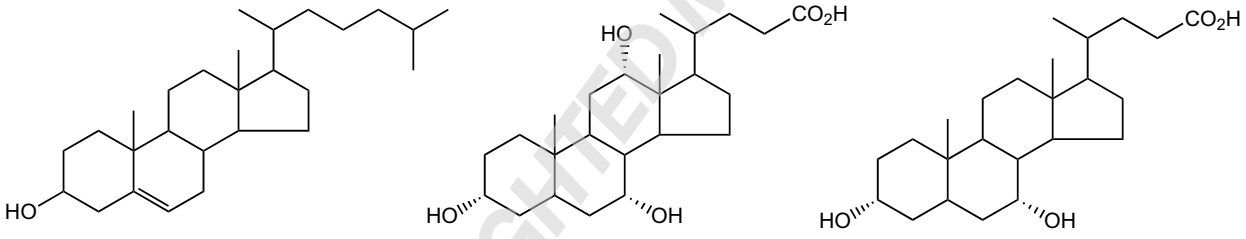
1

Steroids: a Brief History

The early history of steroids devolves almost exclusively about two compounds that had, at the time, been known for decades. These substances, cholesterol (**1-1**) and cholic acid (**1-2**) (Scheme 1.1) are available in large quantities from natural sources. This is arguably explains why these compounds were the first steroids to be obtained in pure crystalline form. Some gallstones, in fact, consist of as much as 90% of the neutral steroid cholesterol. This compound was isolated from those stones as pure crystals well before the birth of organic chemistry. The empirical formula for cholesterol, $C_{27}H_{46}O$, was established as early as 1888 (or alternatively $C_{54}H_{92}O_2$, as the concept of a molecule was at the time still somewhat nebulous and not accepted by all chemists). This compound and many of its derivatives, referred as sterols [from the Greeks *steros* (solid)] are also available from plants. Ox bile from slaughterhouses proved to be a relatively abundant source of bile salts. The acids from acidification of the salts consist largely of cholic acid (**1-2**) and chenodeoxycholic acid (**1-3**). Each of these compounds was obtained as pure crystals at about the same time as cholesterol. The bile acids, it was subsequently found, are formed in the liver by oxidation of cholesterol. They serve as surfactants for absorption of fats from the intestine and also for excretion of cholesterol and other hydrophobic compounds. The relative abundance of pure cholesterol and bile acids focused early research aimed at unraveling the chemical structure of steroids on those two compounds. This research actually preceded the discovery of the hormonal steroids by a good many decades. In view of their lack of biological activity, the investigations aimed at elucidating the structure of cholesterol and of cholic acid was probably undertaken largely as an exercise in structural organic chemistry. Results from these studies markedly facilitated subsequent efforts to assign structures to the so-called sex steroids.



