Tetrahedron Vol. 39, No. 7, pp. 1113 to 1121, 1983
Printed in Great Britain

0040-4020/83/091113-09$03.00/0
Pergamon Press Ltd.

A DISTANCE GEOMETRY STUDY OF RING SYSTEMS

APPLICATION TO CYCLOOCTANE, 18-CROWN-6, CYCLODODECANE AND ANDROSTANEDIONE

PAUL K. WEINER,† SALVATORE PROFETA, JR., GEORGES WIPFF, TIM HAVEL, IRWIN D. KUNTZ, ROBERT LANGRIDGE and PETER A. KOLLMAN*
School of Pharmacy, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143, U.S.A.

(*Received in U.S.A. 20 April 1982*)

Abstract—We present the application of distance geometry methods to the generation of structures of a sampling of organic molecules: cyclooctane, cyclododecane, 18-crown-6, and androstane-3,17-dione. The method provides a simple, convenient method to generate either a random sample of molecular conformations, or to generate specified conformations. Using this approach, we found a new, relatively low-energy conformation of 18-crown-6. We also demonstrated that the lowest energy structure of cyclododecane was not exactly of D_4 symmetry, which had been previously assumed in molecular mechanics optimization.

One of the fundamental problems faced by the chemist is the quantitative construction of realistic 3-dimensional structural models of molecules of interest. For simple, acyclic molecules, the method of internal coordinates works best. Here the bond lengths, bond angles, and diheral angles for each atom relative to other atoms in the molecule are specified. From these values, cartesian coordinates may easily be calculated.

For cyclic systems, however, only certain combinations of dihedral angles will close the ring.[1] The use of arbitrary dihedral angles can easily lead to unrealistic structures with chemically unreasonable bond lengths and angles. While these structures can be energy refined with appropriate constraints to produce a low energy structure, this approach may lead to a minimized structure that does not retain all the properties and symmetry of the model structure. One can solve systems of equations to get allowed combinations of dihedral angles,[2] but this method rapidly becomes complicated and is not general.

A second method, the model projection technique, involves the building of a molecular model and the estimation of cartesian coordinates using model projections. This is conceptually very simple. However, as the complexity of the molecule increases, the ability to accurately estimate coordinates decreases (unless one is willing to build larger and larger models). In addition, it may be difficult to impose a desired symmetry on the resulting coordinates.

A third method, distance geometry,[3] uses the distance matrix (the set of distances between all pairs of atoms) to generate a set of cartesian coordinates that satisfies the original set of distances as closely as possible. Many of these distances are known from the local environment of each atom. Other close non-bonded distances may be obtained from experimental data, such as NOE studies.[4] The remaining distances are obtained by (1) creating upper and lower bound conditions on each intramolecular distance and (2) selecting at random a set of intramolecular distances that meet these boundary conditions. Eigenvalue techniques[3] are then used to find the best three dimensional fit to these distances. Finally, these coordinates are refined against an error function formed from the boundary conditions (see below).

As constraints may easily be converted to distances, this method is ideal for model construction. It can quickly generate many structures that satisfy the original constraints. One may, therefore, be able to avoid the need for a grid search, expensive in terms of computer time, or extensive manipulation of the molecule with computer graphics, expensive in terms of human time.

The method of distance geometry fails to converge if these constraints are inconsistent with a three dimensional structure and can generally be made to converge otherwise. The distance geometry approach has had extensive applications to such varied subjects as: analysis of protein structure,[5] protein–ligand interactions[6] and the structure of the 30S ribosome particle.[7] Here, we present the first extensive application of this method to the construction and analysis of small organic molecules. A preliminary version of this work has been presented.[8]

We chose to study cyclooctane, 18-crown-6, and cyclododecane because of their conformational flexibility and ring constraints. We also examined the number of constraints required for generation of a structure similar to the observed X-ray structure in the more complex molecule androstane-3,17-dione.

