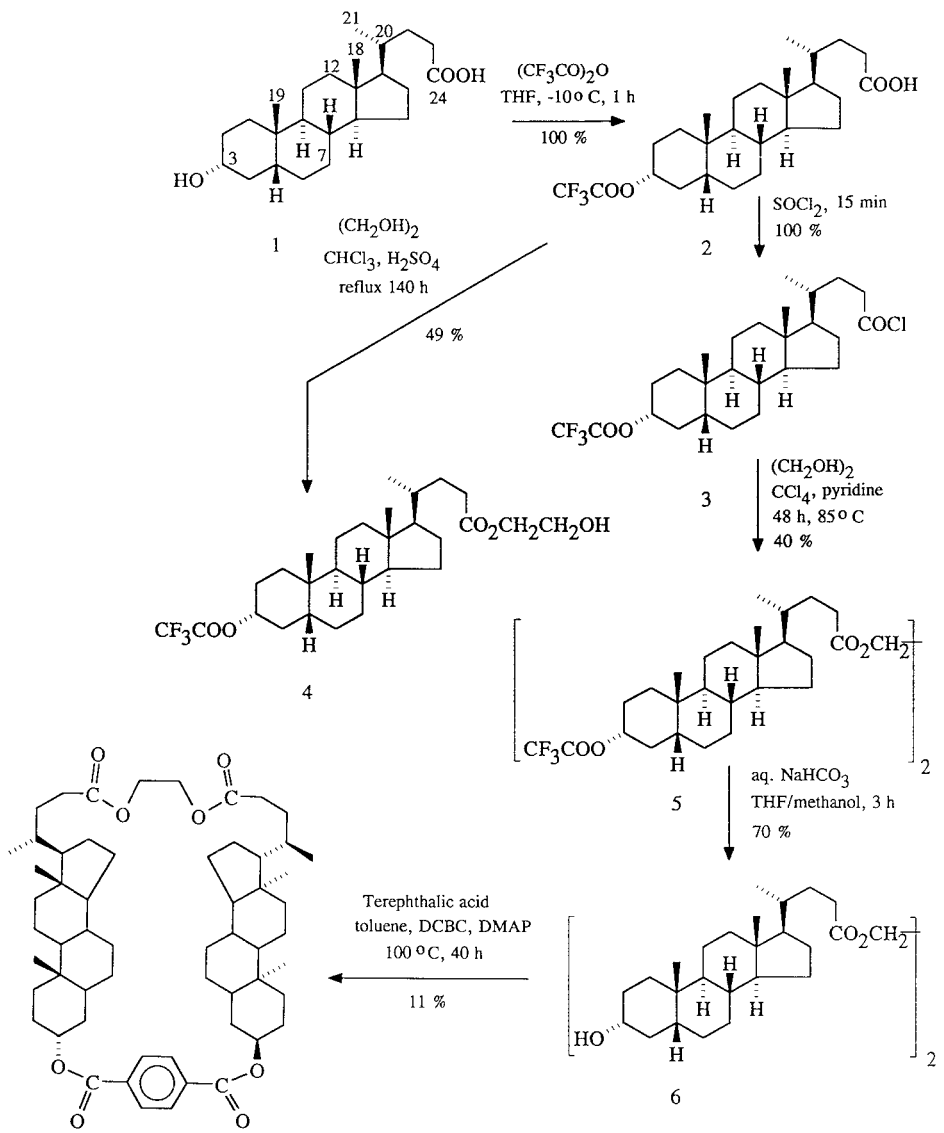
Substituted Methyl 5β-Cholan-24-oates; Part III:† Synthesis of a Novel Cholaphane from Ethylene Glycol Diester of Lithocholic Acid by Cyclization with Terephthalic Acid

Erkki Kolehmainen,* Jari Tamminen, Kari Lappalainen, Tuire Torkkel, Raimo Seppälä
Department of Chemistry, University of Jyväskylä, FIN-40351 Jyväskylä, Finland
Fax +358(41)622501; E-mail ekolehmainen@jylk.jyu.fi
Received 5 February 1996; revised 25 March 1996

A novel cholaphane, containing both an aliphatic and an aromatic spacer, has been prepared from ethylene glycol diester of 3α-hydroxy-5β-cholan-24-oic acid (lithocholic acid) by cyclization (Yamaguchi method) with terephthalic acid.

