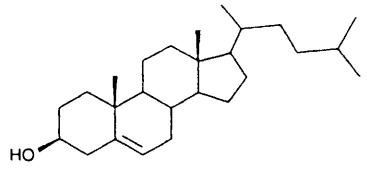
Cholaphanes *et al.*; Steroids as Structural Components in Molecular Engineering

Anthony P. Davis
Department of Chemistry, Trinity College, Dublin 2, Ireland

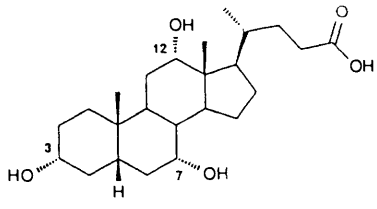
1 Introduction

One of the ultimate aims of Chemistry must be to establish itself as (or give birth to) an *engineering* discipline, within which molecular-scale artefacts may be devised and assembled with the confidence which is currently possible at the macroscopic level. This goal may be some way off, but it provides orientation for an increasing proportion of chemical research in which synthetic targets, often of some complexity, are chosen by rational design to serve theoretical or practical purposes. Some of the most interesting challenges of this type involve the assembly of quite elaborate, extended structures in which spatially separated elements combine to achieve an overall effect. Examples might be molecules which can recognize and bind others, 'artificial enzymes' which can catalyse transformations in bound molecules, and systems which can reproduce themselves or otherwise store and process information at the molecular level.[1] In all of these areas, one of the central problems is the flexibility of most organic molecules. The desired properties will result not only from the presence of the various elements but also from their relative dispositions in space and the three-dimensional shape of the overall assembly. Hence, for most applications, there will be a requirement for molecules with well-defined geometries in which conformational freedom, though probably not eliminated, is kept under close control.

This criterion can be met in principle by designs based on flexible frameworks, in which the necessary restraint is achieved by careful adjustment of non-covalent interactions (the most obvious example being in proteins). However, the difficulty of predicting the end result in such cases suggests that, in practice, there will always be a strong reliance on covalent bonds to provide structural definition. Unfortunately we are supplied with rather few rigid units to use as building blocks, and it is therefore sensible that the potential of each should be explored in depth. When we first surveyed the area a few years ago, it seemed that one fragment which had been under-utilized was the steroid nucleus. It is one of the largest rigid units which is readily available, presents two options for substitution (axial or equatorial) at most positions, and occurs in homochiral form. Moreover, because of the importance of steroids in biochemistry and medicine, their chemistry is understood in great detail. Thus the 'molecular engineer' can access extensive information on potential transformations, spectroscopic properties *etc.*



(1)



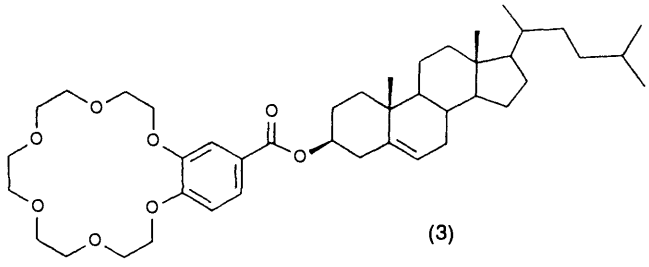
(2)

There are many steroids which are commercially available and might be chosen as starting materials for more elaborate frameworks. However, there are two which stand out on grounds of cost: cholesterol (1) at £0.12/g and cholic acid (2) at £0.16/g (prices from the current Aldrich catalogue). Cholic acid, in particular, will feature strongly in this article. Cholesterol is functionalized at one end and can easily be appended to a structure; moreover oxidative degradation of the side chain gives a second point of attachment, allowing the nucleus to be used as a rigid spacer. However, the lack of functional groups in the central portion of the framework limits cholesterol's potential. In contrast, cholic acid has four functional groups which are fairly evenly spaced around the molecule. Of course, little could be done if it were not possible to differentiate between these groups, and at first sight the fact that three are secondary hydroxyls might appear to present difficulties. However, another valuable feature of the steroid nucleus is its inherent *asymmetry*. No two positions are equivalent, and it is often possible to exert a surprising degree of control in synthetic transformations. In the case of (2), the C3-OH is equatorial while the others are axial, allowing the former to be derivatized selectively. The C7 and C12 hydroxyls are more difficult to distinguish, but, as will be described later, the problem was solved many years ago as part of classical steroid chemistry.

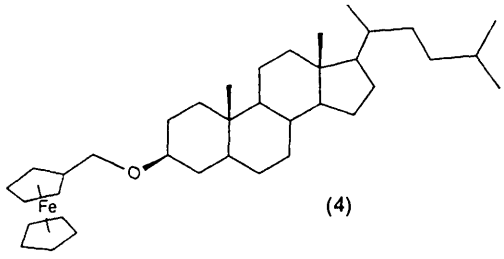
2 Steroids as Appendages and Rigid Spacers

For most of its scientific history (outside the biomedical area) the steroid nucleus has been viewed principally as a 'rigid lump of grease'. It is valuable for promoting liquid crystallinity[2] and is also useful in the study of hydrophobic aggregates such as lipid

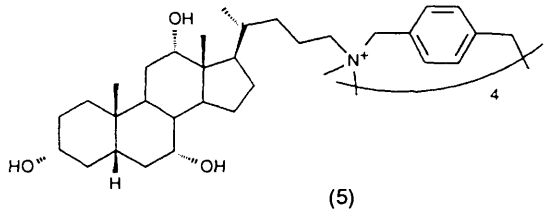
Tony Davis was born in Watford, Hertfordshire in 1954. He received a B.A. in Chemistry from Oxford University in 1977, then stayed on for a D.Phil. under Dr. G. H. Whitham and two years postdoctoral work with Professor J. E. Baldwin. In 1981 he moved to the ETH Zürich as a Royal Society European Exchange Fellow working with Professor A. Eschenmoser, then in 1982 was appointed as a Lecturer in Organic Chemistry at Trinity College, Dublin. In 1993 he was elected to a Fellowship of the College. His research interests include 'molecular engineering' (as discussed in this article) and the application of organosilicon reagents to organic synthesis.



(3)



(4)



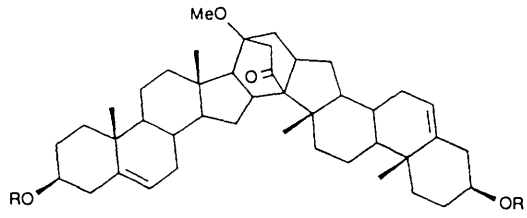
(5)

membranes [3] It continues to be employed for both purposes, sometimes in quite sophisticated systems Thus recent reports describe (a) the cholesteryl crown ether (3) as a component of liquid crystalline phases which responded enantioselectively to the inclusion of chiral anions,[4] (b) redox-responsive vesicle formation by ferrocene derivative (4),[5] and (c) the use of cholanoate units to position the cationic cyclophane (5) in a bilayer [6] An elegant application is the use of the steroidal appendages in (6) to position the porphyrin nucleus at a certain depth within a bilayer membrane, giving a system capable of promoting the regioselective oxidation of matching substrates [7]

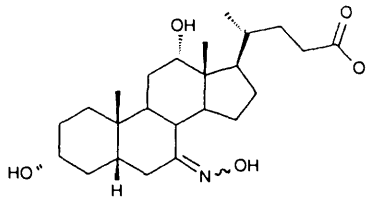
The use of steroid-based spacers to study long-range intramolecular interactions also has a long history As far back as 1965, the rigid dimeric framework in (7) was employed to separate donors and acceptors of excitation energy,[8] and simpler, monomeric spacers have been used in several investigations of intramolecular electron transfer [9]

