Tetrahedron, 1960, Vol. 8, pp. 360 to 365 Pergamon Press Ltd. Printed in Northern Ireland

STEROID NOMENCLATURE

L. F. FIESER and M. FIESER
Dept. of Chemistry, Harvard University, Cambridge, Mass.

(*Received* 21 *September* 1959)

Abstract—Proposals for revision of nomenclature are as follows:

1. In a sapogenin or alkaloid in which the terminal carbon is linked to a hetero atom, this carbon atom, in accordance with the Fischer convention, acquires the number 27.

2. If, in a Fischer projection of an open side chain, methyl or hydroxyl at a given center lies to the left it is β-oriented; if it lies to the right it is α-oriented.

3. In a sapogenin or alkaloid the true stereochemistry at any center in the side chain is described as the configuration at this center in the alcohol or amine derived by opening the side chain and making a Fischer projection. If the orientation so defined as α is down (or rear) it is described as α-oriented. The symbol α_F is used to indicate that the configuration by Fischer projection is α but that the orientation is up (or front).

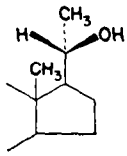
IN our book *Steroids*[1] we employed, without discussion, a system of nomenclature which departs in some respects from current usage, as set forth in the 1957 IUPAC Rules for Nomenclature of Steroids.[2] It is the purpose of this paper to discuss the various problems involved in the hope of stimulating general consideration of the matter and eventual agreement on one system or the other, or on a compromise. Actually the 1957 IUPAC Rules are essentially those formulated at a conference held in 1950 at the Ciba Foundation in London,[3] in which we participated. In view of the rapid advances in the steroid field over the past decade, this would seem an appropriate time to reconsider the proposals of 1950.

One rule recorded without comment and without any evident rationale is that in a sapogenin or related alkaloid the terminal carbon carrying the oxygen or nitrogen function be assigned the number 26. This assignment, however, violates the Fischer convention[4] that in determination of the basic carbon skeleton a terminal carbon carrying oxygen (or nitrogen) takes precedence over a methyl group. This convention was not defined as such by Fischer but merely implied by his usage, and although it is familiar to sugar chemists[5] it may not be generally known. Thus the convention was first called to our attention in very helpful discussions with T. Reichstein and with A. Georg. Since a convention does exist, we see no reason for not assigning the number 27 to the substituted terminal carbon.[6]

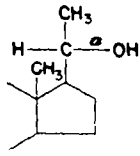
With steroids of the natural (or *d*-) series having an open side chain, assignment of configurations requires merely an extension of existing conventions. One,[7] applicable to pregnane-20-ols, calls for orienting the model with methyl (forming part of

[1] L. F. Fieser and M. Fieser, *Steroids.* Reinhold, New York (1959).
[2] International Union of Pure and Applied Chemistry, *Nomenclature of Organic Chemistry*, 1957. Butterworths, London (1958).
[3] See *Chem. & Ind.* SN 1, June 23 (1951); *Helv. Chim. Acta* **34**, 1680 (1951); *Bull. Soc. Chim. Fr.* No. 3/4 viii-ix (1951).
[4] E. Fischer, *Ber. Dtsch. Chem. Ges.* **24**, 1836, 2683 (1891).
[5] W. Pigman, *The Carbohydrates* p. 23. Academic Press, New York (1957).
[6] In application of the Fischer convention to sugars, the oxygenated carbon is assigned the number 1. With steroids, since it is impossible to use the lowest number, choice of the highest number (C_{27}) seems in keeping with the spirit of the convention.
[7] L. F. Fieser and M. Fieser, *Experientia* **4**, 285 (1948).

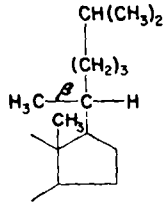
the longest chain) to the rear and hydrogen and hydroxyl to the front, as in I.