During recent years common bile acids (3α-hydroxy-5β-cholan-24-oic acids) and their derivatives have been used as versatile structural components in molecular engineering of steroidal macrocycles such as cholaphanes and cyclocholates[1–5] as well as potential drug-shuttles for liver specific targeting.[6,7] Recently, we have reported synthetic procedures and NMR spectroscopic characterization of several monomeric bile acid derivatives[8,9] as well as cyclic bile acid macrolactones (macrolides)[10,11] and their molecular recognition properties towards aromatic guest molecules such as anisole.[11] Regarding the biomimetic/molecular recognition properties of these macrocycles the size, shape and nature (lipophilicity vs. hydrophilicity) of the molecular cavities are of predominant importance.[1–5,12] One way to influence these critical properties is to vary the spacers between the steroidal building blocks. In this work we wish to report synthetic and analytic (^{13}C NMR and mass spectroscopic) details for



Scheme

a novel steroidal macrocycle containing both a flexible aliphatic spacer derived from ethylene glycol and an aromatic spacer, terephthalic acid, in its structure. The whole synthetic route leading to the desired macrocycle **7** is described in the Scheme.

Our first attempts to use unprotected lithocholic acid **(1)** by sulfuric acid-catalyzed esterification suffered partly from a head-to-tail dimerization and dehydration of bile acid itself. Therefore, in an improved procedure lithocholic acid **(1)** was treated first with trifluoroacetic acid anhydride leading quantitatively to 3α-trifluoroacetoxy-5β-cholan-24-oic acid **(2)**.[1] Further, refluxing **2** with $SOCl_2$ gave 3α-trifluoroacetoxy-5β-cholan-24-oyl chloride **(3)** in ~ 100% yield. The reaction of **3** with ethylene glycol yielded a mixture of 3α-trifluoroacetoxy-5β-cholan-24-oic acid mono- **(4)** and diester **(5)** of ethylene glycol. After optimization of the reaction conditions the amount of diester **5** can be increased up to 40%. The diester **5** was purified by column chromatography (silica gel, acetone/$CHCl_3$, 10:90 as eluent). Hydrolysis of **5** with aqueous $NaHCO_3$ in THF/methanol mixture yielded the lithocholic acid diester **6** of ethylene glycol in 70% yield. A Yamaguchi reaction[14] of **6** with terephthalic acid using 2,6-dichlorobenzoyl chloride (DCBC) and 4-(*N,N*-dimethyl)aminopyridine (DMAP) as cyclization reagents afforded the desired macrocycle **7** in 11% yield.

The formed steroidal macrocycle **7** possessing two different spacers shows: i) three structural units characterized with clearly different conformational properties, ii) several active sites (two sets of carboxylic groups with different spatial distances, an aromatic π-system and two rigid steroid skeletons) for molecular recognition. According to molecular modeling based on Allinger's MM2 force field calculations[15] the cavity dimensions of **7** are 10 × 14 Å.

Varying the steroidal building blocks, i.e. starting the synthesis from other common bile acids (e.g. deoxycholic, chenodeoxycholic and cholic acids) or their derivatives[8–11] and using other related spacers (isomers or homologues) this synthetic route can be modified by structural tailoring for other related host molecules.

In conclusion, we have demonstrated the use of simple building blocks in preparing a novel steroidal macrocycle (cholaphane) characterized by several active sites, structural units with different conformational flexibilities as well as electronic properties.