1-1 1-2 1-3

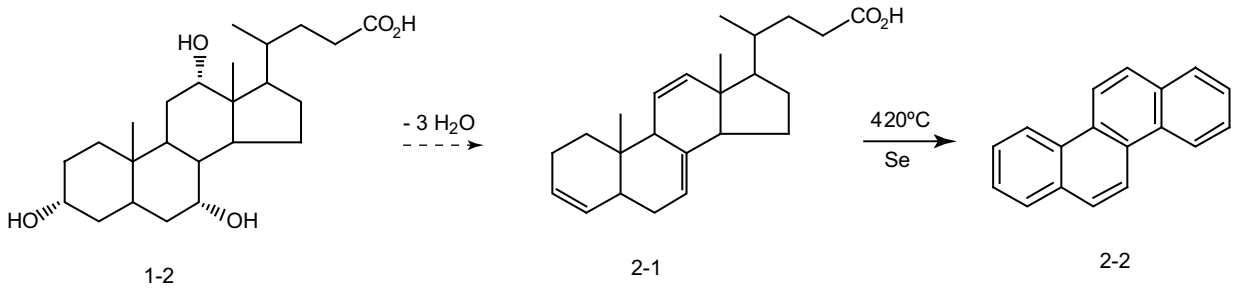
Scheme 1.1

1.1 Structure Determination

1.1.1 Cholesterol and Cholic Acid

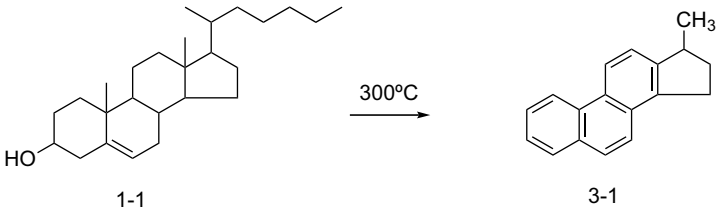
The research aimed at determining the chemical structure of steroids long predated the availability of the instruments that form the backbone of today's work on the determination of the structures of natural products. Although the concept of infrared absorption had already been proposed in the mid-1920s, instruments for determining spectra would not be available until several decades in the future. The phenomenon of nuclear magnetic resonance (NMR) was unknown even to theoretical physicists. Had the concept been proposed, the use of that tool in structural studies awaited the invention of magnetron vacuum tubes as sources of microwave radiation (development of that electronic device was as a direct product of wartime (World War II) research on radar). In the first half of the 20th century, work on structure determination instead relied largely on degradation reactions that would reduce the target to ever smaller molecules until they matched compounds of known structure. The work also relied extensively on combustion analysis for determining elemental composition and Rast molecular weight determinations. Isolation of discrete products from degradation reaction mixtures required great technical skill in those days before the advent of any sort of chromatography. Elegant chemical reasoning played a very large role in interpreting the results of degradation experiments. As an example, the major product from heating the triene **2-1** (Scheme 1.2), likely obtained from dehydration of cholic acid (**1-2**), consists

of the hydrocarbon chrysene (**2-2**), the structure of which had by then been independently established. This result provided early evidence for the presence in steroids of a staggered array of four fused rings.



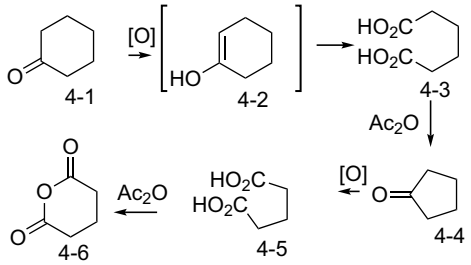
Scheme 1.2

More direct evidence for the gross structure of carbon skeleton of steroids came from the isolation of the hydrocarbon **3-1** (Scheme 1.3) from a mixture of hydrocarbons obtained from heating cholesterol itself with selenium. This product **3-1**, known as the Diels hydrocarbon, proved to be identical with a sample of the compound synthesized from starting materials of known structure by an unambiguous reaction scheme.



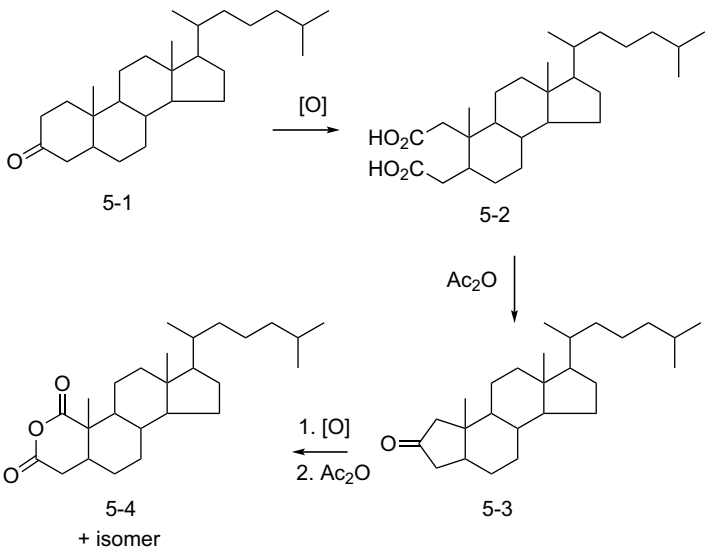
Scheme 1.3

The size of each of the rings present in steroids was established by serial oxidation reactions starting with what would later be dubbed ring A. The empirical, so-called Blanc, rule holds that oxidation of a cyclohexanone (**4-1**) (Scheme 1.4) proceeds to afford a dicarboxylic acid (**4-3**), likely through the enol form, **4-2**. Heating the diacid **4-3** with acetic anhydride proceeds to cyclopentanone **4-4** with loss of one carboxyl carbon. Repeated oxidation of that intermediate again results in a dicarboxylic acid, in this case adipic acid (**4-5**). Exposure to hot acetic anhydride leads to anhydride **4-6**. The strained nature of the cyclobutanone that would result from cyclization as in **4-3** is disfavored over the formation of the anhydride. Reaction thus proceeds to succinic anhydride (**4-6**). In the absence of instruments, anhydrides can be distinguished from ketones by the fact that the former will lead to a dicarboxylic acid on basic hydrolysis. The neutral ketone can be recovered unchanged under the same conditions.



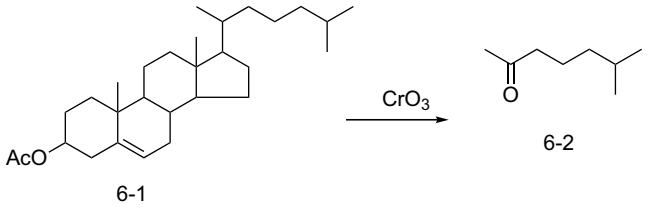
Scheme 1.4

In the case of the reduced derivative of cholesterol, **5-1** (Scheme 1.5), the initial oxidation goes to the highly substituted adipic acid **5-2**. The observation that this leads to a cyclopentanone (**5-3**) can be inferred to indicate that the ring at the start of the scheme was six-membered, Further oxidation of the cyclopentanone again leads to a dicarboxylic acid. On treatment with acetic anhydride, that intermediate leads to a cyclic anhydride. This leads to the inference that the precursor **5-3** was a cyclopentanone.



