Steroid-based receptors with tunable cavities; a series of polyhydroxylated macrocycles of varying size and flexibility

Anthony P. Davis,*[a] Stephan Menzer,[b] John J. Walsh[a] and David J. Williams[b]

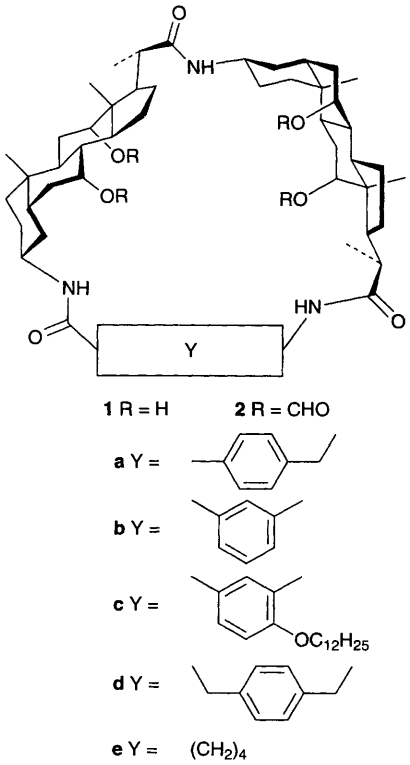
[a] *Department of Chemistry, Trinity College, Dublin 2, Ireland*
[b] *Department of Chemistry, Imperial College of Science and Technology, South Kensington, London, UK SW7 2AY*

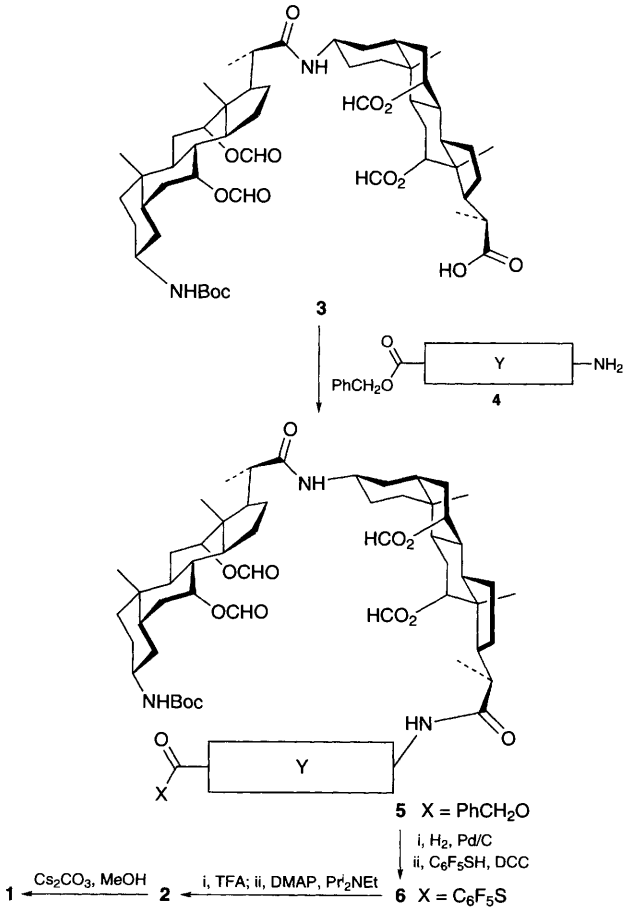
Five 'cyclocholamide' receptors 1 are synthesized and characterised through computer-based molecular modelling and (in one case) X-ray crystallography; a comparison of 1a and 1c as receptors for octyl β-D-glucoside 7 illustrates the potential for tuning recognition properties through variation of group Y.

The preceding communication[1] outlined plans for a new family of receptor molecules **1**, accessible from the inexpensive steroid cholic acid. It also described the synthesis of a C_3-symmetrical prototype in which the macrocycle was completed by a third steroidal unit. We now report (*a*) the extension of this work to the series of non-symmetrical macrocycles **1a–e**, (*b*) the structural characterisation of these molecules by computer-based molecular modelling and (in one case) by X-ray crystallography, and (*c*) preliminary studies indicating that, as expected, the recognition properties of this system may be controlled by adjustment of spacer group Y.

The design of 'cyclocholamides' **1** was influenced by our desire to tune the size of the cavity and, in most cases, to limit conformational freedom as far as possible. Thus rigid spacer units of varying size were employed for **1a–d**, the more flexible analogue **1e** being prepared for comparison purposes. The synthesis of **1c**, with an externally-directed flexible alkyl chain, was prompted by the low solubility of **1b** in $CDCl_3$ and other non-polar organic solvents (*vide infra*).