3α-Hydroxy-5β-cholan-24-oic acid **(1)**, ethylene glycol, and terephthalic acid were purchased from Aldrich Chemical Co. The compounds **2–7** were identified and assigned by their proton BBD (Broad Band Decoupled) ^{13}C NMR and DEPT-135 (Distortionless Enhanced Polarization Transfer) spectra (Table 1) and by EI mass spectra (Table 2) and the compounds **6** and **7** by HRMS as well. The purity and structures of the compounds **2**, **3** and **5** has been checked by elemental analyses and by TLC (silica gel 60 F_{254} Merck) since their melting points often showed a relatively large range owing to the tendency of solvent to be encapsulated in their structures when crystallized.

All ^{13}C NMR spectra were recorded on a Jeol JNM GSX-270 FT NMR spectrometer working at 67.94 MHz. Chemical shifts for ^{13}C NMR are reported in ppm (δ) relative to the central line of $CDCl_3$ (77.07 ppm from internal TMS). EI MS and HRMS were

Table 1. ^{13}C NMR Chemical Shifts of Compounds **1–7**, δ

Assignment	Compound						
	1[a]	**2**	**3**	**4**	**5**	**6**	**7**
C-1	35.6	34.82	34.72	34.75	34.60	35.56	35.13
2	30.4	26.28	26.20	26.23	26.12	30.47	26.90
3	71.8	79.44	79.25	79.40	79.11	71.65	75.68
4	36.5	31.72	31.61	31.65	31.50	36.38	32.42
5	42.5	41.97	41.86	41.90	41.75	42.08	42.01
6	27.5	26.97	26.86	26.90	26.76	27.18	27.14
7	26.7	26.20	26.09	26.13	25.97	26.40	26.37
8	36.1	35.83	35.74	35.76	35.64	35.82	35.87
9	40.6	40.53	40.43	40.44	40.34	40.40	40.55
10	34.8	34.60	34.49	34.53	34.37	34.53	34.75
11	21.1	20.91	20.82	20.86	20.72	20.79	20.97
12	40.5	40.12	40.01	40.05	39.94	40.16	40.17
13	43.0	42.81	42.76	42.74	42.59	42.70	42.86
14	56.7	56.46	56.38	56.39	56.29	56.47	56.51
15	24.4	24.20	24.09	24.15	24.03	24.19	24.48
16	28.4	28.19	28.10	28.17	28.01	28.14	28.26
17	56.5	56.04	55.82	55.99	55.90	55.93	56.11
18	11.8	12.09	11.97	12.02	11.87	12.01	12.15
19	23.2	23.24	23.12	23.18	23.00	23.34	23.44
20	35.6	35.34	34.91	35.33	35.16	35.29	35.44
21	18.1	18.29	18.16	18.27	18.09	18.23	18.36
22	31.2	31.12	31.03	31.14	30.88	31.07	31.23
23	31.1	30.81	44.30	30.93	30.76	30.88	31.00
24	174.5	180.66	173.90	174.65	173.54	173.94	174.04
Gly[b] C-1	–	–	–	61.22	61.85	62.00	62.13
Gly C-2	–	–	–	65.91	–	–	–
TFA CO	–	157.08	156.85	157.01	156.86	–	–
TFA CF_3	–	114.65	114.57	114.58	114.44	–	–
Tere[c] CO	–	–	–	–	–	–	165.46
Tere C-1/4	–	–	–	–	–	–	134.61
Tere C-2/3	–	–	–	–	–	–	129.47

[a] In methanol-d_4.[13]
[b] Glycol.
[c] Terephthalic acid.