3 Functionalized Steroids in Biomimetic Chemistry

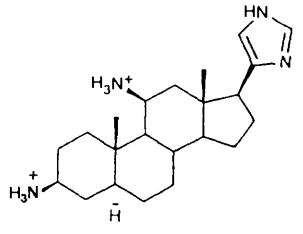
The potential of the steroid nucleus for organizing functional group arrays has perhaps been less widely appreciated However, in the field of biomimetic chemistry it was realized some years ago that by combining both water-solubilizing and catalytic groups on a steroidal framework it might be possible to achieve realistic modelling of enzyme action By mobilizing the framework in the aqueous phase, the water-solubilizing groups would allow the hydrocarbon surfaces to bind substrates by hydrophobic interactions, and correctly positioned catalytic groups would then induce the desired transformations Early examples were the 'synthetic acylases' (8) and (9) reported respectively by the groups of F M Menger[10] and J P Guthrie [11] Both molecules caused substrate-selective accelerations in the hydrolysis of certain *p*-nitrophenyl esters, but were limited by their inability to encapsulate their substrates and a related tendency to aggregate in aqueous solution



(7) R R = donor + acceptor

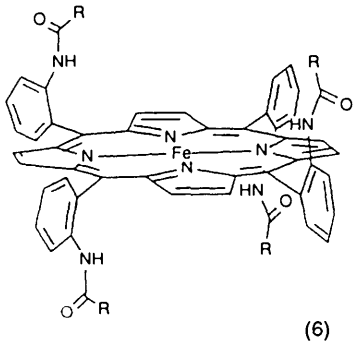


(8)

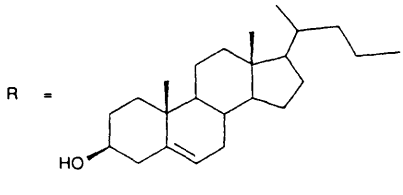


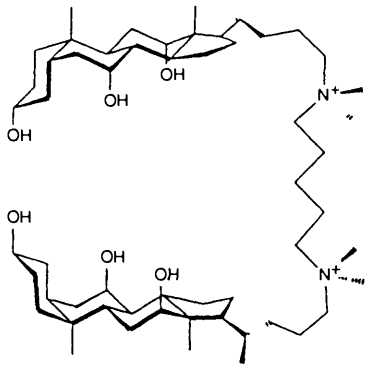
(9)

The use of a dimeric framework was a fairly obvious step forward It would allow a single molecule of the model enzyme to surround its substrate, promoting the formation of 1 1 complexes at the expense of micelles *etc* The complexes would presumably have better structural definition, and provided that the catalytic groups could be positioned appropriately one might expect greatly improved activity and selectivity An initial move in this direction was made by J McKenna and co-workers, who described models derived from the head-to-head linkage of

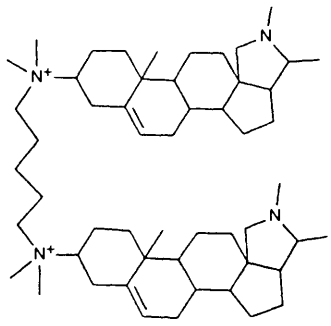


(6)

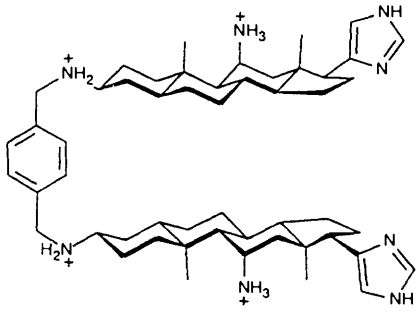




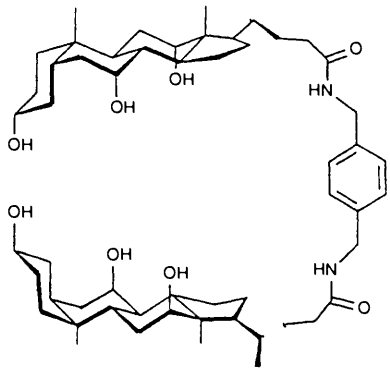
(10)



(11)



(12)



(13)

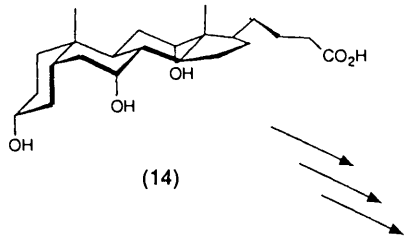
cholic acid, as in (10), or connessine, as in (11) [12] In comparison with monomeric analogues, these structures displayed enhanced abilities for non-micellar binding of organic molecules in aqueous media However, they were not supplied with catalytic functionality and no attempt was made to demonstrate true enzyme mimicry Meanwhile the group of Guthrie extended the work on imidazolyl-substituted steroids by synthesizing dimeric versions, culminating in (12) [13] This system showed considerable improvement over monomeric analogues such as (9), giving impressive accelerations for the hydrolysis of selected esters (up to 550-fold relative to imidazole) It was, on the other hand, clearly quite troublesome to prepare, thus illustrating a principle which has also governed our own experience, while the steroid nucleus has undoubted potential as an 'engineering component', its exploitation requires synthetic effort which, though often rewarding, is rarely trivial

In McKenna's cholanic dimer (10), the cholic acid units were used in much the same way as they are by nature, *i e* to provide a hydrophobic environment for binding organic molecules (cholic acid is the principle ingredient of the bile salts, which are employed as surfactants in living organisms) More recently it was recognized by C J Burrows that it might be used in an inverse fashion in organic media, with the hydrocarbon surface controlling solubility and the cofacial hydroxyl groups providing binding sites for polar molecules The dimeric structure (13) was prepared, and variable temperature NMR studies indicated that it was indeed capable of binding an organic-soluble glycoside, to an unquantified extent, in $CDCl_3$ [14]

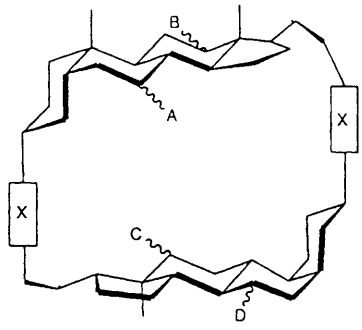
4 Steroid-derived Macrocycles; 'Cholaphanes'

The foregoing section serves as a good introduction to our own contributions We also have been hoping to mimic nature by constructing synthetic analogues of enzymes and receptors, basing our strategy on the use of steroid-derived frameworks to organize functional group arrays We were drawn to cholic acid for the reasons described in Section 1, and also because a 'side-on' view of the molecule [as in (14)] suggested that it might be susceptible to incorporation in macrocyclic structures The linear dimeric systems described above suffer from the lack of an enforced cleft or cavity for substrate binding, limiting their ability to bind strongly and selectively, and also to position catalytic groups accurately on either side of the substrate Macrocyclic structures would have far less freedom, and could be designed to be highly rigid and pre-organized if so desired In cholic acid there is a *cis*-ring junction between rings A and B, which imparts a curvature to its skeleton and means that an equatorial substituent at C3 is directed at *ca* 90° to the main portion of the nucleus It seemed that a rather natural development of the structure would be to introduce a C3α spacer group, and then use the carboxyl group at the far end to form cyclodimers as in (15) By controlling the substitution pattern at C7 and C12 of the monomer units it should be possible to position up to four different binding/catalytic groups around the periphery of the cavity, giving a versatile system for studies in biomimetic and/or supramolecular chemistry