Spacer precursors **4a**, **b** and **e** were prepared from the corresponding amino acids by *N*-Boc-protection, *O*-benzylation and *N*-deprotection. Compounds **4c** and **4d** were synthesized *via* slightly longer sequences from 3-amino-4-hydroxybenzoic acid and *p*-(bromomethyl)phenylacetic acid respectively. Pentafluorothiophenyl (PFTP) active ester methodology[1,2] was used to couple the alkylamino units **4a**, **d** and **e** to the bis-steroidal acid **3**,[1] giving the corresponding 'trimeric' intermediates **5** in yields of 84–90% (Scheme 1). This method was less successful with **4b** and **4c**, presumably due to the combined influences of steric hindrance in the acyl donor and low nucleophilicity of the aromatic amines. However, acceptable alternatives were found, involving Yamaguchi's mixed anhydride[3] in the case of **4b** and the direct use of DCC in the case of **4c**. Debenzylation of **5a–c** and **5e** was followed by PFTP ester formation to give the corresponding macrocycle precursors **6** in yields of 85–90%. These compounds were treated with TFA to





Scheme 1

remove the *N*-Boc groups, then with DMAP and Pr^i_2NEt in CH_2Cl_2 under high dilution (*ca.* 0.8 mmol dm^{-3}) to give the corresponding macrocycles **2**. Compound **5d** was converted to **2d** similarly, except that pentafluorophenyl ester methodology was used for the macrolactamisation. Cyclisation yields varied between 50 and 61% for **2a–d** (with the rigid aromatic spacers) but fell to 35% for **2e** (flexible alkyl spacer). Deformylation of **2a–e** with Cs_2CO_3 in MeOH proceeded smoothly to give tetraols **1a–e** in yields of *ca.* 95%.†

For host molecules designed for recognition by hydrogen bonding in organic media, the issue of solubility in organic solvents (especially $CDCl_3$) is of obvious importance. Tetraols **1b** and **e** were disappointing in this respect, the former being almost completely insoluble in $CDCl_3$ alone (though readily soluble in $CDCl_3$–CD_3OD mixtures) while the latter was just sufficiently soluble for ^1H NMR spectroscopy at 300 MHz. However, compounds **1a** and **1d** could be dissolved to *ca.* 1 mmol dm^{-3} in $CDCl_3$, and as expected the flexible 'tail' on **1c** resulted in a very useful level of chloroform-solubility (≥10 mmol dm^{-3}). The three distinguishable NH protons were well separated in the ^1H NMR spectra of all the macrocycles, and in the case of **2c** were identified unambiguously through H–H COSY and NOESY 2D spectroscopy.

The structure of **1b** was further characterised by X-ray crystallography. Crystals were grown over two weeks from a methanolic solution. The X-ray analysis‡ revealed two crystallographically independent molecules of **1b** in the asymmetric unit. Both had the desired 'open' macrocyclic geometry with the four hydroxy groups directed inwards and the three amide carbonyl oxygen atoms pointing outwards (Fig. 1). Although the conformations of the two molecules were similar, there were torsional differences of up to 13 ° about equivalent bonds within the flexible linking segments,§ resulting in small differences in the cavity dimensions with the respective transannular O(7)–O(7′) and O(12)–O(12′) distances being 5.17 and 5.61 Å in molecule (1) and 5.39 and 5.16 Å in molecule (2). Although water was not deliberately introduced to the crystallisation medium, the crystals were found to be highly hydrated with ten water molecules distributed over 15 different sites within the asymmetric unit. The water molecules lie both within the macrocycle cavities and in the interstitial spaces between molecules. There is no distinctive pattern of hydrogen bonding. The molecules pack with their macrocyclic mean planes oriented approximately normal to the crystallographic *c* direction. Adjacent molecules in the *c* direction are partially overlapped thus preventing the formation of any channels in the crystals.

The macrocyclic tetraols were also studied using computer-based molecular modelling, employing the Monte Carlo search-minimisation routine in MacroModel V4.0 (MM2 force field, chloroform solvation).[4] The predicted ground state structure for **1b** was closely similar to those observed in the crystals, with transannular O–O distances of 5.04 Å [C(7)-oxygens] and 5.75 Å [C(12)-oxygens]. This framework conformation appeared to be quite rigidly enforced, the first significant alternative being at 14.4 KJ mol^{-1} above baseline. The ground state structure for **1a**, shown in Fig. 2, was similarly open, with transannular O–O distances of 5.73 Å [C(7)] and 6.49 Å [C(12)]. In this case the first closed conformation found was at 7.4 kJ mol^{-1} above baseline. The structure predicted for **1c**¶ was somewhat different to that of **1b**, in that formation of an NH–O hydrogen bond to the alkoxy group promoted a twisted conformation with a smaller cavity. Closed conformations stabilised by intra-annular hydrogen bonds were predicted for **1d** and **1e**, although in the former case an open structure with a substantial cavity lay just 2.45 kJ mol^{-1} above baseline.