Table 2. Main Mass Spectral Ions of Compounds **2–7**

Compound	*m/z* ([fragment], intensity %)
2	472 ($[M]^+$, 22), 358 ($[M\text{-}CF_3CO_2H]^+$, 65), 344 (55), 329 ($[M\text{-}143^a]^+$, 100), 215[b] (65)
3	490 ($[M]^+$, 29) , 386 ($[M\text{-}CF_3Cl]^+$, 52), 344 (62), 329 ($M\text{-}161^a]^+$, 100), 215[b] (60)
4	516 ($[M]^+$, 11), 402 ($[M\text{-}CF_3CO_2H]^+$, 99), 257[c] (47), 215[b] (100)
5	970 ($[M]^+$, < 1), 856 ($[M\text{-}CF_3CO_2H]^+$, 44), 742 ($[856\text{-}CF_3CO_2H]^+$, 7), 499 (100), 344 (49), 257[c] (68), 215[b] (56)
6	779 ($[M]^+$, 1), 761 ($[M\text{-}H_2O]^+$, 22), 743 ($[M\text{-}2H_2O]^+$, 24), 401 (45), 384 (53), 341 (45), 257[c] (100), 215[b] (60)
7	909 ($[M]^+$, < 1), 743 ($[M\text{-}C_6H_4(CO_2H)_2]^+$, 35), 386 (100), 341 (48), 257[c] (67), 215[b] (60)

[a] Fragment resulting from D-ring cleavage via fragmentations at bonds $C_{14}-C_{15}$ and $C_{13}-C_{17}$.[16]
[b] Fragment resulting from D-ring cleavage and charge retention on the ABC ring fragment.[17]
[c] Cleavage of the side chain.[17]

recorded on a VG Autospec mass spectrometer by using direct inlet and 70 eV (35 eV for **7**) electron impact (EI) ionization. Elemental analyses were carried out on a CHN 600 Leco apparatus working at 950 °C. The values for carbon deviated more than the accepted limit, probably due to solvent encapsulation.

3α-Trifluoroacetoxy-5β-cholan-24-oic Acid (2):
To a solution of lithocholic acid (**1**; 3.00 g, 8.00 mmol) in sodium dried THF (75 mL) cooled to −10 °C in an ice-salt bath was added trifluoroacetic acid anhydride (19 mL, 130 mmol) in 15 min.[1] The mixture was kept at −10 °C for 1 h and stirred for 2 h at r.t. The mixture was poured into Et_2O/ice (120 mL : 30 g). The organic layer was washed with water (60 mL), with sat. $NaHCO_3$ solution until the water solution gave an alkaline reaction, with brine (60 mL), dried ($MgSO_4$) and evaporated to dryness; yield: 3.76 g (~100%), mp 163–166 °C ($CHCl_3$) (Tables 1 and 2).

$C_{26}H_{39}F_3O_4$	calc.	C 66.08	H 8.32
(472.6)	found	67.71	8.41

3α-Trifluoroacetoxy-5β-cholan-24-oyl Chloride (3):
A mixture of **2** (2.16 g, 4.60 mmol) and distilled $SOCl_2$ (15 mL) was refluxed for 15 min and the excess of $SOCl_2$ was evaporated under vacuo. The crude product was dissolved in CCl_4 (50 mL) and evaporated to dryness under vacuum; yield: 2.25 g (~100%), mp 98–101 °C ($CHCl_3$) (Tables 1 and 2).

$C_{26}H_{38}ClF_3O_3$	calc.	C 63.60	H 7.80
(491.0)	found	65.21	8.27

3α-Trifluoroacetoxy-5β-cholan-24-oic Acid Ethylene Glycol Monoester (4):
To a suspension of **2** (2.27 g, 4.80 mmol) in $CHCl_3$ (30 mL) were added ethylene glycol (15 mL) and conc. H_2SO_4 (0.5 mL) and the mixture was refluxed for 140 h. To the lower layer of the mixture $CHCl_3$ (10 mL) was added and the organic phase was washed with sat. $NaHCO_3$ (3 × 10 mL) and water (1 × 10 mL), and dried ($MgSO_4$). Monoester **4** was purified by column chromatography (silica gel, acetone/$CHCl_3$ 10:90); yield: 1.21 g (49%). The mp of **4** was not possible to determine accurately because all attempts to recrystallize **4** from $CHCl_3$ (and other solvents) gave viscous mixtures of **4** and the solvent. The purity of **4** was ascertained by TLC (silica gel 60 F_{254} Merck, acetone/$CHCl_3$, 5:95) and ^{13}C NMR spectrum of the recrystallized **4**, which did not show any impurity signals except for $CHCl_3$ (Tables 1 and 2). All attempts to obtain **5** from **4** failed.

