COMPUTER CONSTRUCTION OF STANDARD STEROID MODELS

P.SEDMERA,* A.VÍTEK and Z.SAMEK

Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, Prague 6

Received January 4th, 1972

Numerical models of 5α,14α-androstane, 5α,14β-androstane, 5β,14α-androstane, and 5β,14β-androstane were constructed by computer on the basis of geometric considerations using only the weighted averaged values of the natural coordinates. By comparison of these and Dreiding models with some X-ray determined structures it follows that the presented models can serve as a better source of geometric quantities for semiempirical physicochemical calculations than Dreiding models.

Large collections of variously substituted steroid compounds exist in PMR spectroscopy for testing of semiempirical rationalizations on long-range shielding effects and coupling constants. A knowledge of the geometric arrangement of atoms in the molecule is required for their theoretical prediction and also for the estimation of other physical properties of molecules (dipole moments, non-bonding interactions, ORD, *etc.*). Generally speaking, the most direct source of the information about molecular geometry are the methods of the X-ray, electron, or neutron diffraction. However, such data are available for a small number of compounds only, so in the remaining cases it is necessary to use the models. Two kinds of models are used in the present practice: the mechanical models built from solid rods, spheres, *etc.*, and the mathematical ones, represented by a set of position vectors of all the atoms in the molecule. The direct measurement of some quantities, *e.g.* interatomic distances or angles formed by two vectors having a common origin, is possible with the mechanical models constructed on a precise scale (Fieser or Dreiding models)[1,2]. Simple devices were proposed for this purpose[3-5]. However, this procedure is rather awkward, especially with large number of studied quantities and has a limited reproducibility. The advantage of the mathematical models is that any geometric quantities can be obtained by calculation with an accuracy corresponding to the accuracy of the input coordinates. The data can be stored in the computer memory and then manipulated in various manners. A visual check is obtained by special programs[6,7] whose output unit can be the plotter[6,7], or CRT display[8], possibly both these devices[9]. The mathematical models can be divided into two groups. The first type, so-called average structures, is generated from the natural coordinates obtained by averaging of the set of the X-ray determined structures. A number of programs for an easy construction of the models of acyclic compounds is available[10-13]. The difficulties arise with cyclic systems, where (with the exception of idealised molecules possessing high symmetry) the "ring closure" is not assured[12], since the mutual dependence of the natural coordinates (bond lengths, valency and torsional angles) is not taken into account. The second group of the models is represented by so-called equilibrium structures, obtained by minimalization of the conformational energy expressed as a function of the geometric parameters of the molecule[14,15], based on an as-

* Present address: Institute of Microbiology, Czechoslovak Academy of Sciences, Prague 4.

sumed character for the intramolecular interactions. The parameters of the functions describing these interactions are adjusted in such a way as to obtain an agreement with the X-ray diffraction experiments. Full relaxation calculations of this type solve the problem of ring closure[16] but of course, for equilibrium structures only.

In this paper, we have concentrated on the question of the geometry of the steroid skeleton of various types. The process of finding of the equilibrium structure is very time consuming even on the fast computers and the obtained information is often not proportional to the expended effort. Therefore we used the first type of model respecting the interdependence of the natural coordinates. The program,* whose main part is an iterative algorithm for the adjustment of the dependent natural coordinates[17], can be used for the construction of a mathematical model of any steroidal molecule; the results for the four most frequent types, 5α,14α-androstane (*I*), 5α,14β-androstane (*II*), 5β,14α-androstane (*III*), and 5β,14β-androstane (*IV*) are presented in the Tables I and II. These models, based on the averaged values of the natural coordinates derived from the X-ray data, will be further abbreviated as SSM's (Standard Steroid Models). Simultaneously we studied models having an idealized structure and corresponding to the Dreiding models (Calculated Dreiding Models, CDM).**