Methods

The molecular structures are generated in a three step process. First, molecular connectivities and atom information is entered using either an interactive computer graphics template program[8,9] or a user written non-graphics program. Second, EMBED,[10] a distance geometry program, is used to obtain three dimensional coordinates. Finally, these coordinates are refined with molecular mechanics programs, MM2[11] or AMBER.[12]

The template program allows the user to draw a 2-dimensional representation of the molecule using a cursor on a graphics tablet connected to an Evans and Sutherland Picture System 2. The user uses a combination of cursor pointing and typing on the graphics

---

†Present address: E.R. Squibb and Sons, Inc., Princeton, NJ 08540, U.S.A.

keyboard to specify unique atom names, atom types, and any geometrical constraints that should be imposed on the structure. These geometrical constraints consist of distance, angle, or dihedral constraints involving either bonded or non-bonded atoms. Chiral constraints are input as dihedral terms. The constraints can come from prior experimental or theoretical results, as well as from "gedanken" experiments. They consist of both a constraint value (e.g. a specific dihedral of 180°) and an allowed range (e.g. ±10°) from this value.

The data from the template program or a user prepared file is then passed to EMBEDPREP,[13] a program that prepares input for EMBED.[10] The data file need only contain atom names, atom types, connectivities and constraints. All bonds and angles in the molecule are identified from the connectivity information. The specific values for these are determined by matching the atom types in the molecule against an input parameter list of bond lengths and angles. These values become both the upper and lower bound distance matrix elements for the two atoms.

For example, the upper and lower bound matrix elements for two sp^3 carbon atoms bound to each other are set to 1.54 Å. The lower bound distance for an unconstrained dihedral angle is set to the distance between the two end atoms in the *cis* form. The *trans* form is used to determine the upper bound distance. For two atoms separated by more than three bonds, the lower bound element is set to a value smaller than the sum of the Van der Waals radii to allow for H-bond formation, etc. This default lower bound value is a variable parameter in the input and typically a value of ~2 Å has been found to give reasonable results. The upper bound element, also an input parameter, may be set by the program to an estimate of the fully extended chain length. Other explicit non-bonded constraints are converted to appropriate internal distances and then used to set upper and lower bounds as indicated above.

The parameter options and upper and lower distance bounds are then sent to EMBED,[10] the distance geometry program. This program generates sets of three dimensional coordinates that meet the specified constraints. Since the distance geometry technique has been fully described in the literature,[3,5-7] we discuss only the relevant details and modifications.

A novel feature of distance geometry is the use of a random number generator with a uniform distribution to select the internal distances so that they lie between the upper and lower bound values. These boundary conditions are then forced to be internally consistent using the triangle inequalities, as discussed elsewhere.[3,10] Many structures can be generated in this fashion, all meeting the imposed constraints. The differences among the final structures can give insight into how particular constraints affect the 3-dimensional structure. For example, for cyclooctane we were able to use this feature to generate many different structures that span much of its conformational space. For 18-crown-6, the bond length/bond angle data permitted far greater conformational flexibility, and we focused mainly on structures having dihedral angles only in narrow ranges (e.g. 180 ± 30°) for specific bonds.

The methodology for transforming the set of internal distances into a set of 3-dimensional coordinates is straightforward.[3] Essentially one can use the distance matrix to construct a matrix of vector dot products. The principal eigenvalues and eigenvectors of this so-called metric matrix provide the principal axes of the molecular coordinate system. By selecting just the three largest eigenvalues, one obtains the 3-dimensional coordinates that best fit the specific distance matrix being used.

In general, the coordinates found in this manner do not agree completely with the boundary conditions that were entered initially. Optimization procedures using conjugate gradient methods have proved effective in refining the coordinates so that all internal distances fall within the boundary conditions.[14]

The problem of generating appropriate stereochemistry is also handled in the refinement stage of the EMBED program.[3,10] The desired stereochemistry at each center is specified with the signed triple cross product of the bond vectors at the center. Dihedral angle constraints are encoded in the same way. The difference between the desired crossproduct value and that found in a particular conformation is added as a term in the error function. It has been our experience that the stereochemical constraints should be heavily weighted in the early stages of the refinement procedure to avoid unproductive local minima with incorrect stereochemistry, or those which are in the wrong region of dihedral angle space. These weights are then reduced in subsequent refinement.