3α-Trifluoroacetoxy-5β-cholan-24-oic Acid Ethylene Glycol Diester (5):
To a solution of **3** (5.19 g, 11.0 mmol) in CCl_4 (25 mL) were added ethylene glycol (4.0 mL, 72 mmol) and pyridine (1.0 mL, 12 mmol). The mixture was refluxed at 85 °C for 48 h. To the lower layer of the mixture was added $CHCl_3$ (10 mL) and the $CHCl_3$ layer was washed with sat. aq $NaHCO_3$ (3 × 10 mL), with water (1 × 10 mL) and dried ($MgSO_4$). The crude product **5** was purified by column chromatography (silica gel, acetone/$CHCl_3$, 10:90); yield: 2.22 g (40%), mp 70–80 °C ($CHCl_3$). As in the case of **4**, the solvent very strongly tends to be encapsulated in the structure of **5** explaining the large melting interval. The purity of **5** was ascertained by TLC (silica gel 60 F_{254} Merck, acetone/$CHCl_3$, 5:95) and ^{13}C NMR spectrum of the recrystallized **5**, which did not show any impurity signals except for $CHCl_3$ (Tables 1 and 2).

$C_{54}H_{80}F_6O_8$	calc.	C 66.78	H 8.30
(971.2)	found	69.81	8.37

3α-Hydroxy-5β-cholan-24-oic Acid Ethylene Glycol Diester (6):
To a solution of **5** (2.22 g, 2.29 mmol) in THF/MeOH (1:1, 90 mL) was added sat. aq $NaHCO_3$ (22 mL) under stirring. After 1 h an additional portion of $NaHCO_3$ solution (4 mL) was added and the mixture was stirred for 2 h. Then the mixture was poured into a mixture of Et_2O/H_2SO_4 (160 mL : 80 mL, 0.6 M). The Et_2O layer was washed with water (2 × 40 mL), dried ($MgSO_4$) and evaporated under vacuum. The crude product was purified by column chromatography (silica gel, acetone/$CHCl_3$, 10:90); yield: 1.20 g (70%), mp 90–100 °C ($CHCl_3$). As in the cases of **4** and **5**, the solvent very strongly tends to be encapsulated in the structure of **6** explaining the large melting interval. The purity of **6** was ascertained by TLC (silica gel 60 F_{254} Merck, acetone/$CHCl_3$, 5:95) and ^{13}C NMR spectrum of recrystallized **6**, which did not show any impurity signals except for $CHCl_3$. The structure of **6** was ascertained by HRMS which showed the $[M]^{\cdot+}$-ion (int. %): 778.6080 (100) (calc. 778.6111) and the isotopic ions (int. %): 779.6096 (55), 780.6086 (42), 781.6001 (17) (Tables 1 and 2).

Terephthalate of 3α-Hydroxy-5β-cholan-24-oic Acid Ethylene Glycol Diester, Cholaphane (7):
To a suspension of **6** (1.20 g, 1.52 mmol) and terephthalic acid (0.25 g, 1.52 mmol) in sodium dried toluene (150 mL) was added DMAP (1.50 g, 4.0 equiv) and the mixture was heated to 100 °C in an oil bath. After that DCBC (0.70 g, 1.1 equiv) was added and the mixture was kept at 100 °C for 40 h.[14] After the reaction period the solvent was evaporated under vacuum. The crude product was dissolved in $CHCl_3$ (75 mL) and extracted with aq HCl (2 × 60 mL, 2 M) and sat. aq $NaHCO_3$ solution (2 × 60 mL) to remove DCBC and DMAP. Finally, the mixture was washed with water (60 mL), dried ($MgSO_4$) and evaporated to dryness under vacuum. The cholaphane **7** was separated by sequential column chromatography: i) silica gel, EtOAc/CH_2Cl_2 (12:88), and ii) silica gel, EtOAc/CH_2Cl_2 (4:96); yield: 150 mg (11%), mp 143–145 °C. The purity of **7** was ascertained by TLC (silica gel 60 F_{254} Merck, EtOAc/CH_2Cl_2, 4:96) and ^{13}C NMR spectrum of purified **7**. The structure of **7** was ascertained by HRMS which showed the $[M^{\cdot+}]$-ion (int. %): 908.6143 (100) (calc. 908.6166) and the isotopic ions (int. %): 909.6136 (57), 910.6086 (45), 911.6398 (11) (Tables 1 and 2).

