 31.5, 30.5, 29.8, 29.6, 29.5, 27.1, 27.1, 27.0, 26.9, 25.7, 23.7, 20.6, 19.3, 17.2, 17.1, 16.6, 11.7; MALDI-HRMS *m*/*z*: $[M+Na]^+$ calcd for $C_{119}H_{160}O_{32}SNa$, 2156.0506; found: 2156.0489. Compound **6**: $[\alpha]_D^{25}$ −34 (*c* 0.8, $CHCl_3$); ^{1}H NMR (300 MHz, $CDCl_3$): δ 5.41–5.25 (m, 3H), 5.22–5.16 (m, 2H), 5.11–5.01 (m, 2H), 4.97 (s, 1H), 4.84 (s, 1H), 4.74 (s, 1H), 4.57 (d, 1H, $J = 7.5$ Hz), 4.50–4.34 (m, 2H), 4.23 (dd, 1H, $J = 6.0$, 12.0 Hz), 3.98–3.84 (m, 3H), 3.78–3.50 (m, 4H), 2.76–2.56 (m, 3H), 2.48–2.16 (m, 4H), 2.05 (s, 3H), 2.03 (s, 3H), 1.00 (s, 3H), 0.96 (d, 3H, $J = 6.3$ Hz), 0.92 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 177.7, 177.1, 177.0, 177.0, 176.8, 176.7, 176.2, 176.2, 171.1, 155.7, 140.1, 139.4, 135.6, 126.9, 121.9, 99.1, 97.6, 97.0, 78.8, 77.2, 76.0, 75.8, 71.9, 70.7, 70.3, 69.7, 69.2, 69.0, 69.0, 68.6, 68.0, 67.0, 63.1, 61.4, 50.6,

44.2, 38.8, 38.7, 38.6, 38.4, 37.1, 36.8, 35.2, 34.9, 32.1, 31.5, 30.4, 29.8, 29.5, 27.1, 27.0 (2C), 26.9, 25.7, 20.9, 20.5, 19.2, 17.2, 17.1, 16.7, 11.8; MALDI-HRMS *m/z*: $[M+Na]^+$ calcd for $C_{87}H_{136}O_{24}SNa$, 1619.9035; found: 1619.9040. Compound **7**: $[\alpha]_D^{25}$ −54 (*c* 0.8, $CHCl_3$); ^1H NMR (300 MHz, $CDCl_3$): δ 5.40 (d, 1H, $J = 3.9$ Hz), 5.36–5.26 (m, 2H), 5.23–5.16 (m, 2H), 5.12–5.02 (m, 2H), 4.97 (s, 1H), 4.85 (s, 1H), 4.74 (s, 1H), 4.57 (d, 1H, $J = 7.5$ Hz), 4.50–4.36 (m, 2H), 4.24 (dd, 1H, $J = 5.4$, 11.7 Hz), 4.00–3.88 (m, 3H), 3.86 (s, 3H), 3.83 (s, 3H), 3.78–3.51 (m, 4H), 3.01 (dd, 1H, $J = 6.1$, 5.9 Hz), 2.33 (s, 3H), 2.06 (s, 3H), 1.02 (s, 3H), 1.00 (s, 3H), 0.99 (d, 3H, $J = 7.2$ Hz); ^{13}C NMR (75 MHz, $CDCl_3$): δ 177.8, 177.1, 177.1, 177.0, 176.8, 176.8, 176.3, 176.2, 171.2, 170.4, 167.6, 154.1, 142.4, 140.1, 137.1, 136.1, 132.4, 123.8, 122.0, 99.3, 97.6, 97.1, 79.0, 76.1, 75.9, 72.0, 70.8, 70.4, 69.8, 69.3, 69.1, 68.9, 68.7, 68.1, 66.8, 62.9, 56.5, 52.2, 52.0, 50.0, 47.2, 38.8, 38.8, 38.7, 38.5, 37.1, 36.7, 36.5, 34.1, 33.1, 32.7, 31.5, 30.7, 29.9, 29.7, 28.2, 27.2, 27.1, 27.1, 27.1, 27.0, 21.1, 20.9, 19.3, 17.3, 16.7, 16.2, 14.9; MALDI-HRMS *m/z*: calcd for $[M+Na]^+$ $C_{93}H_{142}O_{28}Na$, 1729.9580; found: 1729.9602. Compound **12**: $[\alpha]_D^{25}$ −33 (*c* 0.3, $CHCl_3$); ^1H NMR (300 MHz, $CDCl_3$): δ 8.05–7.80 (m, 8H), 7.58–7.24 (m, 12H), 5.90 (t, 1H, $J = 9.6$ Hz), 5.69 (t, 1H, $J = 9.6$ Hz), 5.53 (t, 1H, $J = 