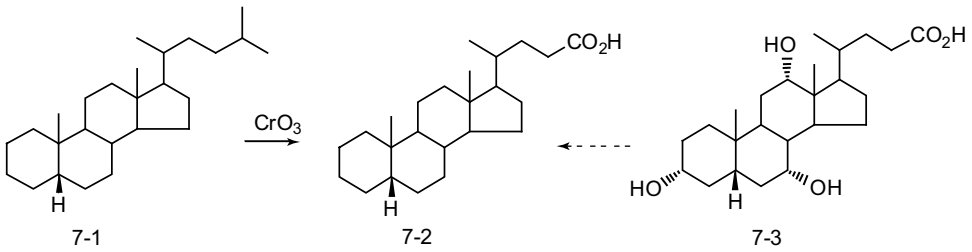
Scheme 1.5

Serendipity played a role in establishing the structure of the side chain in cholesterol. Some investigators had noted that a sweet, perfume-like odor accompanied the vigorous oxidation of cholesterol acetate (**6-1**) (Scheme 1.6). The odorous substance was finally isolated from a very large-scale (500 g) oxidation run and converted to its semicarbazone. This proved to be identical with the same derivative from 6,6-dimethylhexan-2-one (**6-2**).



Scheme 1.6

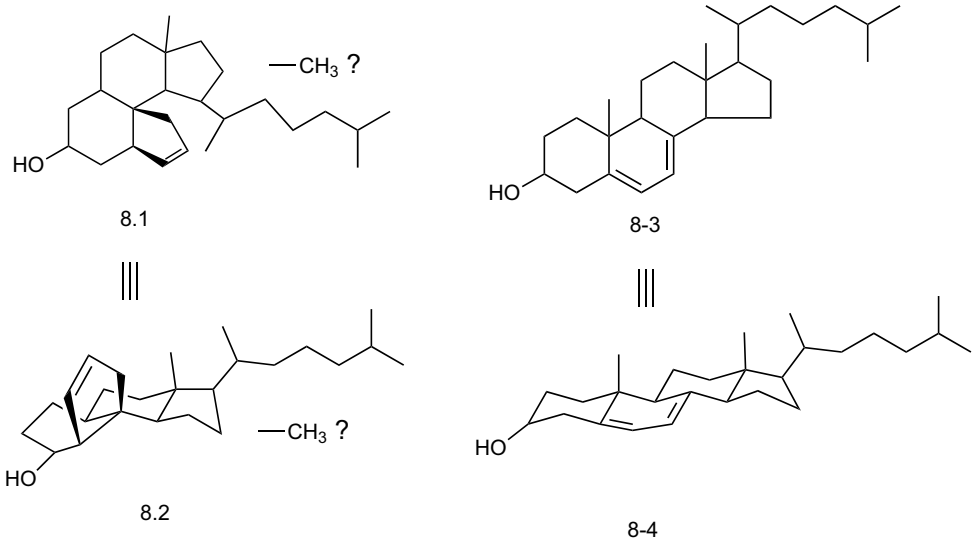
Many chemists, principally those in Adolf Windaus's group at the University of Göttingen, worked on unraveling the intricacies of the structures in the cholesterol series. Another group, led by Heinrich Wieland at the University of Munich, studied the structures of the bile acids. Suspecting that these two natural products shared a common carbon nucleus, they each sought to relate the two by preparing a common derivative. In brief, they established that the cholanic acid **7-2** from exhaustive reduction of cholic acid (**7-3**) was identical to the product from oxidation of coprostane (**7-1**) (Scheme 1.7). The latter was obtained by exhaustive reduction of cholesterol. The common derivative, it should be noted, incorporates the less common cis A–B ring fusion. The reactions that lead to the common intermediate are unlikely to alter the configuration of chiral centers present in the natural products. The identity of the derivatives obtained from each starting material thus established that the product related to cholesterol and those derived from bile acids shared the same nucleus and overall stereochemistry.



Scheme 1.7

The two groups and also other investigators who worked on the problem felt that enough data had been accumulated to propose a structure in 1928. Most of the carbon atoms had been accounted for and the results, they deemed, supported

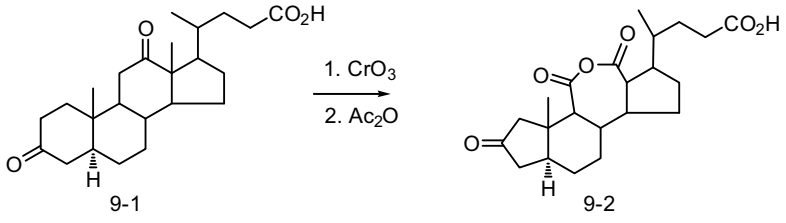
8-1 (Scheme 1.8) as the structure of cholesterol. Some ambiguity existed as to the attachment of one of the two methyl groups in cholesterol. Some, it is said, referred to this fragment as the 'floating methyl'. Depiction of the proposed structure in three dimensions (**8-2**), instead of the common two-dimensional notation (**8-1**), makes it clear that the proposed structure would have consisted of a relatively thick, congested molecule.