RESULTS AND DISCUSSION

A. *Cyclooctane*

The cyclooctane molecule represents a special challenge owing to its coupled ring inversion and pseudorotation processes. 8-Member rings form an extensive family of structures of interest to both organic and medicinal chemists and much spectroscopic work has been published. However, while much time has been invested in these studies they have been, quoting Anet[15] "... singularly unrewarding". The cyclooctane system itself contains too many atoms for a reliable analysis by IR and Raman spectroscopy[16] and X-ray diffraction work gives little information due to crystal disorder.[17] Electron diffraction[18] and NMR studies[19] conclusively that the biangular[26] "boat-chair" concoordinate analysis by Strauss[20] indicates that the boat-chair conformer predominates, supporting the previous conclusions of Anet.[19] X-ray analyses of the oxime and phenylsemicarbazone of cyclooctanone by Groth[21] show conclusively that the biangular [26], "boat-chair" conformation is favored.

While we are by no means the first group to carry out conformational analysis of cyclooctane,[11,22-28] we have taken a unique approach to generating the individual conformers. We suggest that, with proper intuitive gudiance, a distance geometry approach to conformer generation can be both efficient and comprehensive. Hendrickson's papers[22-25] detailed a protocol for examining symmetrical ring conformations for 6- to 10-membered rings without recourse to models. His extensive examination of these ring systems was aimed at a comprehensive survey of: (1) detailed ring geometries and their relative energies; (2) modes of conformational interconversions between forms and the associated activation energies; and (3) substituent strain energies of the geometrically distinguishable substituted forms. To facilitate the generation of families of ring structures, Hendrickson devised a simple, yet effective notation which represented the ring forms with either a plane or axis of symmetry.[22] While this notation does allow for direct translation into molecular models (e.g. Dreiding), it does place symmetry restrictions on the possible

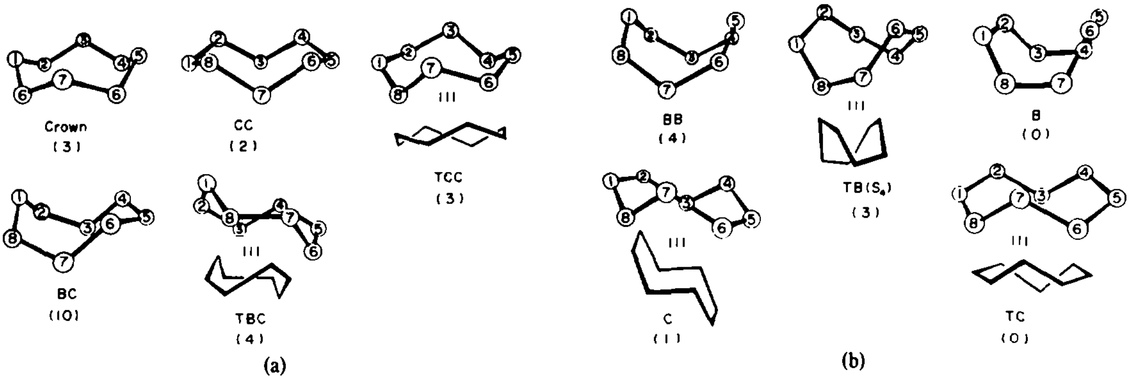


Fig. 1. Cyclooctane conformations. Values in parentheses represent frequency of generation of this form using 30 random numbers.

combinations of dihedral signs and the reader is referred to the literature for a complete review of this methodology.[22]

We attempted to generate cyclooctane conformations using distance geometry. For the initial trials, we used the following geometric constraints: C–C bond lengths equal 1.54 Å, C–C–C angles equal 109.5°. Torsional angle constraints were not specified. Using thirty (30) random numbers the distribution seen in Figs. 1(a) and (b) is obtained. It is interesting that even though there is no energy term in these geometrical determinations, the lowest energy boat-chair form (BC) was the most frequently generated. In addition, one can clearly see a correlation between the frequency of a conformer and the "extended" or "non-congested" nature of the conformer. This relationship is more clearly evidenced in the near exclusion of the most congested forms such as the Boat (B), Chair (C) or twist-chair (TC).