9.0$ Hz), 5.42 (d, 1H, $J = 3.6$ Hz), 5.39–5.28 (m, 2H), 5.25–5.19 (m, 2H), 5.15–5.05 (m, 2H), 5.00 (s, 1H), 4.88 (s, 1H), 4.84 (d, 1H, $J = 7.8$ Hz), 4.77 (s, 1H), 4.65 (dd, 1H, $J = 3.0$, 12.3 Hz), 4.60 (d, 1H, $J = 7.2$ Hz), 4.56–4.38 (m, 3H), 4.26 (dd, 1H, $J = 5.4$, 12.0 Hz), 4.21–4.12 (m, 1H), 3.99–3.90 (m, 1H), 3.84 (s, 3H), 3.82 (s, 3H), 3.79–3.54 (m, 4H), 3.42–3.34 (m,1H), 3.00 (dd, 1H, $J = 6.6$, 16.2 Hz), 2.21 (s, 3H), 1.04 (s, 3H), 0.96 (s, 3H), 0.84 (d, 3H, $J = 6.6$ Hz); ^{13}C NMR (75 MHz, $CDCl_3$): δ 177.7, 177.1, 177.0, 176.8 (2C), 176.3, 176.2, 170.4, 167.7, 166.1, 165.8, 165.2, 165.1, 153.8, 142.0, 140.1, 137.4, 136.4, 133.4, 133.2, 133.0, 133.0, 132.2, 129.8, 129.7, 129.6, 129.4, 128.9, 128.9, 128.4, 128.3, 128.3, 128.2, 123.6, 122.0, 101.5, 99.4, 97.4, 97.1, 79.0, 76.1, 75.9, 75.0, 73.0, 72.2, 72.1, 72.0, 70.9, 70.4, 69.9, 69.8, 69.4, 69.1, 68.7, 68.1, 66.8, 63.3, 63.0, 56.5, 52.1, 51.9, 50.1, 47.1, 38.8, 38.8, 38.7, 38.5, 37.1, 36.7, 36.4, 34.0, 33.8, 32.7, 31.5, 30.7, 29.9, 29.7, 28.4, 27.6, 27.2, 27.1 (3C), 27.0, 21.0, 19.3, 17.3, 16.6, 16.1, 14.8; MALDI-HRMS *m/z*: $[M+Na]^+$ calcd for $C_{125}H_{166}O_{36}Na$, 2266.1051; found: 2266.1071. Compound **8a/b**: ^1H NMR (300 MHz, $CDCl_3$): δ 7.57 (s, 0.3H), 7.44 (s, 0.7H), 5.45–5.26 (m, 3H), 5.24–5.16 (m, 2H), 5.13–5.03 (m, 2H), 4.98 (s, 1H), 4.86 (s, 1H), 4.75 (s, 1H), 4.58 (d, 1H, $J = 7.8$ Hz), 4.52–4.36 (m, 2H), 4.24 (dd, 1H, $J = 5.7$, 12.0 Hz), 4.05–3.88 (m, 3H), 3.86 (s, 1H), 3.84 (s, 2H), 3.80–3.50 (m, 4H), 3.26 (dd, 0.3H, $J = 6.0$, 16.8 Hz), 2.31 (s, 3H), 2.05 (s, 3H), 1.05–0.99 (m, 9H); MALDI-HRMS *m/z*: $[M+Na]^+$ calcd for $C_{91}H_{140}O_{26}Na$, 1671.9525; found: 1671.9501. Compound **13a/b**: ^1H NMR (300 MHz, $CDCl_3$): δ 8.05–7.80 (m, 8H), 7.56–7.23 (m, 13H), 5.90 (t, 1H, $J = 9.3$ Hz), 5.69 (t, 1H, $J = 9.9$ Hz), 5.54 (t, 1H, $J = 7.8$ Hz), 5.42 (br s, 1H), 5.38–5.27 (m, 2H), 5.25–5.18 (m, 2H), 5.14–5.04 (m, 2H), 4.99 (s, 1H), 4.87 (s, 1H), 4.83 (d, 1H, $J = 7.8$ Hz), 4.77 (s, 1H), 4.69–4.56 (m, 2H), 4.55–4.38 (m, 3H), 4.25 (dd, 1H, $J = 5.1$, 12.3 Hz), 4.21–4.12 (m, 1H), 3.99–3.89 (m, 1H), 3.86 (s, 1.3H), 3.81 (s, 1.7H), 3.79–3.53 (m, 4H), 3.49–3.36 (m, 1H), 3.27 (dd, 0.3H, $J = 6.9$, 16.8 Hz), 2.22 (s, 1.3H), 2.21 (s, 1.7H), 1.03 (s, 3H), 0.97 (s, 3H), 0.86 (m, 3H); MALDI-HRMS *m/z*: $[M+Na]^+$ calcd for $C_{123}H_{164}O_{34}Na$, 2208.0996; found: 2208.0993.