I

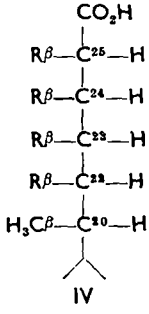


II

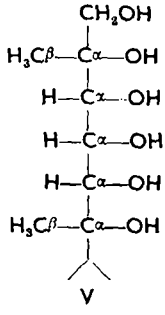


III

The Fischer projection is then made as in II, and the hydroxyl lying to the right is defined as α-oriented. A second convention, proposed by Pl. A. Plattner in an appendix to the Ciba conference report,[3] is that cholesterol, known to have the configuration III, be defined as having a 20β-methyl group. Obvious extensions for definition of the orientations of alkyl and hydroxyl groups at positions in the side chain are shown in IV and V. An alkyl group that is on the left, like the 20β-methyl



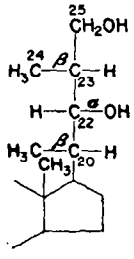
IV



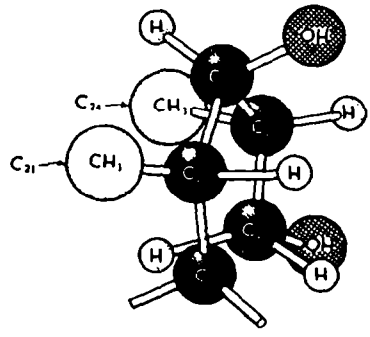
V

group, is described as β-oriented, one on the right is α-oriented. Hydroxyl groups lying to the right are α, to the left, β. Thus V has a 20β-methyl and a 20α-hydroxyl group. The diol derived from cholesterol by replacement of the 20α-hydrogen by hydroxyl without inversion is 20α-hydroxycholesterol; its 20-epimer is best described as 20β-hydroxy-20-isocholesterol.

In 20α-hydroxycholesterol the methyl at C_{20} is β-oriented and the hydroxyl is α-oriented, and it does not seem proper to describe carbon atom 20 as having either the α- or the β-configuration. The configuration is adequately defined by the orientation of groups, considered as substituents. In diol VI the configuration at C_{20} is indicated by description of the substance as of the 20β-methyl orientation, and the number (21)



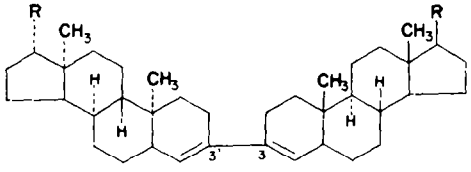
VI



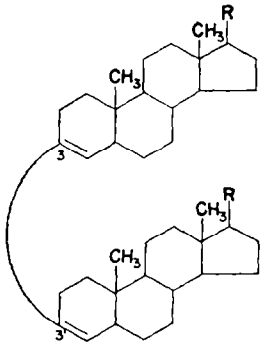
of the carbon of the methyl group need not appear in the name. The methyl group attached to C_{23} can similarly be regarded as a substituent, comparable to the substituent hydroxyl at C_{22}. Thus VI is a 20β, 24β-dimethyl-22α, 25-diol.

The proper way to view a model is shown in the drawing VII of the diol VI. The carbon atoms forming the basic chain are arranged in a plane so that the chain curves around to the rear. One views C_{20} from the front as usual, but to view other carbons one must float along the chain and look to the right and left as one passes over each station. When one reaches C_{23} one is upside down, but appropriately sees methyl to the left (β) and hydrogen to the right (α).