The three-dimensional array of polar functionality in **1**, and in other macrocycles derived from cholic acid, is well suited to the recognition of carbohydrate nuclei in non-polar solvents.[5,6] Preliminary studies indicated that variation of spacer group Y in **1** can, as expected, be used to moderate the carbohydrate binding properties of this series of hosts in an understandable fashion. Thus, addition of octyl β-D-glucoside **7** to **1a** in $CDCl_3$ resulted in downfield shifts of all three N*H* NMR resonances. The motion of each signal could be analysed in terms of 1 : 1 complex formation with K_a = 750 dm^3 mol^{-1} (limiting Δδ of 0.76, 1.10, and 0.72 ppm respectively). Molecular modelling of **1a** and methyl β-D-glucoside indicated that the glycoside head-group should indeed be able to enter the cavity to form a complex stabilised by a number of intermolecular hydrogen

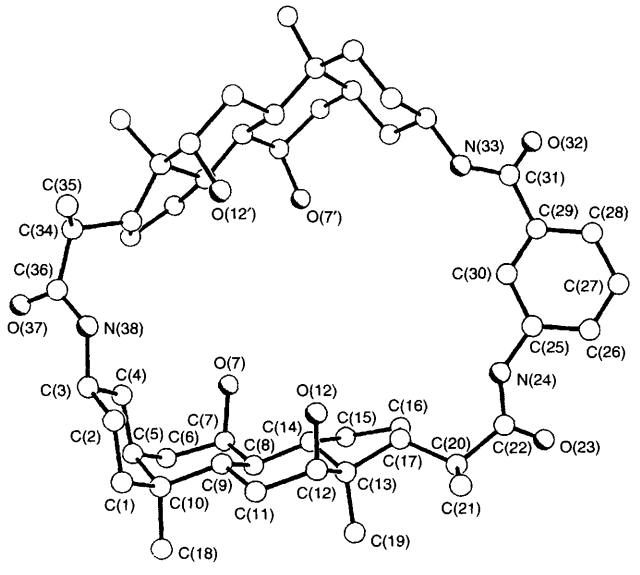


Fig. 1 One of the two independent molecules of **1b** as found in the crystal. Hydrogen atoms are omitted for clarity.

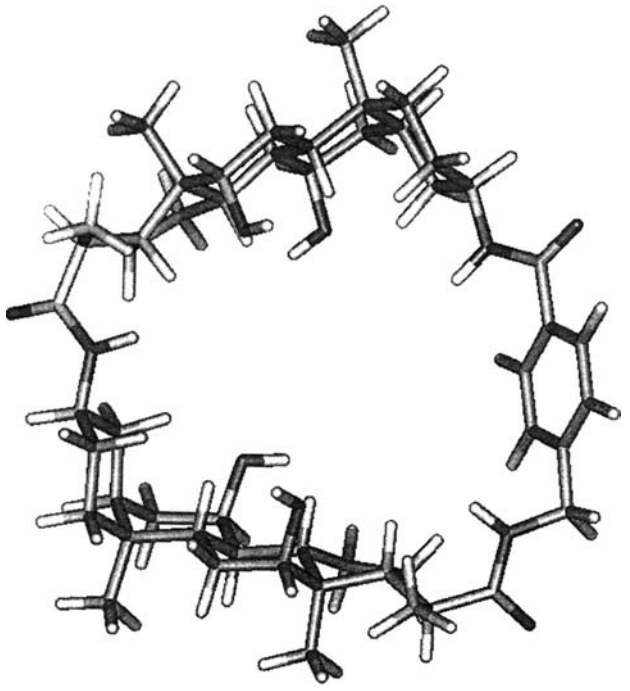
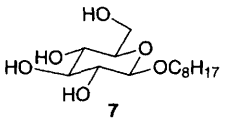


Fig. 2 Structure predicted for **1a** using computer-based molecular modelling (see text)



7

bonds.‖ By contrast, the cavity predicted for **1c** is clearly too small to encapsulate a carbohydrate nucleus. Addition of the octyl glucoside to **1c** in $CDCl_3$ did again cause downfield motions of the NH signals, but the increase of $\Delta\delta$ with concentration were almost linear in this case, indicative of a very low binding constant.

In conclusion, we have demonstrated a versatile method for synthesizing host molecules with well-defined structures and a high level of inward-directed functionality, and have shown that cavity size and other properties can be controlled through choice of the variable module Y. Many further modifications can be envisaged, suggesting that this 'tunable' system could prove a valuable asset in studies of biomimetic molecular recognition and catalysis.