At this point we introduced dihedral angle constraints to help complete the generation of all ten forms in Figs. 1(a) and (b). We find that at least two dihedral angles need be specified to generate the twist-chair (TC) conformer and at least four dihedrals are needed to generate the Boat (B) and Chair (C) conformers with the proper symmetries. Additionally, a minimum valence angle of 115.5° must be used to permit successful embedding[10] and the generation of structures with chemically reasonable C–C bond lengths and C–C–C angles. It can be seen, however, that even with these additional constraints, the method does not require the overdetermination of structure common to model projection or internal coordinate techniques.

The coordinates generated were energy refined (after placing hydrogens on the molecule) using MM2CDC[29] and typical results for the chair conformation of cyclooctane are presented in Table 1. In the table we list the initial steric energies of the structures generated by the internal coordinate, distance geometry, and model projection techniques.[30] We used no symmetry constraints in any of the coordinate generation methods or during the subsequent energy refinements and this is clear from the geometry of the minimized structure which is reached in all three cases (Table 1). The optimized structure (E = 21.1 kcal/mole) corresponds to a twist-boat-chair (TBC) conformation, which has C_2 symmetry. This result should not be surprising considering the very high steric energy of the chair conformer relative to the boat-chair ($\Delta E = 7.5$ kcal/mole), and that the twist-boat-

Table 1. Typical steric energies and convergence data for cyclooctane conformers: the chair conformation

	Coordinate generation method			
Energy components (kcal/mole)	Internal coordinate method[a,b]	Distance geometry method[a]	Model projection[a,c]	MM2 optimized energies[d]
Compression	4.2	5.1	3.5	0.6
Bending	4.0	8.8	4.4	5.0
Stretch–bend	0.4	0.6	0.3	0.4
1,4-van der Waals	10.9	11.1	11.2	7.9
Torsional	12.3	11.0	12.3	8.1
Total initial/final energy	35.4	37.7	33.3	21.1
Number of iterations	61	41	66	—

[a] Initial energies.

[b] Internal Coordinate Method as available in programs COORD or PROXYZ.

[c] See Ref. 27.

[d] The three structures minimized to the same steric energy using MM2CDC. It should be noted that no C_{Sh} symmetry constraints were used during the energy refinement and that the final energies listed correspond to the twist-boat-chair conformer.

chair is most likely a relative minimum ($\Delta E = 1.67$ kcal/mole) along the pseudorotation pathway between the chair and boat-chair forms.

The results in Table 1 highlight the usefulness of the distance geometry generated structures in terms of efficient generation of structures which are close to optimal. In terms of energies, the distance geometry structures for cyclooctanes were generally somewhat higher than those of the other methods; however, they took on the average 30–40% fewer iterations to converge to an optimum structure, a factor which becomes more significant as the system size increases. It also required considerably less effort to generate such structures, since the procedure is automatic once the user draws in the two dimensional connectivity of the molecule.

B. 18-*Crown*-6

A second molecule to which we applied the distance geometry approach was 18-crown-6, $(CH_2CH_2O)_6$. This molecule is of interest because it and its derivatives are probably the most studied crown ethers, and there are X-ray structures for a number of different conformations of this molecule. In particular, 18-crown-6 has a different conformation when crystallized alone (C_i),[31] when crystallized with K^+, Rb^+ or Cs^+ (D_{3d}),[32] when crystallized with Na^+ (C_1),[33] or when crystallized with $(NH_4)_2SO_4(C_i')$.[34] We have previously carried out molecular mechanics studies of 18-crown-6 and metal complexes, using X-ray coordinates as input geometries.[35] We thus wished to compare such an approach to one in which a distance geometry approach is used to generate input geometries.