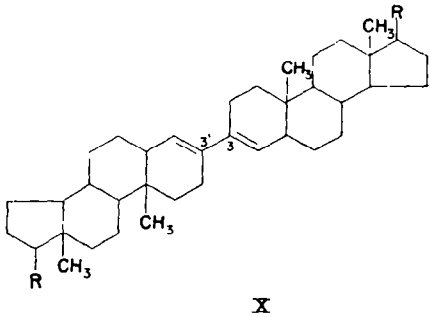
The α and β designations defined for an open side do not carry the implication of down (or rear) and up (or front) that is usually associated with similar designations at nuclear positions but they have the same status of defining the actual relative configurations. The configuration at a given center, once determined, is fixed and, in our view, is not altered by making rotations about single bonds, by turning a model upside down, or by the closing of rings in the side chain. This is the key point on which our view differs from that of the IUPAC. The problem is particularly pressing in the sapogenin and alkaloid series, but it may come up also with respect to positions in the ring system. In the bisteroid VIII the groups at positions 10, 8, 13, and 17 are β, and so, we think, are those at positions 10′, 18′, 13′, and 17′, even though they happen to be to the rear. J. Mathieu has pointed out in a letter that all the β-groups can be shown as oriented to the front by formulating the bisteroid as in IX, and Georg has noted that Chopin[8] and others achieve the same end with formula X. However VIII might be the formula of choice for certain discussions because it shows that the hydrocarbon has the potentiality for Diels–Alder reaction as a cisoid diene, and it



VIII



IX

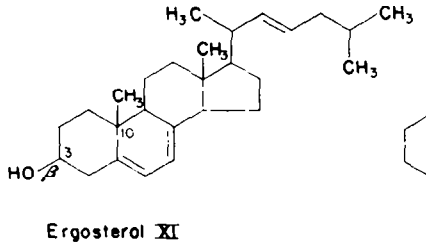


X

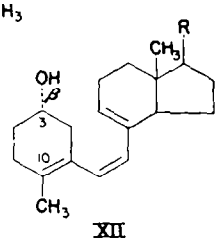
[8] J. Chopin, *Bull. Soc. Chim. Fr.* 258 (1956).

seems to us perfectly valid; one must simply recognize that in this case α and β denote configuration alone and not orientation in space.

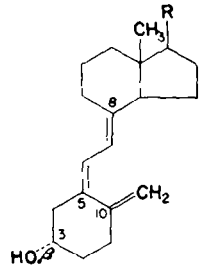
In the case of previtamin D_2 (XII) and vitamin D_2 (XIII) there is no alternative way of writing correct formulas that show the 3-hydroxyl group oriented to the front. One colleague, commenting on a first draft of this paper, suggested that *for nomenclature purposes* formulas (XII) and (XIII) be rotated about appropriate single bonds to give arrangements in which the carbon atoms are placed in the same way as in the conventional formulas of the steroids and in which the hydroxyl group is to the front. But vitamin D_2 would then be represented as having a cisoid diene system extending from C_8 to C_5 whereas X-ray and chemical evidence indicates that it is transoid. If it is indeed necessary to choose between correct chemistry and convenience of nomenclature, the former surely should prevail. To describe these isomers of ergosterol (XI) as having the 3α-configuration seems to us incorrect. The configuration at C_3 is



Ergosterol XI



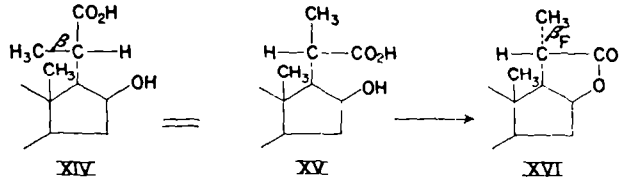
XII



XIII

the same as in ergosterol, or β, and the compounds are adequately described as 3β-alcohols with a rear-oriented hydroxyl group. A β-label and a dotted bond show both the configuration and the orientation. Since bonds at ring positions ordinarily are not labeled, the appearance of the β on a bond which is dotted and hence to the rear shows at once that the case is a special one.