In the initial work it made sense to retain the 7α- and 12α-oxygens, and aim for organic-soluble macrocycles with converging polar functionality Our first instinct was to keep the 3α-oxygen atom as well, and introduce the spacer by making an ether linkage (*e g* to a benzyl unit) However, we soon recognized that any such plan would result in a rather flexible framework, and that the *direct* attachment of a rigid spacer to the steroid would be far preferable The obvious choice for the spacer was a *para*-substituted benzene ring, apart from being structurally suitable, one might hope that the ring would induce useful NMR effects when we came to study the properties of the system This raised questions perhaps more typical of a project in natural product synthesis than in supramolecular or biomimetic chemistry The aryl unit would need to carry functionality for use in the cyclodimerization, and a complementary group would be required in the steroidal side-chain Bearing these



(14)



(15) A D = binding/catalytic functionality
X = spacer

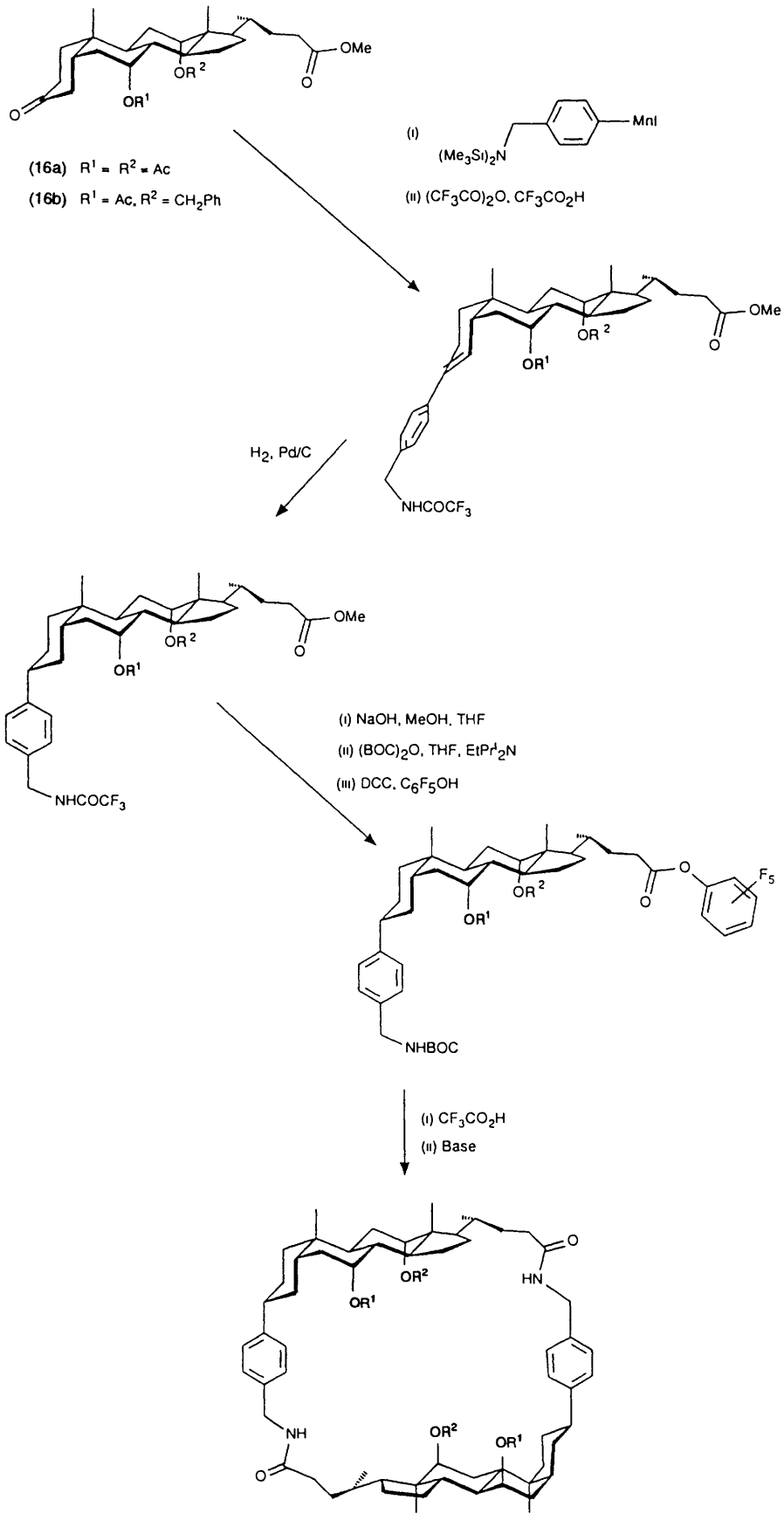
points in mind, how might the spacer be introduced at C3 with (a) the correct stereochemistry, and (b) acceptable chemoselectivity?

One of our answers is shown in Scheme 1 [15] We chose to use amide bond formation for the cyclodimerization, partly because it is easily accomplished and partly because the macrocycle would be chemically robust This required retention of the side-chain carboxyl in suitably protected form, and incorporation of a protected amino group in the spacer The new carbon–carbon bond was formed with the help of one of the newer types of organometallic reagent, an organomanganese derivative [16] This had the advantage of being completely inert to ester groups, so that acetyl protection could be used at positions 7 and 12 (*vide infra*) and, critically, the side-chain carboxyl could be carried through as its methyl ester The stereochemistry was controlled by a traditional method, *i e* catalytic hydrogenation of the less-hindered face of a double bond The final step, cyclodimerization to 'cholaphanes' (17), was the one which initially caused us the greatest concern The shortest route around the macrocyclic framework encompasses 38 atoms Although ring closure would be assisted by the rigidity of the monomeric units (reducing the activation entropy for cyclization), we were still afraid that much hard-won material might be lost as polymer In fact the transformation proved to be remarkably easy It could be accomplished in stepwise fashion *via* a linear dimer (demonstrating that products lacking C_2 symmetry could be made if required) or, as shown, directly from a monomer unit The most effective method employed pentafluorophenyl ester intermediates (as illustrated) and gave up to 90% yield of crystalline macrocycle

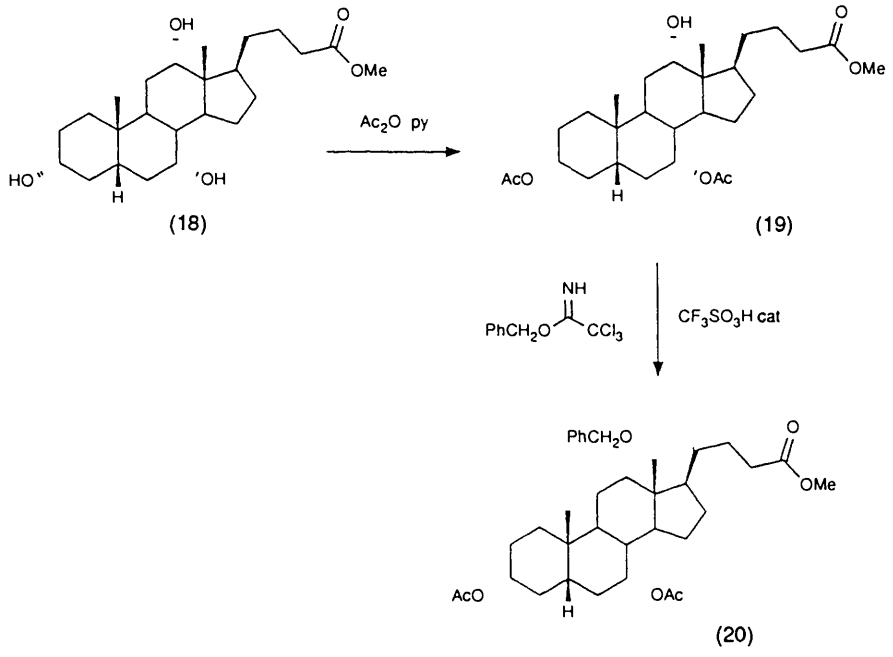
The sequence in Scheme 1 was carried out for two series of compounds Initially we worked from the diacetoxy ketone (16a) through to tetraacetoxycholaphane (17a), and then (after treatment with hydroxide) to tetrahydroxycholaphane (17b) It is worth noting that, although several steps were involved, the overall yield of (17a) from cholic acid was nearly 40% For the second series, our aim was to show that we could control the substitution pattern at carbons 7 and 12, and thus differentiate between the two faces of the macrocycle As mentioned in Section 1, the problem of distinguishing between the axial 7α- and 12α-OH groups in cholic acid is non-trivial, but had been solved in the 'classical' period of steroid chemistry Acetylation of methyl cholate (18) with acetic anhydride/pyridine gives the 3,7-diacetate (19) with reasonable selectivity,[17] allowing isolation of the crystalline product in *ca* 65% yield (Scheme 2) The preferential acylation at position 7 as opposed to 12 is actually quite curious, as the former is apparently the more hindered due to the axial orientation of C4 (with respect to ring B) Indeed, under other acylation conditions the 12α-OH is the *more* reactive (a fact which we were later able to exploit[18]), and the formation of (19) seems to be a particular quirk of the pyridine-catalysed reaction [19] Benzylation of (19) gave (20), which was converted into (16b), and thenceforth to cholaphane (17c) Selective deprotection could be accomplished at either face, giving two further macrocycles (17d) and (17e) (Scheme 1)