Scheme 1.8

The use of X-ray crystallography for solving the structures of organic compounds was still in its infancy in the late 1920s. The use of that tool was hindered by the need to perform an enormous amount of data reduction; the mechanical calculating machines employed for that were then just coming into wider use. Atom-by-atom mapping of a complex structure such as a steroid was, at that time, still beyond the then state-of-the-art. Resolution of an X-ray crystallographic study of ergosterol (**8-3**) was however sufficient to indicate that this steroid consisted of a long, flat molecule (**8-4**) rather than a thick, congested entity such as **8-2**. Re-examination of all the data from degradation studies revealed that an exception to the Blanc rule caused assignment of the wrong structure (**8-2**) to ergosterol. This second look also led to the correct formulation of the steroid nucleus as depicted by **8-3**.

One set of degradation studies on cholanic acid (**9-1**) (Scheme 1.9) led to scission of what are now known as ring A and ring C. One pair of the four new carboxylic acids led to a cyclopentanone and the other to an anhydride. On the basis of this, it was then inferred that ring A was six-membered whereas ring C comprised a cyclopentane (see also **8-1**). It was later recognized that carboxylic acids attached directly to rings as in **9-2** cannot form a cyclopentanone. This exception was later attributed to steric strain in the hypothetical product.



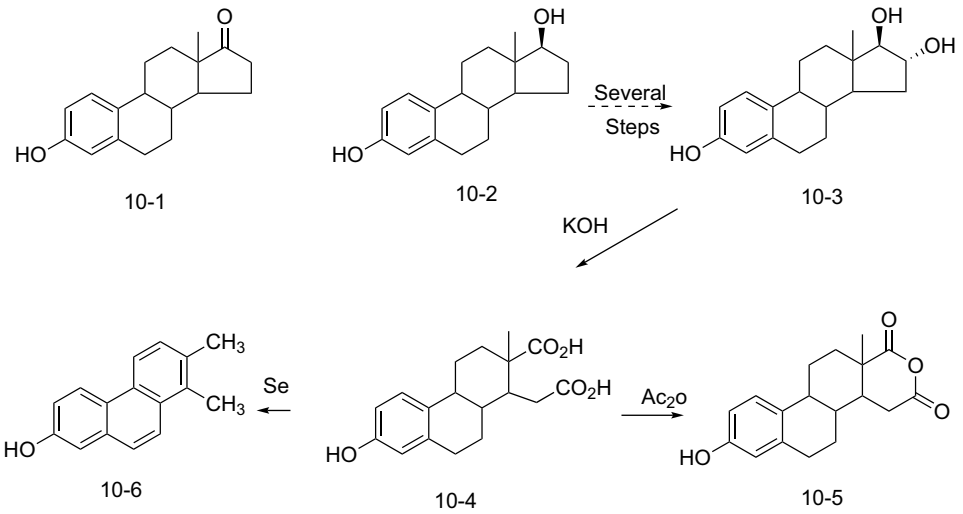
Scheme 1.9

1.1.2 The Sex Steroids

By the early 1930s, it was clear that the reproductive function in mammals was directed by a group of potent discrete chemical substances. These compounds, dubbed sex hormones, consist of three distinct classes, the estrogens, the progestins and the androgens; these substances differ from each other in both biological activity and structure. The very small amounts of these compounds found in tissues posed a major challenge to investigations aimed at defining their chemical structure.

1.1.2.1 Estrogens

The first of the three classes of hormones that regulate reproductive function in both females and males of the species, the estrogens, progestins and androgens, were isolated in 1929. This marked contrast to the dates for the first isolation of bile acids and of cholesterol is due in no small part to the minute amounts of those hormones that were available for structural studies. Isolation of those substances from mammalian sources, such as mare's urine, was guided by bioassays, increasing potency of a sample signaling higher purity. This work culminated in the isolation in 1929 of a weakly acidic compound, estrone (**10-1**) (Scheme 1.10). The acidity indicated the presence in the molecule of a phenol and hence an aromatic ring. Various chemical tests pointed to the steroid nature of estrone and also several closely related compounds. The principal accompanying compound, estradiol (**10-2**), and estrone comprise the primary estrogens and are freely intraconvertible both *in vivo* and *in vitro*. The former occurs as two isomers that differ at position 17; one isomer features the alcohol at position 17 above and the other below the plane of the molecule. Estradiol-β that carries the hydroxyl above the plane is the more potent than its 17α-hydroxy epimer. The closely related compound estriol (**10-3**) often accompanies estradiol *in vivo*. The compound can also be prepared from estradiol by a straightforward sequence of reactions not likely to change the carbon skeleton.