We first generated structures in which we constrained the dihedral angles to a particular *range* of conformation space. For example, the D_{3d} structure was generated by specifying only that the OCCO angles be alternating g^+ and g^- ($70 \pm 20°$) and the CCOC angles all be near *trans* (180 ± 20). The "generation" of the C_i' structure was similarly straightforward.

In the case of the C_2 structure, which was originally generated from C_i by changing a few key torsional angles,[35] there is an angle not near either *trans* or gauche ($\varphi_7 = -138°$); we generated a structure that successfully refined to the "X-ray" geometry by specifying the φ_7 be -150 ± 20 (Table 2).

In the case of the C_1 structure, we generated the initial geometry using only the appropriate three possible angle choices (g^+, g^-, t) as we had done for the C_i and D_{3d} geometries; upon energy refinement, this led to a *lower* energy structure than found by refining from the X-ray structure. The relative lack of stability of the X-ray C_1 structure is not surprising, since it is found only in the crystal structure with Na^+ and involves significant strong $Na^+ \ldots O$ attractions to compensate for the rather higher strain energy of the crown. The dihedral angles which are most "stained" (φ_{10}, φ_{16} and φ_{17}) remain more "strained" upon refinement starting with the X-ray geometry than upon refinement starting from the distance geometry generated structure. This is to be expected since the "initial" distance geometry structure did not contain highly strained angles either before or after energy refinement. Its internal strain energy is almost identical to that of the C_i structure, but its electrostatic energy (mainly coming from $O \ldots O$ repulsions) is significantly higher, as one might expect for a structure which seeks to optimize its $O \ldots Na^+$ interactions.

A comparison of the distance geometry and X-ray initial and energy refined structures is given in Table 2 and suggest that distance geometry is indeed a powerful approach to generating reasonable, low energy structures without the imposition of artifical ring closure constraints or the necessity of X-ray structures. In particular, one can build approximate structures with wire models, estimate the dihedral angles from these, and use these as input in the distance geometry calculation.

We then generated other structures that did not come from X-ray determined models. The two lowest energy structures previously had been of D_{3d} and C_i symmetry and the lowest non centrosymmetric structure is ~3 kcal/mole higher in energy. This is inconsistent with the observation of a dipole moment of ~1D for 18-crown-6,[36] which increases with increasing temperature. We then used distance geometry methods to generate structure C_i', using as the initial conformation those dihedrals from the C_i structure, with the exception of φ_7, φ_8 and φ_9, which we changed from the values of t, t, g^+ in the C_i structure to the g^+, t, t of the D_{3d} structure. Upon energy refinement, this structure was only 0.7 kcal/mole above C_i and 0.4 kcal/mole more stable than D_{3d}. The calculated dipole moment for this structure was 1.9D. Thus, at 298K, the calculated dipole moment for an equilibrium mixture of D_{3d}, C_i and C_i' structures would be ~0.6D, in qualitative agreement with experiment. In the above calculation, we have assumed a double degeneracy of C_i', since one could equally well change φ_{16}, φ_{17} and φ_{18} to t, t, g^+. We thus predict that this C_i' structure contributes significantly to the average solution properties of 18-crown-6.

We also generated a number of structures with no dihedral constraints; although most, upon energy refinement, had relatively high energies (50–53 kcal/mole), one structure had an energy comparable to the D_{3d}, C_i and C_i' structure (~48 kcal/mole). It was a noncentrosymmetric structure with a dipole moment comparable to the C_i' structure and should be considered in evaluating the overall solution properties of 18-crown-6.

C. *Cyclododecane*

In our study of cyclododecane, we found an example of the failure of internal coordinate generation to properly close a ring structure, thus causing the loss of symmetry of the starting conformation and subsequent energy refinement to another minimum.