During the early phases of this work, our main concern was to show that the cholaphane framework could be assembled with reasonable ease, and was thus a viable starting point for artificial enzymes *etc* However, we came to realize that our 'demonstration models' (17) might in themselves have quite interesting properties Molecular modelling indicated that if they adopted an open conformation, they would enclose a cavity of cross-sectional area 40—50 Å^2 (see *e g* Figure 1) Including the annular amides, the cavity would be surrounded by six polar functional groups in a fully three dimensional arrangement which might be likened to a (highly) distorted octahedron The cholaphanes might therefore be well suited to act as receptors for small molecules with a 3D array of divergent functionality The obvious targets were carbohydrate nuclei In spite of the importance of these units as carriers of biological information,[20] there was little work reported on modelling their recognition The only successful system was the resorcinol–aldehyde tetramer (21) investigated by Aoyama and co-workers,[21] however this molecule was clearly incapable of *encapsulating* a carbohydrate and apparently operated *via* face-to-face interactions (the report of Burrows,[14b] referred to above, appeared soon after we started our work)

Accordingly we used NMR to investigate the interaction of (17a—e) with the organic-soluble glucoside (22) in $CDCl_3$ [22] Although none of the acetylated macrocycles showed any sign of binding, we were delighted to find that addition of the glucoside caused significant changes in the spectra of (17b) and (17d) Spectra from an experiment involving tetraol (17b) are shown in Figure 2 As glucoside is added (moving up the Figure) the signal at δ 5 67 (N*H*) moves sharply downfield, the AB system at δ 4 3—4 5 (CH_2N) separates and the HN–CH vicinal couplings change dramatically Analysis of the movements was consistent with 1 1 complex formation, with a binding constant of 1740 (± 200)M^{-1} Similar effects were observed with dibenzyloxy diol (17d), yielding a binding constant of 700 (± 100) M^{-1} In the latter case, it was interesting to note that the AB quartet due to the *O*-benzyl methylene protons also moved appreciably While we have no direct proof that the glucoside head-group was entering the cavities of the macrocycles, several factors point in this direction First, the size of the binding constants suggests that several hydrogen bonds are being formed Secondly, the NMR movements suggest major changes in the conformation of the macrocycles which, in the case of (17d) at least, involve widely-separated parts of the molecule Finally, computer-based molecular modelling confirmed that the hypothesis is reasonable Making the assumption that both annular amides were acting as H-bond donors (supported by the N*H* NMR movements and also by IR data), we were able to devise the configuration shown in Figure 3 Based on a relatively unstrained

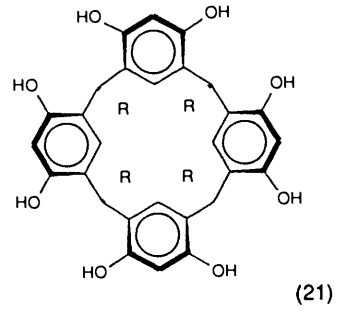

(16a) R¹ = R² = Ac
(16b) R¹ = Ac, R² = CH₂Ph
(i) (Me₃Si)₂N MnI
(ii) (CF₃CO)₂O, CF₃CO₂H
H₂, Pd/C
(i) NaOH, MeOH, THF
(ii) (BOC)₂O, THF, EtPrⁱ₂N
(iii) DCC, C₆F₅OH
(i) CF₃CO₂H
(ii) Base
NaOH
(17a) R¹ = R² = Ac
(17b) R¹ = R² = H
NaOH
(17c) R¹ = Ac, R² = CH₂Ph
(17d) R¹ = H, R² = CH₂Ph
(17e) R¹ = Ac, R² = H
H₂, Pd/C

Scheme 1



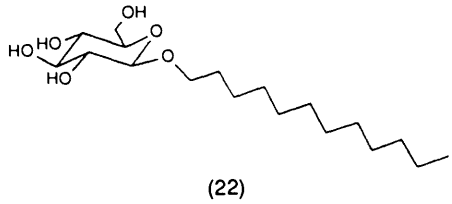
(18) (19) (20)

Scheme 2



(21)

R = $(CH_2)_{10}CH_3$



(22)

conformation of the macrocycle, the modelled complex is held together by six intermolecular hydrogen bonds Although it may well be a figment of our (computer-aided) imagination, with no basis in reality, one can reasonably argue that it can only be superseded by an even more favourable arrangement

If the carbohydrate was indeed entering the cavities of (17b/d), an expected consequence was that the binding constants should be quite sensitive to a change in substrate We felt that an interesting test might be to investigate binding to a complete set of stereoisomeric glucosides (α/β and D/L) Results from experiments employing (17b) and the octyl glucosides are shown in Table 1 [23] Not only was there significant diastereoselectivity (*ca* 5 5 1, compare entries for β-D and α-D) but also appreciable enantioselectivity (*ca* 3 1, β-D *vs* β-L) The latter is of course made possible by the chirality of the macrocyclic framework The level of selectivity compares poorly with that found in natural systems but, considering the flexibility of (17), has to be seen as encouraging

Considering the next phase of our cholaphane programme, we focused on two medium-term objectives One was to develop carbohydrate receptors with improved potency and controllable selectivity, and the other was to carry the macrocycle into aqueous solution where realistic enzyme modelling might be attempted For both purposes, it was clear that major alterations would be required to framework (17) First, we would need to increase the rigidity of the structure so that derivatives would be more pre-organized to bind target substrates, less able to adapt themselves to other guests, and less able to find collapsed conformations in which guests are excluded (relevant to aqueous solution, where hydrophobic forces would tend to bring the organic surfaces together) We were also aware that unless conformational flexibility could be reduced, there was little prospect of predicting binding/catalytic behaviour in a rational fashion

Secondly, we needed a method of controlling the solubility of the framework For studies in chloroform the tetraol (17b) had barely sufficient solubility ($\leq$ 1mM), and there was every likelihood that increasing the rigidity would make the problem worse A rigid molecule moving from the solid state into solutions gains just translational and rotational entropy, a flexible molecule acquires other freedoms and thus gains more from undergoing such a transition Any plan for increasing cholaphane rigidity should therefore make provision for the introduction of flexible solubilizing substituents, oriented in such a way as not to interfere with the organized core of the framework For studies in water it was clear that polar or ionic substituents would be necessary, and again it would be desirable that they should be directed away from the centre