Scheme 1.10

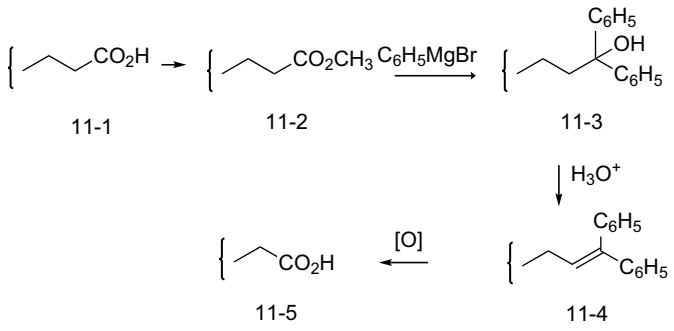
Fusing estriol (**10-3**) with potassium hydroxide cleaves the bond between the two hydroxyl-bearing carbon atoms in the five-membered ring. Those atoms are oxidized to carboxylic acids under reaction conditions to afford **10-4**, dubbed marrianolic acid (this derivative, named after its discoverer, interestingly retains significant estrogenic activity). On treatment with acetic anhydride, this gives an anhydride, **10-5**; the Blanc rule indicates that the precursor ring is five-membered. Heating the diacid **10-4** with selenium causes the carboxyl groups to leave as carbon dioxide; under reaction conditions, the six-membered rings lose hydrogen, leaving behind the phenanthrol **10-6**. This molecule proved to be identical with a sample of **10-6** synthesized by an unambiguous route. By 1933, the detailed structure of estradiol was firmly established. Several total syntheses have been published since then (see Chapter 3).

1.1.2.2 Progestins

Biological studies carried out at roughly the same time identified another hormone, this one a neutral substance whose concentration in body fluids fluctuated in synchrony with the menstrual cycle; the hormone was further found in blood at high levels during pregnancy. This compound, progesterone, was not obtained in crystalline form until 1934 as it was often accompanied in extracts by other closely related compounds such as pregnenolone. Instrumental tools in this case played a role in deducing the chemical structure of progesterone. Although X-ray crystallography did not as yet enable atom-by-atom mapping, it did provide evidence that progesterone possessed a four-ring sterol-like structure. The ultraviolet spectrum showed an absorption spectrum typical for a conjugated unsaturated ketone. Much of the structural work was carried out using about 2 g of pregnanediol (**12-4**) obtained by extracting in excess of 1000 L of pregnancy urine.

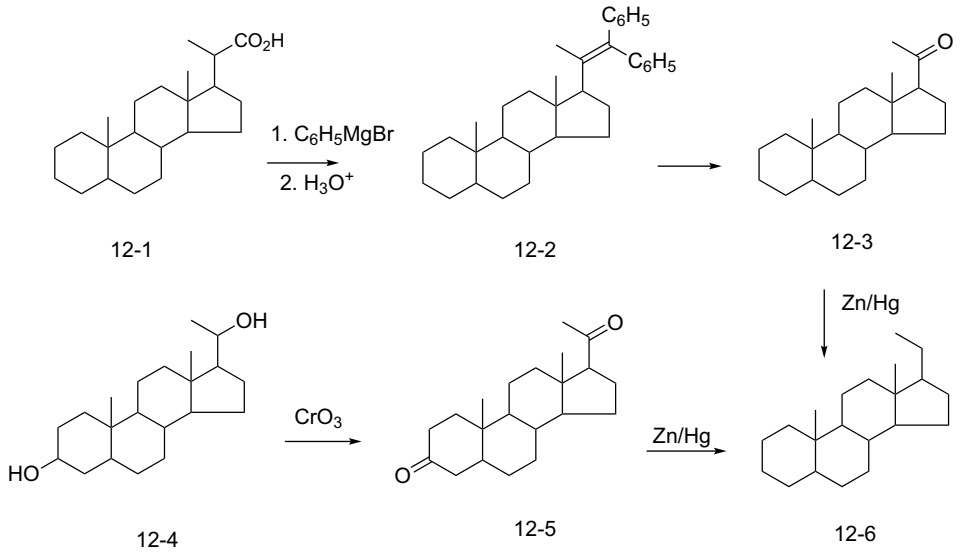
A fairly straightforward scheme, called the Barbier–Wieland degradation, was at that time used for determining the number of carbon atoms in an aliphatic acid fragment. This involves first converting the carboxylic acid **11-1** to the corresponding ester **11-2** (Scheme 1.11). Reaction of the ester with phenylmagnesium bromide leads to carbinol **11-3**. Treatment with acid leads to dehydration and formation of the olefin **11-4**. Oxidation by one of several methods cleaves

the double bond with formation of a new carboxylic acid (**11-5**), shorter by one carbon atom than the starting acid **11-1**. The sequence would be repeated until degradation met a branch in the chain and afforded a ketone instead of an acid.