We attempted to generate the square, [3333] structure, which is not "diamond lattice" owing to its repetitive $tg \pm g \pm t$ units, using both distance geometry and the internal coordinate method. The task of generating the structure using the internal coordinate method proved much harder than first thought. Owing to the larger internal bond angles in the 12-membered ring, it was thought that using the experimental values (~114.5)[37–39] would help close the ring more accurately than if the default values of 109.5 or 112.3° were used. This did not prove to be correct, as the structures generated using all three bond angle choices were similarly deformed. All attempts to derive a structure with at least D_2 symmetry using the internal coordinate method failed. The structures were of C_1 symmetry exclusively.

Conversely, the distance geometry technique generated structures which were minimally of C_2 symmetry (nearer to D_2) and which, when energy refined with MM2VAX,[41] became ~D_4 symmetry. In generating this structure, we used default bond lengths of 1.53 Å,

Table 2. Structures and energies for distance geometry and crystal structures of 18-crown-6

	Conformation					
	D_{3d}		C_i		C_i'	
	X-ray	Distance geometry	X-ray	Distance geometry	X-ray	Distance geometry
E_T (initial)[b]	55.2	63.8	53.4	60.4	57.0	68.9
E_T (refined)[b]	48.3	48.3	47.2	47.0	50.2	50.2
φ_1^c	72 (65)	68 (70)	−179 (−175)	165 (−179)	66 (68)	63 (64)
φ_2	−169 (177)	180 (−172)	−179 (−170)	173 (173)	−168 (−173)	180 (−165)
φ_3	169 (−171)	−178 (168)	−79 (−80)	−90 (−81)	167 (177)	180 (167)
φ_4	−72 (−65)	−79 (−72)	72 (75)	53 (69)	−69 (−66)	−60 (−69)
φ_5	169 (179)	−176 (176)	−154 (−155)	−176 (−171)	−166 (−158)	180 (−168)
φ_6	−170 (178)	−178 (−165)	158 (166)	−175 (164)	−176 (−179)	180 (−176)
φ_7	72 (70)	83 (71)	−72 (−68)	−78 (−73)	−65 (−67)	−58 (−66)
φ_8	−170 (−176)	177 (−174)	161 (176)	175 (158)	173 (180)	180 (174)
φ_9	169 (−177)	172 (−193)	178 (175)	−172 (−176)	−76 (−73)	−57 (−73)
φ_{10}	−72 (−66)	−93 (−72)	179 (175)	178 (179)	−66 (−68)	−64 (−64)
φ_{11}	169 (−178)	180 (173)	179 (170)	−179 (−173)	167 (173)	180 (166)
φ_{12}	−170 (171)	−172 (−170)	79 (80)	86 (82)	−167 (−177)	180 (−167)
φ_{13}	72 (65)	90 (71)	−72 (−75)	−68 (−70)	69 (66)	62 (69)
φ_{14}	−170 (−179)	180 (−170)	156 (155)	−176 (169)	166 (158)	180 (168)
φ_{15}	169 (−178)	173 (172)	−158 (−166)	−177 (−168)	176 (179)	180 (176)
φ_{16}	−72 (−70)	−80 (−73)	72 (68)	73 (72)	65 (67)	59 (66)
φ_{17}	169 (175)	−176 (169)	−161 (−176)	−166 (−153)	−173 (180)	180 (−174)
φ_{18}	169 (177)	−175 (−168)	−178 (−175)	177 (179)	76 (73)	57 (74)

[a] Generated from half of C_i structure, setting φ_2 and φ_{11} gauche.

[b] Energy refinement as described in Ref. 35, using the molecular mechanics software AMBER.[12] Energy in kcal/mole.

[c] The torsional angles for 18-crown-6, in degrees starting with an OCCO angle. The initial dihedral angles are listed, with energy refined values in parentheses.