Calculation of the Steroid Model

a) *Algorithm description.* The values of natural coordinates are given for every ring in the molecule as input data; six of them are chosen to be dependent. These coordinates are adjusted using the algorithm[17], which yields besides their corrected values a set of transformation matrices describing the transformations between the local coordinate systems of the individual skeletal atoms. The local coordinates of all substituents bonded to the skeletal atoms are calculated in the next step using simple geometric relations[18,19] and are transformed by a sequence of transformations[20] into one chosen coordinate system (identical with the local coordinate system of the first skeletal atom in this case) using the above mentioned matrices. This procedure is repeated for all rings. A special transformation matrix should be used for the transition from the coordinate system of one ring to the coordinate system of the neighbouring ring and for the transition from the coordinate system of the side chain to the coordinate system of the ring to which this chain is attached. The order of transformations must by chosen before the calculation so that the sense of rotation in the coordinate systems during the transformations is preserved. In order to discriminate the α- and β-substituents and to obtain a clearer presentation of the final coordinates, the transformation into a new coordinate system with origin at C_{10} and the *z*-axis in the C_{10}—C_{19} direction was performed.

* Written in Elliott 503 Algol Mk.1; the listing is available from P.S. upon request.

** Tables of coordinates for these molecules can be obtained from P.S.

TABLE I

Coordinates of the Carbon Atoms and Unit Vectors of the C—H Bonds in Standard Steroid Models SSM-*I* and SSM-*II*

No	Carbon atoms			Substituents			
	x	*y*	*z*	Orient.	*i*	*j*	*k*
				5α,14α-androstane (SSM-*I*)			
1	−0·633	−1·306	−0·514	α	0·034	0·003	−0·999
				β	−0·929	−0·068	0·364
2	0·117	−2·570	−0·051	α	−0·418	−0·797	−0·437
				β	−0·036	−0·045	0·998
3	1·604	−2·560	−0·452	α	0·075	−0·096	−0·993
				β	0·470	−0·757	0·454
4	2·255	−1·232	−0·024	α	0·907	0·061	−0·416
				β	0·032	0·033	0·999
5	1·459	0·000	−0·493	α	−0·033	−0·004	−0·999
6	2·174	1·331	−0·195	α	0·911	−0·038	−0·410
				β	0·051	0·110	0·993
7	1·405	2·524	−0·792	α	0·084	−0·069	−0·994
				β	0·417	0·829	0·372
8	−0·090	2·523	−0·427	β	−0·089	0·199	0·976
9	−0·724	1·151	−0·721	α	0·062	−0·149	−0·987
10	0·000	0·000	0·000				
11	−2·203	1·106	−0·293	α	−0·371	−0·894	−0·250
				β	−0·044	0·143	0·989
12	−3·019	2·205	−0·997	α	0·030	−0·156	−0·987
				β	−0·944	−0·037	0·327
13	−2·426	3·592	−0·684				
14	−0·941	3·620	−1·095	α	0·020	−0·144	−0·989
15	−0·573	5·101	−0·887	α	0·828	0·314	−0·465
				β	0·189	0·275	0·943
16	−1·830	5·801	−1·438	α	0·173	0·322	−0·931
				β	−0·232	0·782	0·579
17	−3·000	4·801	−1·447	α	−0·313	−0·266	−0·912
				β	−0·837	0·334	0·434
18	−2·438	3·883	0·827	Me	0·541	−0·692	0·478
					−0·942	−0·026	0·336
					0·391	0·905	0·169
19	0·000	0·000	1·540	Me	0·530	−0·780	0·333
					−0·940	−0·069	0·333
					0·411	0·848	0·334

TABLE I
(*Continued*)

No	Carbon atoms			Substituents			
	x	*y*	*z*	Orient.	*i*	*j*	*k*
				5α,13β-androstane[a] (SSM-*II*)			
14	−0·941	3·620	−1·095	β	0·353	0·891	0·285
15	−1·029	3·685	−2·631	α	0·826	0·308	−0·472
				β	−0·177	−0·867	−0·465
16	−2·191	4·677	−2·823	α	−0·580	−0·302	−0·756
				β	0·361	0·909	−0·208
17	−3·020	4·735	−1·527	α	−0·984	−0·124	−0·130
				β	0·053	0·871	0·488
18	−2·438	3·883	0·827	Me	0·541	−0·692	0·478
					−0·942	−0·026	0·336
					0·391	0·905	0·169
19	0·000	0·000	1·540	Me	0·530	−0·780	0·333
					−0·940	−0·069	0·333
					0·411	0·848	0·334

[a] The entries for carbon atoms 1 to 13 are equivalent to those for 5α,14α-androstane.