There was one alteration which appeared synthetically feasible, and which seemed likely to improve matters in both the above respects Considering the sequence in Scheme 1, it was apparent that we were being somewhat wasteful of the functionality at C3 of our starting materials (16) In principle there was an opportunity to introduce two substituents at this centre an aryl spacer in the equatorial orientation and a solubilizing group in the axial orientation If the latter were fairly bulky it would restrict rotation about the C–aryl bond, such that the spacer would be positioned as shown in partial structure (23) This would remove some of the conformational freedom of the macrocycle (calculations indicated that rotation about the

Figure 1 Computer-generated space-filling model of tetrahydroxycholaphane (17b) This structure is one of 36 energy minima within 4 5 kcal mol $^{-1}$ of baseline located in a search employing the QUANTA/CHARMm software package Although their shapes varied considerably all were open conformations enclosing substantial cavities

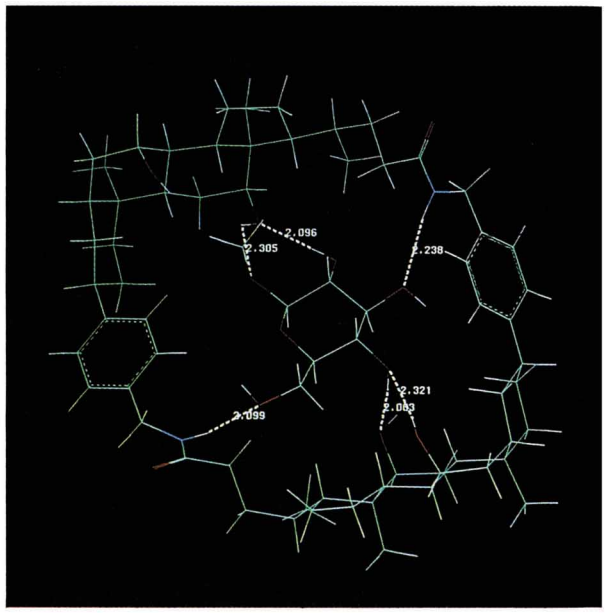


Figure 3 A possible conformation for the complex of (17b) and methyl β-D-glucopyranoside (acting as a model for (22)) Intermolecular hydrogen bonds are shown as white dotted lines, accompanied by the corresponding distances

C–aryl bonds in (17) should be relatively unrestricted), and also give a well-defined surface to the cavity

The idea was first realized as shown in Scheme 3 [24] Satisfying features of this sequence were the chemo- and stereoselectivity of the reactions used to introduce the spacer The Knoevenagel reaction to give (24) could be performed under very mild

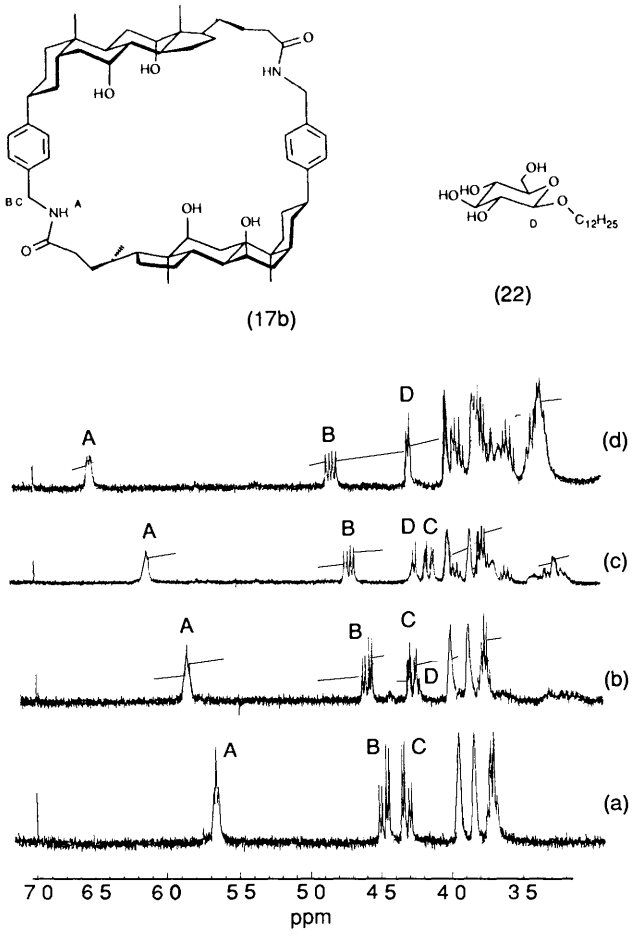
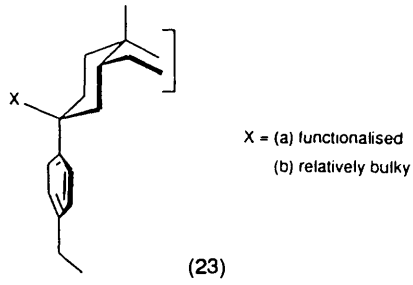


Table 1 Binding constants for stereoisomeric octyl glucosides with cholaphane (17b) in $CDCl_3$ at 25 °C, as determined by ^{1}H NMR titration experiments

	K_a (M $^{-1}$)
β-D	3100 (± 14%)
α-D	582 (± 7%) 545 (± 4%)*
β-L	1000 (± 14%)
α-L	1030 (± 10%)

* Separate experiment conducted at end of series to check reproducibility

Figure 2 ^{1}H NMR spectra from a titration experiment involving cholaphane (17b) and glucoside (22), with $CDCl_3$ as solvent In the initial spectrum of (17b) [1 1 mM, spectrum (a)], the amide NH A' appears at 5 67 p p m , while the benzylic protons 'B,C' are quite closely grouped around 4 4 p p m and show roughly equal couplings to the NH In the presence of (22) at increasing concentrations [0 31 mM in spectrum (b), 0 86 mM in (c), and 2 92 mM in (d)] signal A moves downfield towards an estimated limiting value of 6 75 p p m Signal B moves downfield (estimated limiting $\Delta\delta$ 0 32 p p m) and shows increased coupling to A, while signal C moves upfield (est $\Delta\delta$ 0 44 p p m) and shows decreased coupling to A The shifts clearly suggest major conformational changes in the macrocyclic skeleton, consistent with complex formation The doublet 'D' due to the anomeric proton of the carbohydrate, moves very slightly downfield during the experiment As the proportion of carbohydrate bound is at its greatest at *low* concentrations, it can be inferred that this proton experiences a weak shielding effect within the complex



(23)

conditions, the organocuprate (25) was inert to the ester groups in (24), and the aryl group was introduced almost exclusively from the equatorial direction. It was also pleasing that the final product (26) could be analysed by *X*-ray crystallography, which showed the spacer groups in the expected orientation and confirmed the ability of the framework to encompass small guest molecules (Figure 4). We were less happy about the use of oxygen-based functionality in organometallic (25). This arose because we were unable to find a form of *N*-protection compatible with both the cuprate and its Grignard precursor (for further discussion, see Section 6). However, the replacement of oxygen for nitrogen proceeded smoothly and in good yield.