Scheme 1.11

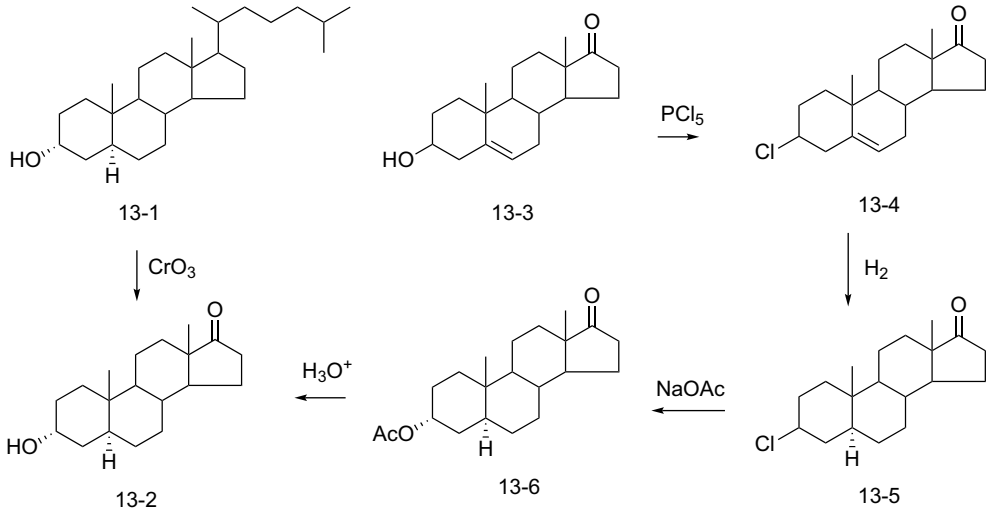
The sequence that established the structure of the pregnan nucleus starts with the chain length probing sequence depicted in Scheme 1.12. The carboxylic acid derivative **12-1**, which can, in concept, be prepared from cholanic acid by initial exhaustive reduction to remove the hydroxyl groups followed by two rounds of sequence depicted in Scheme 1.11. The carbonyl group at position 20 (**12-3**) was reduced by means of amalgamated zinc to give the pregnane nucleus **12-6**. In a convergent sequence, pregnanediol (**12-4**) was oxidized to pregnane-2,17-dione (**12-5**) with chromium trioxide. The carbonyl groups at C_3 and C_{20} were then reduced with amalgamated zinc to give a sample of **12-6** identical in all respects with that obtained from cholanic acid (**12-1**). The assignment of one of the oxidized carbon atoms at position 3 was based on analogy with cholesterol and the other at position 20 relied on intra-conversion with the C_{17} androgens.



Scheme 1.12

1.1.2.3 C_{19} Androgens

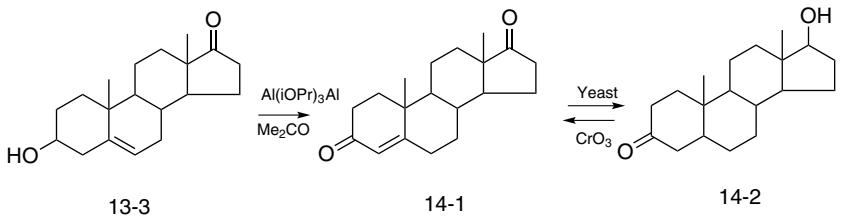
Androgens, the male sex hormones, proved far more elusive that either the estrogens and progestins since they occur at much lower concentrations in biological fluids. The bioassay used to track the isolation in this case comprised the 'capon unit'. This was the amount of extract that produced a 20% increase in the surface of a rooster's comb. The 15 mg of pure crystalline testosterone isolated in 1931 came from about 15 000 l of urine. The structural investigations of this series relied on the then newly discovered side chain oxidations of cholestanol (**13-1**) (Scheme 1.13). This method in essence comprised fairly drastic oxidation of reduced cholesterols of known stereochemistry at the A–B junction to afford in fairly low yield products in which the side chain at C_{17} had been consumed to leave behind a carbonyl group. One of these products proved to be identical with androsterone (**13-2**). That compound had in turn been obtained from a sequence of reactions starting from dehydroepiandrosterone (**13-3**) that had been isolated from male urine.



Scheme 1.13

Treatment of dehydroepiandrosterone (**13-3**) with phosphorus pentachloride replaces the hydroxyl at position 3 with retention of configuration (**13-4**). It had been established prior to this work that catalytic reduction of unsaturation in steroids proceeds almost invariably from the bottom side to afford reaction products as their 5α epimers, as for example **13-5** [this is also the case for cholestanol (**13-1**), obtained from hydrogenation of cholesterol]. Sodium acetate then displaces chlorine in **13-5** to afford the acetoxy derivative **13-6** with inverted configuration at C_3. Mild hydrolysis of the acetoxy group affords the corresponding alcohol **13-2** that is identical with that of a sample produced by oxidation of cholestanol (**13-1**).