This work was supported by Eolas, the Irish Science and Technology Agency, with additional funding from the EU Human Capital and Mobility Programme and BioResearch Ireland. We thank Mr L. Lawless for molecular modelling, Mr J. Perry for the binding studies, Mr P. Ashton (University of Birmingham) for mass spectra and Diamalt GmbH for generous gifts of cholic acid. J. W. acknowledges financial support from Mayo County Council and D. J. W. thanks the Wolfson Foundation for equipment and computing grants.

Footnotes

† All the macrocyclic systems were characterised by FABMS at either the tetraformate or tetraol stage.

‡ *Crystal data* for 2[**1b**]. $10H_2O$, $C_{102}H_{170}N_6O_{24}$, M = 1864.4, monoclinic, a = 15.283(7), b = 17.243(5), c = 21.056(6) Å, β = 97.28(1)°, U = 5547(3) Å^3, space group $P2_1$, Z = 2, D_c = 1.12 g cm^{-3}, μ(Cu-$K\alpha$) = 6.4 cm^{-1}. F(000) = 2032. Data were measured on a Siemens P4/PC diffractometer (2θ < 120°), with Cu-$K\alpha$ radiation (graphite monochromator) using ω-scans. 8576 independent reflections were measured and of these 6215 had $|F_o| > 4\sigma(|F_o|)$ and were considered to be observed. The data were corrected for Lorentz and polarisation factors; no absorption corrections were applied. The structure was solved by direct methods and all full weight non-hydrogen atoms were refined anisotropically. The positions of the C(7) and C(12) hydroxy hydrogen atoms of the steroid components were determined from a ΔF map, those of the included water molecules could not be located. The positions of the hydrogen atoms were idealised C–H = 0.96 Å, O–H = 0.90(3) Å, assigned isotropic thermal parameters U(H) = 1.2 U_{eq}(C), allowed to ride on their parent carbon or oxygen atoms. Refinement was by full-matrix least squares based on F^2 to give R_1 = 0.0838, wR_2 = 0.227. Computations were carried out on a SGI IRIS XZ 4000 Workstation using the SHELXTL program system Version 5.03. Atomic coordinates, bond lengths and angles, and thermal parameters have been deposited at the Cambridge Crystallographic Data Centre. See Information for Authors, Issue No. 1.

§ The largest torsional difference occurs about the C(3)–N(38) bond. The plane of the aromatic spacer unit is inclined by 7 ° to the mean plane of the macrocycle in molecule (1) and by 10 ° in molecule (2).

¶ In this case the analogous methyl ether was investigated in order to simplify the calculations.

‖ The fit of guest in host is, however, rather tight, which may explain the modest binding constant (*cf.* ref 5b).

References

1 A. P. Davis and J. J. Walsh, *J. Chem. Soc., Chem. Commun.*, 1996, 449.

2 A. P. Davis and J. J. Walsh, *Tetrahedron Lett.*, 1994, **35**, 4865.

3 J. Inanaga, K. Hirata, H. Saeki, T. Katsuki and M. Yamaguchi, *Bull. Soc. Chem. Jpn.*, 1979, **52**, 1989.

4 F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, C. Caufield, G. Chang, T. Hendrickson and W. C. Still, *J. Comput. Chem.*, 1990, **11**, 440.

5 Carbohydrate receptors derived from cholic acid: (*a*) R. P. Bonar-Law, A. P. Davis, and B. A. Murray, *Angew. Chem., Int. Ed. Engl.*, 1990, **29**, 1407; (*b*) K. M. Bhattarai, R. P. Bonar-Law, A. P. Davis and B. A. Murray, *J. Chem. Soc., Chem. Commun.*, 1992, 752; (*c*) R. P. Bonar-Law and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1995, **117**, 259.

6 Other carbohydrate receptors operating *via* hydrogen bonding in organic solvents: Y. Kikuchi, Y. Tanaka, S. Sutaro, K. Kobayashi, H. Toi and Y. Aoyama, *J. Am. Chem. Soc.*, 1992, **114**, 10302 (and references cited therin); R. Liu and W. C. Still, *Tetrahedron Lett.*, 1993, **34**, 2573; P. B. Savage, S. K. Holmgren, J. M. Desper and S. H. Gellman, *Pure Appl. Chem.*, 1993, **65**, 461; G. Das and A. D. Hamilton, *J. Am. Chem. Soc.*, 1994, **116**, 11 139; J. Cuntze, L. Owens, V. Alcázar, P. Seiler and F. Diederich, *Helv. Chim. Acta*, 1995, **78**, 367; For relevant work on cyclitols, see: C.-Y. Huang, L. A. Cabell and E. V. Ansyln, *J. Am. Chem. Soc.*, 1994, **116**, 2778.

Received, 7th August 1995; Com. 5/05297A