A more serious problem concerned the dicyanomethyl groups in cholaphane (26). There were a number of ways in which we had hoped to elaborate them into either flexible or ionic substi-

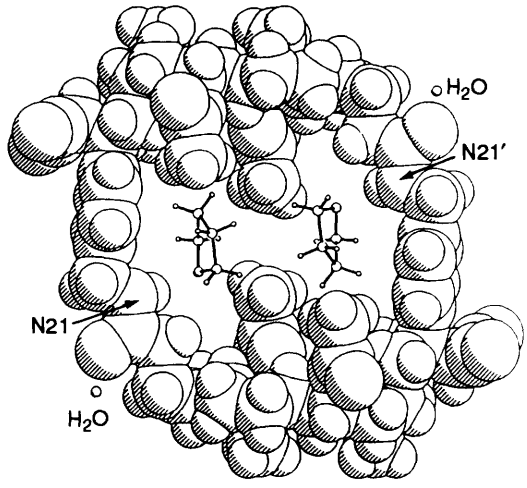
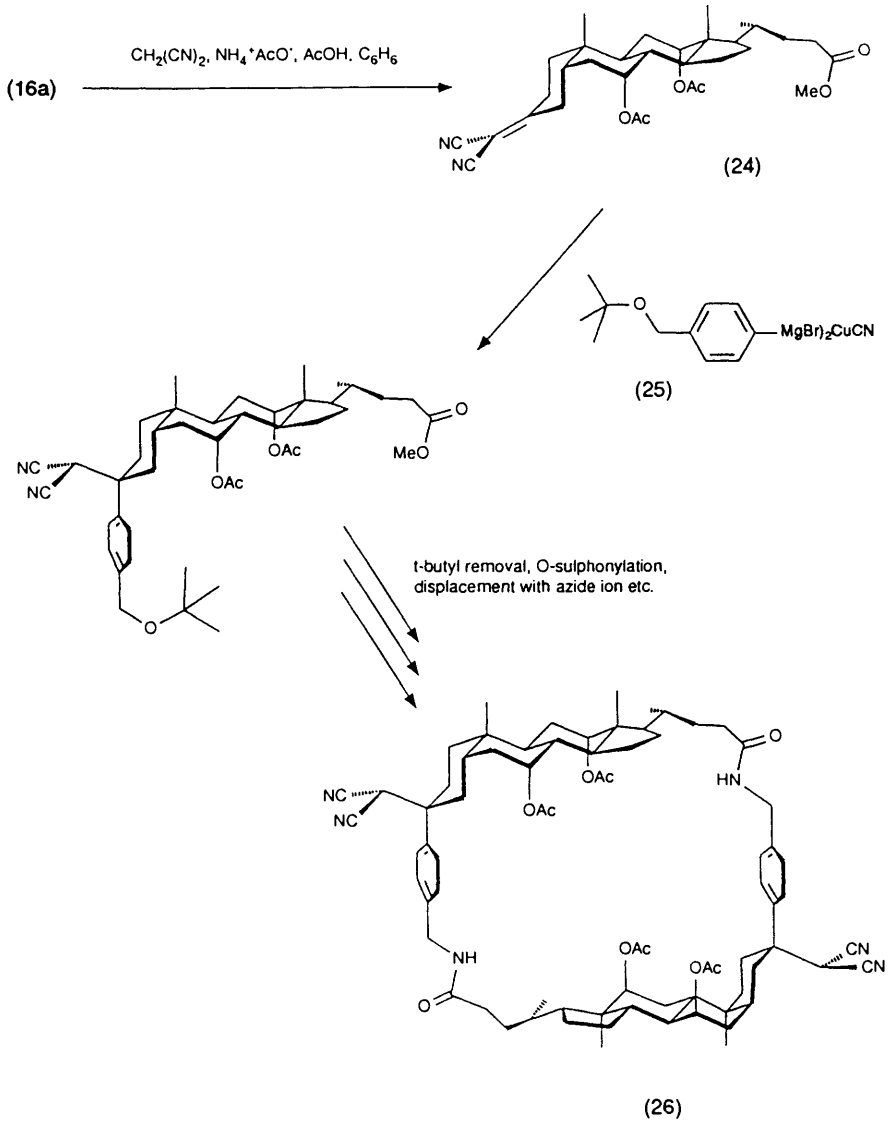


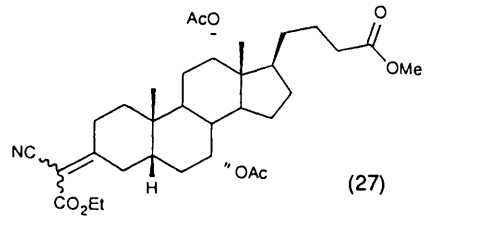
Figure 4 *X*-Ray crystal structure of cholaphane (26), shown in space-filling mode. The cavity includes two molecules of THF, positioned by N–H···O hydrogen bonds.



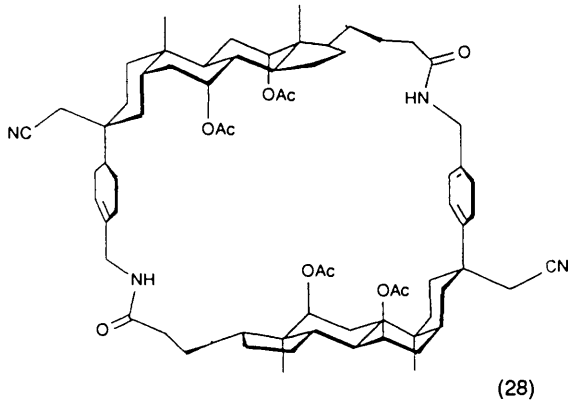
(26)

Scheme 3

tuents, but they proved disappointingly inert (possibly because of their unusually hindered environment) The obvious solution was to employ an alternative to malononitrile in the initial condensation Here again we met with unexpected difficulties, (re)discovering that high-yielding Knoevenagel condensations on cyclohexanones are the exception rather than the rule However, one of the exceptions is provided by cyanoacetate reagents, and recent work has shown that (16a) may be converted into (27) and thenceforth *via* arylcuprate addition, deethoxycarbonylation *etc* to cholaphane (28) [25] Although the cyanomethyl groups in (28) are less bulky than their dicyano relatives, they seem to be more tractable and we are hopeful that the methodology will prove viable for the synthesis of a range of externally-functionalized cholaphanes



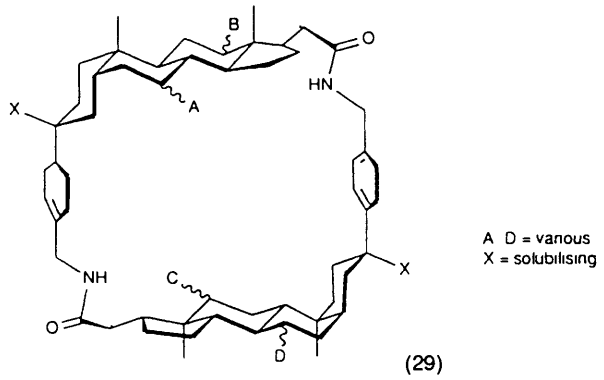
(27)



(28)

In (26) and (28), which might be described as 'second-generation' cholaphanes, only modest progress has been made towards a rigid, predictible macrocyclic framework However, for the third generation we plan a modification which should effectively complete the journey There are practical, large-scale methods for shortening the side-chain of cholic acid by two carbons, and with such 'bis-nor' derivatives as starting materials macrocycles of the general form (29) should be accessible Computer-based molecular modelling suggests that these structures should have very little flexibility indeed, and could prove very informative in studies of molecular recognition and enzyme action

A final point for this section concerns the move into aqueous solution The 'second-generation' methodology allows us to attach water-solubilizing functionality at two points on the framework, but it is likely that more will be required An option which seems attractive in principle is to replace the 7,12α-OH groups in cholic acid by β-directed NH_3^+ units, *via* S_N2 displacements Although this transformation has proved less easy than we might have hoped, we have been able to achieve a workable procedure using azide as the nucleophile and sulphonate leaving groups [26] Water-soluble cholaphanes are thus a realistic goal for the future



(29)