The stereochemical argument can be closed with the observation that oxidation of dehydroepiandrosterone by the Oppenauer reaction (aluminum isopropoxide in the presence of a ketone) yields the oxidation product androst-4-ene-3,17-dione (**14-1**) (Scheme 1.14). The same diketone is formed from oxidation of testosterone (**14-2**). Going in the reverse direction, androst-4-ene-3,17-dione can be converted to testosterone by treatment with fermenting yeast.



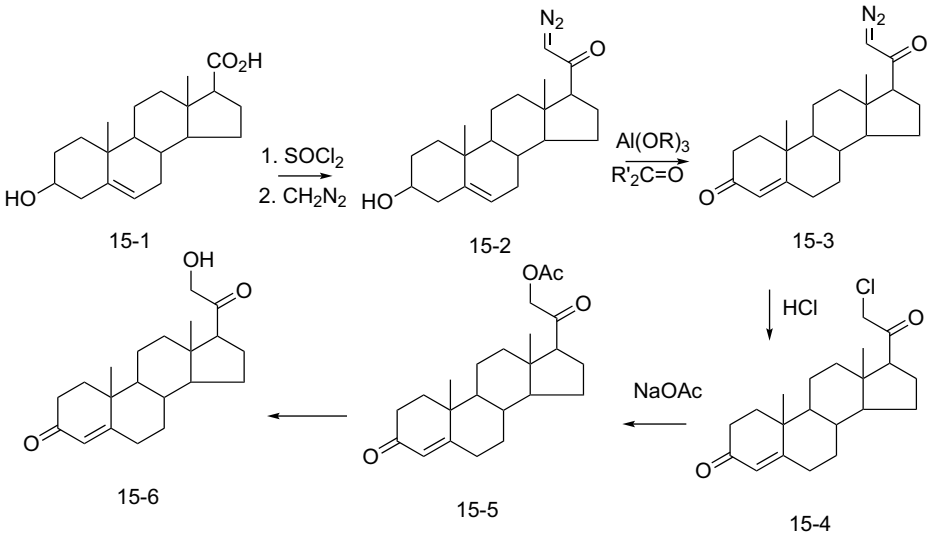
Scheme 1.14

1.1.3 Corticosteroids

1.1.3.1 Glucocorticoids

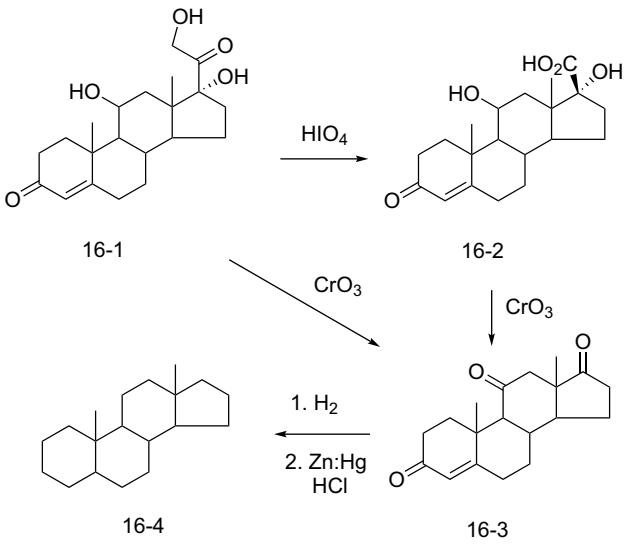
The realization in the 1930s that substances secreted by the so-called endocrine glands play a major part in various life processes led to a major effort to determine the chemical structure of those secretions, as was the case for the set of sex hormones described in the preceding sections. Investigation of the adrenal glands, located atop the kidneys, revealed that animals whose adrenals had been removed died within a few days. Administration of adrenal extracts increased their lifespan. Identification of the active ingredient was complicated by the 30 or so compounds, now known to be steroids, secreted by the adrenal outer layer, known as the cortex. The principal products, hydrocortisone (formerly cortisol) and aldosterone, account for most of the activity of the extracts. Hydrocortisone regulates carbohydrate, fat and protein metabolism whereas aldosterone acts on electrolyte balance via the kidneys. These hormones, like the sex hormones, occur at low concentrations: about 450 kg of beef adrenals yielded only about 300 mg of hydrocortisone. One of the compounds accompanying hydrocortisone proved to be identical with 20-hydroxyprogesterone (**15-6**) (Scheme 1.15) that had been prepared from the known carboxylic acid **15-1** in studies on the structure of progesterone.