TABLE II
Coordinates of the Carbon Atoms and Unit Vectors of the C—H Bonds in Standard Steroid Models SSM-*III* and SSM-*IV*

No	Carbon atoms			Substituents			
	x	*y*	*z*	Orient.	*i*	*j*	*k*
				5β,14α-androstane (SSM-*III*)			
1	−0·694	−1·275	−0·514	α	0·519	−0·794	0·317
				β	−0·948	0·005	0·317
2	−0·743	−1·364	−2·051	α	−0·534	0·770	−0·349
				β	−0·389	−0·885	−0·256
3	0·652	−1·277	−2·696	α	−0·099	0·094	−0·991
				β	0·516	−0·829	0·215
4	1·411	−0·055	−2·147	α	−0·454	0·830	−0·323
				β	0·954	−0·057	−0·294
5	1·415	0·000	−0·608	β	0·463	−0·816	0·347
6	2·290	1·137	−0·046	α	0·873	−0·010	−0·487
				β	0·127	−0·162	0·978
7	1·620	2·508	−0·245	α	0·011	0·189	−0·982
				β	0·529	0·661	0·532

TABLE II
(*Continued*)

No	Carbon atoms			Substituents			
	x	*y*	*z*	Orient.	*i*	*j*	*k*
				5-β,14α-androstane (SSM-*III*)			
8	0·160	2·544	0·243	β	−0·003	−0·059	0·998
9	−0·639	1·361	−0·334	α	−0·018	0·110	−0·994
10	0·000	0·000	0·000				
11	−2·082	1·336	0·202	α	−0·484	−0·759	−0·435
				β	0·036	−0·119	0·992
12	−2·817	2·651	−0·120	α	−0·051	0·106	−0·993
				β	−0·919	−0·038	0·392
13	−2·055	3·846	0·483				
14	−0·607	3·850	−0·041	α	−0·060	0·120	−0·991
15	−0·065	5·184	0·504	α	0·825	0·351	−0·443
				β	0·280	−0·001	0·960
16	−1·270	6·114	0·269	α	0·146	0·539	−0·829
				β	−0·107	0·620	0·778
17	−2·542	5·260	0·111	α	−0·398	0·013	−0·917
				β	−0·764	0·282	0·580
18	−1·938	3·728	2·013	Me	0·493	−0·838	0·233
					−0·915	−0·028	0·403
					0·499	0·789	0·358
19	0·000	0·000	1·540	Me	0·492	−0·804	0·334
					−0·942	−0·025	0·334
					0·452	0·828	0·333
				5β,14β-androstane[a] (SSM-*IV*)			
14	−0·607	3·850	−0·041	β	0·467	0·749	0·470
15	−0·786	4·325	−1·495	α	−0·302	−0·694	−0·653
				β	0·822	0·347	−0·451
16	−1·843	5·432	−1·320	α	−0·658	−0·038	−0·752
				β	0·446	0·895	0·000
17	−2·574	5·219	0·018	α	−0·998	0·004	−0·067
				β	0·180	0·703	0·688
18	−1·938	3·728	2·013	Me	0·493	−0·838	0·233
					−0·915	−0·028	0·403
					0·499	0·789	0·358
19	0·000	0·000	1·540	Me	0·492	−0·804	0·334
					−0·942	−0·025	0·334
					0·452	0·828	0·333

[a] The entries for carbon atoms 1 to 13 are equivalent to those for 5β,14α-androstane.

b) *Input data.* The weighted averaged values of valency and torsional angles, derived by Geise, Altona, and Romers[21,22] from the X-ray data, were found to be a suitable basis for the calculation of SSM's. Standard values of the C—C and C—H bond lengths (1·54 and 1·09 Å) were used. The ring D was calculated in the most probable conformation[22]. The magnitudes of valency angles were fixed in the six-membered rings whereas the torsional angles were adjusted; in the ring D a reverse procedure was employed owing to the sensitivity of the resulting conformation to the small changes in the torsional angles. The same values of valency angles and the same values of torsional angles but with an opposite sign were taken for the skeletons with the *cis*-junction of the rings. Staggered position was assumed for the hydrogen atoms of the angular methyl groups. A value 109·467° was taken for the angle between the hydrogen atoms in the CH_2-group. The same bond lengths were used in the calculation of CDM's; all valency angles had the tetrahedral value 109·467°, the torsional angles in the six-membered rings were $\pm 60°$; the ring D was calculated in the same way as for SSM.