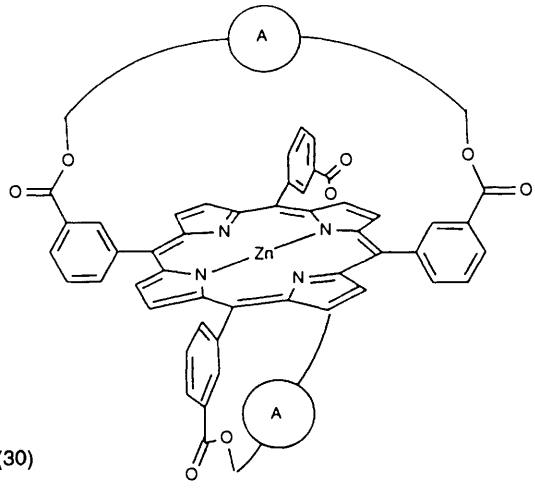
5 Steroid-derived Macrocycles; Directly-linked Oligo-cholane Units

Formula (15) does not, of course, represent the only way of assembling cholane units into macrocyclic structures An alternative is to make a direct linkage between two steroidal components, and complete the macrocycle with a third fragment (which may or may not be derived from the steroid) We have been exploring some possibilities of this type, as discussed below However, most of the running has been made by R P Bonar-Law, who synthesized cholaphanes (17) as his Ph D work in Dublin, and then moved to Cambridge to collaborate with J K M Sanders Structures investigated by the Cambridge group are the steroid-capped metalloporphyrin (30),[27] and the 'cyclocholates' (31)[28] and (32) [29] In (30) the two cholic acid units are joined by formation of a bis-lactone, the resulting cap having two hydroxyl groups which can interact with substrates bound to the metal The molecule showed interesting selectivity for binding of hydroxyamines, and was rapidly monoacylated by a 3-carboxypyridine derivative Macrocycles (31) were formed by direct cyclo-oligomerization of 7,12-diprotected cholic acid derivatives, and applied to the binding of alkali metal cations in organic media Cyclocholates (32) were prepared in an analogous fashion from starting materials which had been subjected to a Beckmann rearrangement at the steroidal C12 It was shown that they could self-associate into tubular dimers by amide–amide H-bond formation It has also proved possible to construct a tetrameric cyclocholate bridged on one face by a metalloporphyrin, giving an elegant, bowl-shaped structure with inward-directed hydroxyl groups This molecule has been found to bind morphine by a combination of H-bonding and nitrogen–metal ligation [30]

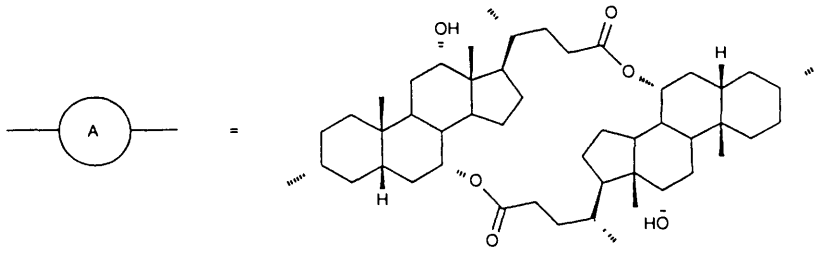
Although the simple cyclocholate frameworks are readily accessible, they are rather flexible for use in pre-organized receptors (unless supplied with extra constraints, as in the porphyrin described above) However, analogous molecules derived from 'bis-nor' cholic acid have been prepared by the Cambridge group, and appear to have considerable potential [31] An approach of our own also involves the bis-nor steroid The 3α-OH has been replaced by an amino group, and two units have then been linked by amide bond formation The resulting dimers may be seen as consisting of two rigid blocks connected by a fairly flexible hinge We hope that by completing the macrocycle with spacers of varying lengths we will be able to generate a family of host molecules of the form (33) Spacers B will be used to introduce functionality and control solubility as well as to tune the size of the cavity

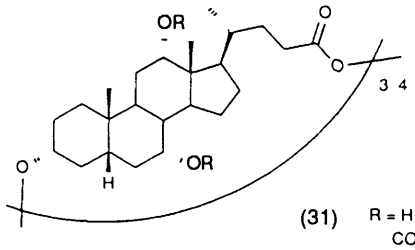
6 Concluding Remarks; Spin-offs and Future Prospects

Molecular engineering as illustrated in this article consists of a collaboration between organic synthesis and physical or physical–organic chemistry For most chemists the ultimate justification would probably lie in the latter area – it is the demonstration of a novel and interesting property which gives the work its meaning However, as progress continues, it is likely that the synthetic aspect will be of increasing importance Structures will become more elaborate, their designs will be subject to more precise constraints, and their syntheses will present greater difficulty To the physical-organic chemist this may seem a bleak prospect, but to the synthetic chemist it is an exciting challenge

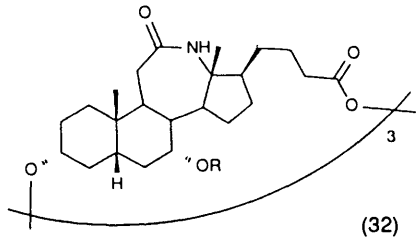


(30)

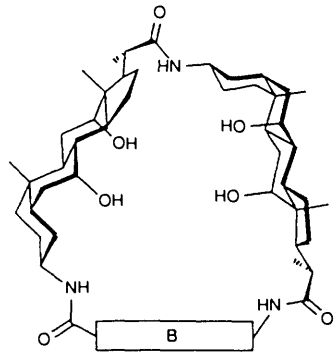




(31) R = H $CH_2OCH_2CH_2OCH_3$
$COCH_2O(CH_2CH_2O)_2CH_3$



(32)

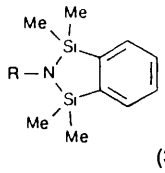


(33)

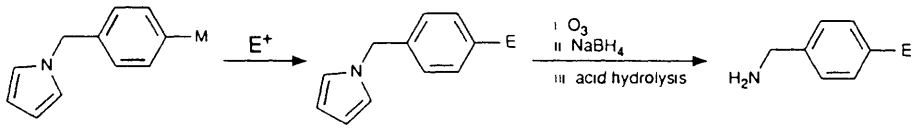
If, as seems likely, work in the area begins to converge with natural products synthesis, it will also yield some of the benefits of the latter In particular it will highlight general problems of methodology and will stimulate the finding of solutions In a small way this has already happened within our own programme For example, our experience in introducing *p*-aminomethylphenyl spacer groups (Schemes 1 and 3) has made us aware of the paucity of *N*-protecting groups which are compatible with strongly basic reagents *N,N*-Bis(trimethylsilyl) was satisfactory in the first sequence, but was too unstable to be truly convenient This prompted the development of the 'Benzostabase' protecting group, as in (34) [32] In the synthesis of (26), we found that Si-based groups were incompatible with the organocuprate (probably because of N-basicity) and we were forced to resort to the *O*-substituted reagent (25) However, we later developed the use of the pyrrole ring, as in Scheme 4, this probably stands as the only group which is compatible with organometallic centres, essentially non-basic, and removable under reasonably non-aggressive conditions [25b]

As illustrated in the foregoing sections, the steroid nucleus can be employed to construct a variety of complex, extended molecular frameworks It hardly needs saying that its potential is not yet exhausted, and that many further applications may be expected However, two points are worth making concerning prospects for the future The first is a general one relating to the use of steroids to construct molecular receptors As mentioned in the introduction, a particular feature of the steroidal framework is its asymmetry Thus, in a C_2-symmetric cholaphane such as (17) or (26), there are 24 steroidal carbons which are different from each other, and which therefore give distinct, identifiable NMR signals Because of the prevalent sp^3 hybridization, these carbons carry *ca* 30 distinguishable proton types, almost all of which should be assignable to ^{1}H NMR signals (after performing this exercise for (17a), the only uncertainties were between pairs of protons in four methylene groups) Many of these protons are directed into the cavity, and may be seen as sensors positioned within the interior wall Provided that the receptor is able to bind its substrate in a single, well-defined orientation, one should be able to obtain remarkably precise structural information on the complex using intermolecular NOEs and other effects