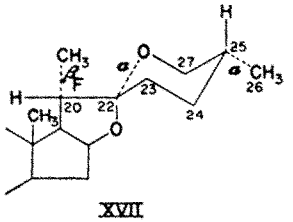
The same treatment seems to us a practical and accurate way of dealing with compounds having cyclized side chains, for example, the lactone XVI derived from the 16β-hydroxybisnor acid XIV, or its equivalent XV. The acid has a 20β-methyl group like cholesterol, and the configuration is retained on cyclization. To follow IUPAC and say that cyclization of a 20β-methyl hydroxy acid gives a 20α-methyl lactone does not seem rational. In this system a Greek letter indicates only that a group is up or down, (or front or rear) which can be seen in any case from the nature of the bond, and it has no reference to the actual configuration relative to other centers and other compounds. This abandonment of configurational indication seems to us a serious shortcoming of the system, and we see distinct advantages in an alternative system from which it is clear at a glance, for example, whether or not a given



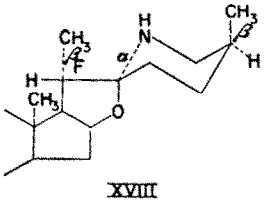
XIV XV XVI

alkaloid corresponds in configuration at a specific site to a sterol or a sapogenin. With a complicated molecule it may be necessary to build a model and open it up for projection in order to establish the configurations, but this operation need be performed only once. However, we are aware, particularly from comments by R. S. Cahn and by J. Fried, that many chemists have become too familiar with the simple "α-below and β-above the plane" nomenclature to be willing to give it up.

To resolve the impasse, we now suggest a compromise utilizing the symbols α_F and β_F suggested by J. Fried to indicate configurations as defined by the Fischer convention. We propose that, where the IUPAC and that the Fischer assignments are identical, the configuration (and orientation) be indicated by α or β, but that where a difference exists the designation be α_F or β_F. Thus cyclization of the 16β-hydroxybisnor acid XIV of 20β-configuration gives a lactone with an α-oriented $20\beta_F$-methyl group. Projection

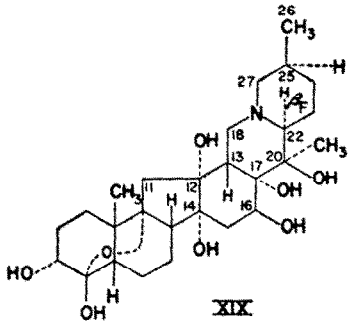


XVII



XVIII

of the opened-up forms of smilagenin (XVII) and of tomatidine (XVIII) shows that the bond extending from the hetero atom to C_{22} is α (rear) and that the 25-methyl group is α (down) in (XVII) and β (up) in (XVIII). Both the sapogenin and the alkaloid, however, have a rear-oriented $20\beta_F$-methyl group. In the case of cevine (XIX) and related alkaloids[9] the IUPAC system defines configurations identical with those derived by Fischer convention except at C_{22}; here the hydrogen is down but

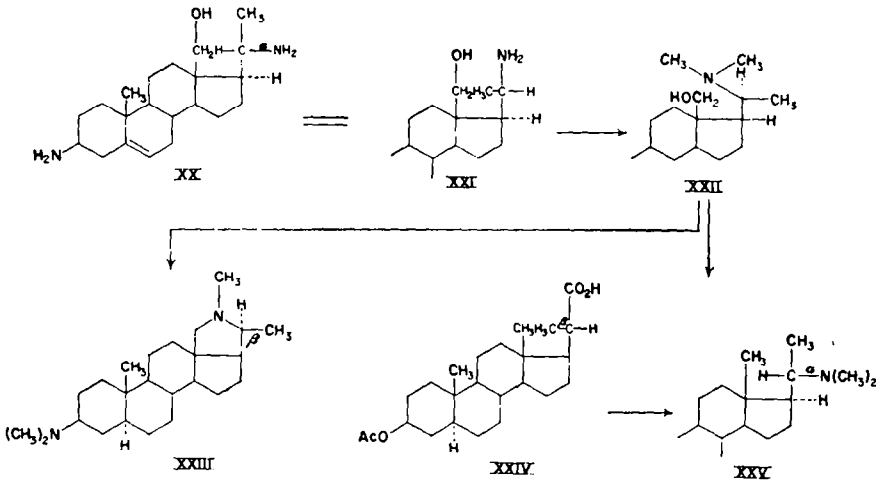


XIX

the configuration is $22\beta_F$-H. The alkaloid is unusual in that the configuration at C_{20} is the opposite of that of cholesterol and of other steroid alkaloids and of the sapogenins.