RESULTS AND DISCUSSION

The coordinates of carbon atoms and the unit vectors of the C—H bonds for SSM's of the four principal steroidal skeletons 5α,14α-androstane (*I*), 5α,14β-androstane (*II*), 5β,14α-androstane (*III*), and 5β,14β-androstane (*IV*) are given in the Tables I and II.

The comparison of the properties of CDM's and SSM's was made using the X-ray determined structures (X-Ray Structures, XRS). SSM and CDM of 5α,14α-androstane (*I*) were compared with XRS of 2β,3α-dichlorocholestane[23] (*V*) and with XRS of 3α-hydroxy-5α-androstan-17-one[24] (*VI*). Table III serves for the comparison of the values of adjusted angles of SSM-*I* and CDM-*I* with XRS-*V* and XRS-*VI*, respectively. It is evident that the difference between the weighted averaged values and the final SSM-*I* parameters is not large. SSM's have – in analogy with real molecules – a slightly flattened shape of the rings with respect to the idealised geometry of the CDM's; the absolute values of the torsional angles are in all cases less than 60°. The vectors C_{10}—C_{19} and C_{13}—C_{18} (the axes of the angular methyls) are parallel in CDM-*I* whereas, they are at angles of 9°, 14°, and 11° in XRS-*V*, XRS-*VI*, and SSM-*I*, respectively. The difference between SSM-*I* and XRS-*V* might be due to the deformation of the ring A caused by the two bulky substituents; the difference between XRS-*VI* and SSM-*I* by the different conformation of the ring D forced by 17-keto group. Further differences between CDM's and SSM's follow from the comparison of 50 randomly selected interatomic distances including both distances of different magnitude and of different type (C···C, C···H, H···H) with corresponding distances in XRS-*V* and XRS-*VI*. SSM's (and also XRS's) are lacking the internal symmetry of CDM's, shown by the equality of certain interatomic

TABLE III

Comparison of Values of the Torsional Angles and Valency Angles in the Ring D of X-Ray Determined Structures XRS-*V* and XRS-*VI* with Standard Model SSM-*I* and Dreiding Model CDM-*I*

Ring	Angle	XRS-*V*	XRS-*VI*	Weighted average (input)	SSM-*I*	CDM-*I*
			Torsional angles			
A	1— 2	—42·9	—55·7	—55·8	—56·5	—60·0
	2— 3	40·9	51·9	52·3	50·7	60·0
	3— 4	—49·7	—52·3	—52·6	—50·9	—60·0
	4— 5	59·2	56·3	56·5	56·9	60·0
	10— 5	—56·7	—55·3	—55·1	—57·6	—60·0
	10— 1	49·4	55·3	54·9	54·3	60·0
B	10— 5	58·0	57·7	58·0	59·5	60·0
	5— 6	—57·8	—57·3	—57·2	—55·7	—60·0
	6— 7	54·3	53·5	53·4	50·2	60·0
	7— 8	—53·1	—51·7	—51·9	—50·2	—60·0
	8— 9	55·7	54·2	54·6	56·1	60·0
	9—10	—56·9	—56·7	—56·9	—60·0	—60·0
C	11—12	—55·0	—54·8	—55·0	—58·0	—60·0
	11— 9	53·1	53·8	53·7	57·7	60·0
	9— 8	—53·4	—52·9	—52·8	—58·0	—60·0
	8—14	58·9	56·3	56·7	56·3	60·0
	14—13	—60·0	—59·0	—59·2	—56·6	—60·0
	13—12	55·8	55·4	55·6	55·3	60·0
D	13—14	46·8	44·0	46·3	46·3	46·3
	14—15	—33·6	—38·6	—41·0	—41·0	—41·0
	15—16	7·6	16·6	20·1	20·1	20·1
	16—17	20·6	11·5	8·5	8·5	8·5
	17—13	—40·4	—34·3	—33·9	—33·9	—33·9
			Valency angles			
D	13—14—15	104·2	104·3		104·9	101·7
	14—15—16	103·6	102·6		102·9	103·2
	15—16—17	106·8	106·0		108·2	106·8
	16—17—13	103·7	107·8		103·2	105·0
	17—13—14	99·8	99·2		100·6	101·8