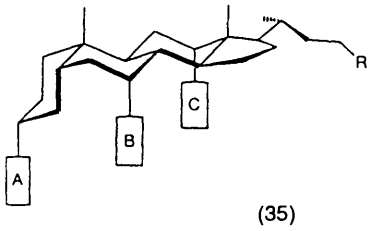
The second and final point relates specifically to cholic acid Most of the work on this molecule has involved its assembly into large, oligomeric structures capable of encapsulating substrate molecules However, there may well be unrealized potential in the single cholic acid unit, given its supply of differentiable, co-directed functionality A wide range of molecules of the general



(34)



Scheme 4



(35)

form (35) should be readily accessible, and various applications can be envisaged A recent paper by Kahne demonstrates one option, describing the glycosylation of the axial hydroxyls to give 'facially amphiphilic' molecules [34] We are actively considering the possibilities of systems (35) in molecular recognition, with particular emphasis on the achievement of enantioselectivity *via* three-point binding

Acknowledgements For the work performed in Dublin, I have to thank my co-workers Richard Bonar-Law, Michael Orchard, Brian Murray, Brian Dorgan, Tom Egan, Khadga Bhattarai, and a number of others who are currently 'on steroids' We are grateful to the groups of Pat McArdle (Galway) and David Williams (Imperial College, London) for *X*-ray crystallography, to Manfred Reetz for collaboration on the Benzostabase *N*-protecting group, to Jeremy Sanders for help obtaining mass spectra, and to Andy Derome for advice concerning NMR spectra Funding has been provided by Eolas (the Irish Science and Technology agency), BioResearch Ireland and (in support of our molecular graphics facility) Loctite Corporation Finally, thanks are due to Diamalt Gmbh for a generous supply of cholic acid

6 References

1 For leading references, see, J Rebek, Jr , *Ang Chem Int Ed Engl*, 1990, **29**, 245, J -M Lehn, *Angew Chem Int Ed Engl*, 1990, **29**, 1304, F N Diederich, 'Cyclophanes', Royal Society of Chemistry, Cambridge, 1991, P L Anelli, P R Ashton, R Ballardini, V Balzani, M Delgado, M T Gandolfi, T T Goodnow, A E Kaifer, D Philp, M Pietraszkiewicz, L Prodi, M V Reddington, A M Z Slawin, N Spencer, J F Stoddart, C Vicent, and D J Williams, *J Am Chem Soc* , 1992, **114**, 193, C Seel and F Vogtle, *Angew Chem Int Ed Engl*, 1992, **31**, 528

2 P J Collings, 'Liquid Crystals', Adam Hilger, Bristol, 1990

3 Leading reference, J -H Fuhrhop and J Mathieu, *Angew Chem Int Ed Engl*, 1984, **23**, 100

4 S Shinkai, T Nishi, and T Matsuda, *Chem Lett* , 1991, 437

5 J C Medina, I Gay, Z Chen, L Echegoyen, and G W Gokel, *J Am Chem Soc* , 1991, **113**, 365

6 J Kikuchi, C Matsushima, K Suehiro, R Oda, and Y Murakami, *Chem Lett* , 1991, 1807

7 J T Groves and R Neumann, *J Am Chem Soc* , 1987, **109**, 5045, *J Org Chem* , 1988, **53**, 3891, *J Am Chem Soc* , 1989, **111**, 2900

8 S A Latt, H T Cheung, and E R Blout, *J Am Chem Soc* , 1965, **87**, 995

9 see *e g* , J R Miller, L T Calcaterra, and G L Closs, *J Am Chem Soc* , 1984, **106**, 3047, Y Kobuke, M Yamanishi, I Hamachi, H Kagawa, and H Ogoshi, *J Chem Soc Chem Commun* , 1991, 895, S -z Zhou, S -y Shen, Q -f Zhou, and H -j Xu, *J Chem Soc Chem Commun* , 1992, 669

10 F M Menger and M J McCreery, *J Am Chem Soc* , 1974, **96**, 121

11 J P Guthrie, *Can J Chem* , 1972, **50**, 3993, J P Guthrie and Y Ueda, *J Chem Soc Chem Commun* , 1973, 898

12 J McKenna, J M McKenna, and D W Thornthwaite, *J Chem Soc Chem Commun* , 1977, 809

13 J P Guthrie, J Cossar, B A Dawson, *Can J Chem* , 1986, **64**, 2456, and refs cited therein

14 (*a*) C J Burrows and R A Sauter, *J Inclusion Phenomena*, 1987, **5**, 117, (*b*) J F Kinneary, T M Roy, J S Albert, H Yoon, T R Wagler, L Shen, and C J Burrows, *J Inclusion Phenomena*, 1989, **7**, 155

15 R P Bonar-Law and A P Davis, *J Chem Soc Chem Commun* , 1989, 1050

16 G Cahiez and J F Normant, in 'Modern Synthetic Methods', Vol 3, ed R Scheffold, Wiley, Chichester, 1983, p 173

17 L F Fieser and S Rajagopalan, *J Am Chem Soc* , 1950, **72**, 5530

18 R P Bonar-Law, A P Davis, and J K M Sanders, *J Chem Soc Perkin Trans 1*, 1990, 2245

19 R T Blickenstaff and J Baker, *J Org Chem* , 1975, **40**, 1579, and refs cited therein

20 see *e g* , T W Rademacher, R B Parekh, and R A Dwek, *Ann Rev Biochem* , 1988, **57**, 785, S -i Hakomori, *Adv Cancer Res* , 1989, **52**, 257, N Sharon and H Lis, *Chem Br* , 1990, **26**, 679

21 Y Aoyama, Y Tanaka, H Toi, and H Ogoshi, *J Am Chem Soc* , 1988, **110**, 634, Y Kikuchi, Y Tanaka, S Sutarto, K Kobayashi, H Toi, and Y Aoyama, *J Am Chem Soc* , 1992, **114**, 10303, and refs cited therein

22 R P Bonar-Law, A P Davis, and B A Murray, *Angew Chem Int Ed Engl* , 1990, **29**, 1407

23 K M Bhattarai, R P Bonar-Law, A P Davis, and B A Murray, *J Chem Soc Chem Commun* , 1992, 752

24 A P Davis, M G Orchard, A M Z Slawin, and D J Williams, *J Chem Soc Chem Commun* , 1991, 612, A P Davis and M G Orchard, *J Chem Soc Perkin Trans 1*, 1993, 919

25 (*a*) A P Davis, T J Egan, M G Orchard, D Cunningham, and P McArdle, *Tetrahedron*, 1992, **48**, 8725, (*b*) A P Davis and T J Egan, *Tetrahedron Lett* , 1992, **33**, 8125 and refs cited therein

26 A P Davis and M G Orchard, *Tetrahedron Lett* 1992, **33**, 5111

27 R P Bonar-Law and J K M Sanders, *J Chem Soc Chem Commun* , 1991, 574

28 R P Bonar-Law and J K M Sanders, *Tetrahedron Lett* , 1992, **33**, 2071

29 R P Bonar-Law and J K M Sanders, *Tetrahedron Lett* , 1993, **34**, 1677

30 R P Bonar-Law, L G Mackay, and J K M Sanders, *J Chem Soc Chem Commun* , 1993, 456

31 R P Bonar-Law, personal communication

32 R P Bonar-Law, A P Davis, and B J Dorgan, *Tetrahedron Lett* , 1990, **31**, 6721, R P Bonar-Law, A P Davis, B J Dorgan, M T Reetz, and A Wehrsig, *Tetrahedron Lett* , 1990, **31**, 6725

33 Y Cheng, D M Ho, C R Gottlieb, D Kahne, and M A Bruck, *J Am Chem Soc* , 1992, **114**, 7319