Holarrhimine (XX) has been correlated with conessine by cyclization of tetramethyldihydroholarrhimine (XXII) to a product identical with dihydroconessine (XXIII) and the configuration at C_{20} of both alkaloids was established by conversion of the 19-hydroxymethyl group of XXII to methyl (XXV) and synthesis of the latter compound from 3β-acetoxybisnorallocholanic acid (XXIV).[10] By the well established

[9] S. M. Kupchan, W. S. Johnson and S. Rajagopalan, *Tetrahedron* 7, 47 (1959).
[10] V. Černý and F. Šorm, *Coll. Czech. Chem. Comm.* 20, 1473 (1955); L. Lábler, V. Černý and F. Šorm, *Ibid.* 20, 1484 (1955); V. Černý, L. Lábler and F. Šorm, *Ibid.* 22, 76 (1957).



convention for 20-alcohols, holarrhimine (XX) is described as having a 20α-oriented amino group, the 20-hydrogen is β, and the 20-methyl group has no assignment. To meet the case at hand, we suggest that the formula be arranged as in (XXI) with the amino group at the top or rear, corresponding to the carboxyl group of the acid (XXIV) or C_{22} of the original cholesterol side chain. The 20-methyl group then corresponds to that of cholesterol, and in dihydroconessine (XXIII) retains the β-orientation in both the Fischer and the IUPAC sense.

In arriving at the present proposals, which supplement those of our book,[11] we have been guided, and in some cases corrected, by very helpful discussions with the commentators already mentioned as well as with R. E. Beyler, R. K. Callow, J. W. Cornforth, M. E. Wall, and O. Wintersteiner.

A further recommedation is with respect to chemical names in English. In the interest of preservation of some beauty in the language, we protest against current usage of such names as cholest-5-ene, cholest-5-en-3β-ol, pregna-5,16-dien-20-one, and urge adoption of the rule that a chemical name should be divided only when each part is a correctly pronounceable word in its own right. Cholest and pregna are not words. En and dien are not words and should be pronounced én and dién. On the other hand, ene diene, polyene are acceptable as properly pronounced words, as are ol, diol, triol. Suitable names for the compounds mentioned are: 5-cholestene or Δ⁵-cholestene, 5-cholestene-3β-ol, or Δ⁵-cholestene-3β-ol, 5,16-pregnadiene-20-one or Δ⁵,¹⁶-pregnadiene-20-one. We prefer use of the Δ, particularly since this prevents confusion of numbers; thus 3β-chloro-Δ⁵-cholestene seems better than 3β-chloro-5-cholestene. We are opposed also to the prevalent cholestan-3β-yl acetate. Cholestan is not a word and we see little to support the contention that yl is a word, certainly not a good word; diyl and triyl do not exist. Cholestanyl acetate is assuredly correct, but for a specific name we prefer cholestane-3β-ol acetate, to be read in the sense "cholestane-3β-ol, acetate of," comparable to estrone acetate, cortisol 21-acetate, cholestane-3β, 5α, 6β-triol 3,6-diacetate. An alternative method is that currently used by sugar chemists: O-acetyl-Δ⁵-cholestene-3β-ol. Our only contention is that the yl-system should be abandoned.

[11] The compromise scheme is stated on p. 889 of *Steroids*; it was added, in proof, too late for application in earlier parts of the book.