distances. SSM's are also closer to these real molecules (deviations from XRS-*V* are smaller with SSM-*I* in 30 cases and with CDM-*I* in 12 cases; the deviations from XRS-*VI* are smaller with SSM-*I* in 28 cases and with CDM-*I* in 19 cases). Statistical

evaluation of these deviations (Table IV) shows unequivocally that SSM's describe the XRS's significantly better than CDM's. Since the usual accuracy of the roentgenographic determination of the atom position is 0·01 Å, these results can be regarded as satisfactory. Greater value of the mean deviation in the case of the H···H distances in XRS-*VI* reflects probably the effect of substantially shorter C—H bonds calculated in the ref.[24].

TABLE IV

The Mean Differences in Interatomic Distances for Different Models

Compared models[a]	C···C	H···H	all types[b]
	mean value of difference in distances[c], Å		
SSM-*I*—XRS-*V*	+0·009 (0·094)	−0·010 (0·065)	+0·010 (0·117)
SSM-*I*—XRS-*VI*	+0·018 (0·108)	−0·049 (0·099)	−0·014 (0·103)
CDM-*I*—XRS-*V*	+0·116 (0·331)	−0·017 (0·104)	+0·069 (0·268)
CDM-*I*—XRS-*VI*	+0·127 (0·333)	−0·062 (0·126)	+0·036 (0·261)

[a] Sign of difference corresponds to the subtraction shown in this column. [b] Including mixed type C···H. [c] With the standard deviation (1σ) in parentheses.

Therefore these averaged structures can be considered as a suitable basis for the semiempirical calculations. The presented method allows one to overcome the computationally difficult minimalization procedures, which is an advantage especially when more structures are to be compared with the experiment; furthermore, this method allows one to construct any geometrically possible structure.

REFERENCES

1. Tori K., Kitahonoki K.: J. Am. Chem. Soc. *87*, 386 (1965).
2. Bell R. A., Osakwe E. N. C.: Chem. Commun. *1968*, 1093.
3. ApSimon J. W., Demarco P. V., Raffler A.: Chem. Ind. (London) *1966*, 1792.
4. ApSimon J. W.: Can. J. Chem. *46*, 808 (1968).
5. ApSimon J. W., Craig W. G., Demayo A., Raffler A.: Can. J. Chem. *46*, 809 (1968).
6. Johnson C. K.: *Program ORTEP*, Oak-Ridge Natl. Lab. No 3794 (1965).
7. Vitek A.: *Program MGP-5*, Gier Program Library PA-1290, Inst. Nucl. Res., Prague-Řež 1969.
8. Levinthal C.: Sci. American *214*, 42 (1965).
9. Corey E. J., Wipke W. T.: Science *166*, 178 (1969).
10. Freeman J. M., Henshall T.: J. Mol. Struct. *1*, 172 (1967—8).
11. Dewar M. S.: *Program COORD*, QCPE No 136.

12. Gordon M. S., Pople J. A.: *Program MBLD*, QCPE No 135.
13. Houlden S. A.: *Program CORCAL*, QCPE No 130.
14. Altona C., Hirschmann H.: Tetrahedron *26*, 2173 (1970).
15. Cohen N. C.: Tetrahedron *27*, 789 (1971).
16. Scott R. A., Vanderkooi G., Tuttle R. W., Shames P. M., Sheraga H. A.: Proc. Natl. Acad. Sci. U.S. *58*, 2204 (1967).
17. Vítek A.: This Journal *33*, 1601 (1968).
18. Hendickson J. B.: J. Am. Chem. Soc. *83*, 4537 (1961).
19. McCullough R. L., McMahon P. E.: Trans. Faraday Soc. *60*, 2089 (1964).
20. Eyring H.: Phys. Rev. *39*, 746 (1932).
21. Geise H. J., Altona C., Romers C.: Tetrahedron *23*, 439 (1967).
22. Altona C., Geise H. J., Romers C.: Tetrahedron *24*, 13 (1968).
23. Geise H. J., Romers C., Rutten E. W. M.: Acta Cryst. *20*, 249 (1966).
24. High D. F., Kraut J.: Acta Cryst. *21*, 88 (1966).

Translated by the author (P. S.).