The sequence for preparing the hydroxyketone started by conversion of the acid **15-1** to its chloride with thionyl chloride. Reaction of that acid halide with diazomethane gives the diazoketone **15-2**. The hydroxyl group at C_3 is then oxidized to the corresponding ketone by means of an Oppenauer reaction. Treatment of the product **15-3** with gaseous hydrogen chloride replaces nitrogen in that intermediate by chlorine. Displacement of chlorine by acetate then leads to the 21-acetate **15-5**. Saponification of the ester completes the sequence.



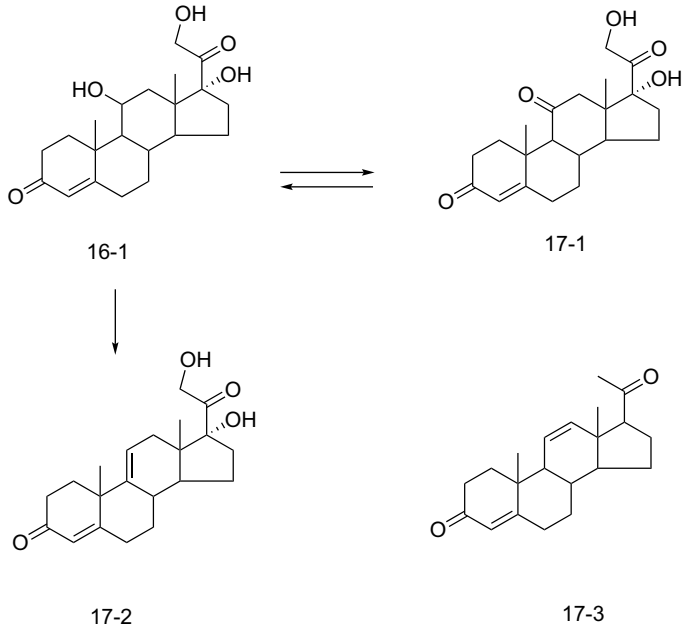
Scheme 1.15

Elemental analysis of hydrocortisone indicated the presence in the molecule of two additional oxygen functions compared with progesterone. Treatment of hydrocortisone **16-1** with periodic acid cleaves the side chain at position 20 to afford hydroxy acid **16-2**, a reaction that indicates the presence of a 1,3-dihydroketone (Scheme 1.16); this is also evidence that one of those additional oxygen atoms is at C_{17}. Further oxidation of **16-2** with chromium trioxide then causes the hydroxy acid to lose carbon dioxide to leave behind a ketone (**16-3**). The hydroxyl group at position 11, which is virtually inert to other reactions, is oxidized to a ketone under these reaction conditions; that product, known as adrenosterone, can also be formed by direct oxidation of **16-1** with chromium trioxide. Reduction of the double bond in **16-3** followed by treatment of the product with zinc amalgam in acid gives the hydrocarbon androstane (**16-4**). This proved to be identical with a sample prepared from dehydroepiandrosterone. It might be noted in passing that androstane emits a very strong odor, similar to that of poorly maintained urinals.



Scheme 1.16

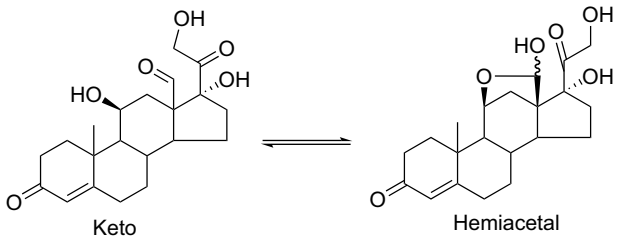
Assignment of the remaining hydroxyl group to position 11 rests in large part on the lack of reactivity of the hydroxyl group in hydrocortisone (**16-1**), or for that matter the 11-ketone (**17-1**) in cortisone (Scheme 1.17). Molecular models show that the 18- and 19-methyl groups effectively shield those positions from attack from the β side. The ketone at position 11 will form normal derivatives only under the most forcing conditions. Reduction of the suitably protected form of the ketone readily gives the corresponding hydroxyl (**16-1**). This is assigned as β on the basis of the approach of hydride or hydrogen from the more accessible α face of the steroid. Additional support for placing oxygen at position 11 comes from the finding that dehydration results in the formation of a double bond at position 9(11) (**17-2**), whereas the 12-hydroxyl in 12-hydroprogesterone gives an 11,12-olefin (**17-3**).



Scheme 1.17

1.1.3.2 Aldosterone

The work that led to the identification of cortisone in extracts of the adrenal cortex led, as noted above, to the isolation of a host of closely related steroids, There remained, however, a fraction that defied crystallization. This material, the amorphous fraction, exhibited fairly respectable activity in regulation blood volume and serum electrolytes. This steroid, aldosterone, can exist in either the keto or lactal form (Scheme 1.18). Degradation and synthesis studies are beyond the scope of this book.



18

Scheme 1.18