Ms. M. Lahtiperä and Mr. R. Kauppinen are gratefully acknowledged for their kind help in recording the mass and NMR spectra and Mrs. M. Salo for her kind help in running the elemental analyses. Our thanks also are due to the Academy of Finland for the financial support to this project (Grant No. 8581).

† For Parts I and II, see References 8 and 9.

(1) Bonar-Law, R.P.; Davis, A.P.; Sanders, J.K.M. *J. Chem. Soc., Perkin Trans. 1* **1990**, 2245.
(2) Davis, A.P. *Chem. Soc. Rev.* **1993**, *22*, 243 and references cited therein.
(3) Bonar-Law, R.P.; Davis, A.P. *Tetrahedron* **1993**, *49*, 9829.
(4) Bonar-Law, R.P.; Davis, A.P. *Tetrahedron* **1993**, *49*, 9845.
(5) Bonar-Law, R.P.; Davis, A.P.; Dorgan, B.J. *Tetrahedron* **1993**, *49*, 9855.
(6) Wess, G.; Kramer, W.; Enhsen, A.; Glombik, H.; Baringhaus, K.-H.; Bock, K.; Kleine, H.; Schmitt, W. *Tetrahedron Lett.* **1993**, *34*, 819.
(7) Wess, G.; Kramer, W.; Schubert, G.; Enhsen, A.; Baringhaus, K.-H.; Glombik, H.; Müllner, S.; Bock, K.; Kleine, H.; John, M.; Neckermann, G.; Hoffmann, A. *Tetrahedron Lett.* **1993**, *34*, 817.
(8) Kolehmainen, E.; Kaartinen, M.; Kauppinen, R.; Kotoneva, J.; Lappalainen, K.; Lewis, P.T.; Seppälä, R.; Sundelin, J.-P.; Vatanen, V. *Magn. Reson. Chem.* **1994**, *32*, 441.
(9) Lappalainen, K.; Kolehmainen, E.; Kaartinen, M.; Kauppinen, R.; Seppälä, R.; Vatanen, V. *Magn. Reson. Chem.* **1994**, *32*, 786.
(10) Lappalainen, K.; Kolehmainen, E.; Kotoneva, J. *Magn. Reson. Chem.* **1996**, *34*, 316.
(11) Lappalainen, K.; Kolehmainen, E.; Šaman, D. *Spectrochim. Acta. Part A51.* **1995**, 1543.
(12) Albert, D.; Feigel, M. *Tetrahedron Lett.* **1994**, *35*, 565.
(13) Blunt, J.W.; Stothers, J.B. *Org. Magn. Reson.* **1977**, *9*, 439.
(14) Inanaga, J.; Hirata, K.; Saeki, H.; Katsuki, T.; Yamaguchi, Y. *Bull. Chem. Soc. Jpn.* **1979**, *52*, 1989.
(15) Allinger, N.L.; MM2 (85), QCPE (1985) in PCMODEL (Serena Software).
(16) Edenharder, R.; Slemr, J. *J. Chromatogr.* **1981**, *222*, 1 and references cited therein.
(17) Lawson, A.M.; Setchell, K.D.R. In *The Bile Acids. Chemistry, Physiology, and Metabolism*; Setchell, K.D.R.; Kritchevsky, D.; Nair, P.P. Eds.; Plenum: New York, 1988; p 167.