Table 2 (*Contd*)

	Conformation				
	C_2		C_1		C_i'
	X-ray[a]	Distance geometry	X-ray	Distance geometry	Distance geometry
E_T (initial)	53.7	69.4	63.0	70.7	58.1
E_T (refined)	49.7	49.3	56.6	52.8	47.9
φ_1	−179 (−179)	−179 (180)	−62 (−59)	−70 (−68)	69 (72)
φ_2	−70 (−65)	−70 (−78)	158 (167)	−165 (−172)	180 (182)
φ_3	−67 (−56)	−86 (−74)	−174 (173)	−165 (−142)	179 (177)
φ_4	75 (80)	65 (75)	68 (61)	84 (74)	−91 (−73)
φ_5	−138 (−142)	−152 (−125)	−158 (−171)	175 (−176)	179 (176)
φ_6	155 (176)	177 (160)	164 (−177)	177 (171)	178 (−178)
φ_7	−69 (−174)	−59 (−70)	−58 (−59)	−85 (−73)	78 (73)
φ_8	−179 (−178)	180 (173)	180 (−173)	177 (157)	−173 (−165)
φ_9	−170 (−139)	178 (−174)	−166 (−174)	177 (−179)	180 (171)
φ_{10}	−179 (−178)	179 (180)	37 (52)	77 (67)	−53 (−69)
φ_{11}	−70 (−65)	−79 (−78)	79 (71)	74 (88)	176 (164)
φ_{12}	−67 (−56)	−73 (−73)	−155 (−172)	−175 (−167)	174 (−166)
φ_{13}	74 (81)	92 (75)	67 (63)	76 (73)	55 (71)
φ_{14}	−138 (−141)	−156 (−128)	−170 (−178)	176 (178)	−165 (−140)
φ_{15}	155 (126)	177 (159)	77 (77)	82 (83)	175 (168)
φ_{16}	−68 (−74)	−77 (−69)	37 (47)	63 (59)	−165 (180)
φ_{17}	−179 (−176)	−168 (173)	114 (115)	85 (81)	171 (175)
φ_{18}	−170 (−137)	−172 (−171)	−75 (−74)	−103 (−111)	−82 (−85)

Table 3. A comparison of distance geometry and internal coordinate generated cyclododecanes

Feature	Internal coordinate Initial	Internal coordinate Final*	Distance geometry Initial	Distance geometry Final*
Bond lengths (Å)				
1–12	3.630	1.539	1.530	1.530
All others	1.530	1.542(Ave)	1.537(Ave)	1.541(Ave)
Bond angles*				
1–2–3	112.5	117.4	113.9	114.3
1–12–11	59.0	113.6	104.8	114.9
2–3–4	112.5	117.4	105.9	115.0
2–1–12	144.8	118.0	119.4	114.2
3–4–5	112.5	115.1	110.8	114.3
4–5–6	112.5	114.3	110.9	114.2
5–6–7	112.5	115.9	105.7	114.9
6–7–8	112.5	114.6	118.9	114.2
7–8–9	112.5	114.9	115.4	114.3
8–9–10	112.5	116.3	106.2	115.0
9–10–11	112.5	114.1	111.6	114.3
10–11–12	112.5	113.4	111.1	114.2
Dihedrals*				
1–2–3–4	60.0	82.5	72.7	67.8
1–12–11–10	150.8	166.4	56.8	68.3
2–3–4–5	60.0	65.8	59.7	68.9
2–1–12–11	−94.1	−55.1	86.2	67.4
3–4–5–6	180.0	−175.6	−174.4	−171.3
3–2–1–12	−8.2	−61.2	−163.5	147.4
4–5–6–7	60.0	61.7	57.5	68.3
5–6–7–8	60.0	67.3	83.3	68.9
6–7–8–9	180.0	−148.4	−160.9	147.4
7–8–9–10	60.0	58.2	74.0	67.8
8–9–10–11	60.0	54.2	61.0	67.1
9–10–11–12	180.0	−161.9	−172.6	−171.3
Energy (kcal/mol)	169.0	21.69	73.12	20.61
Iterations		87.0		56.0

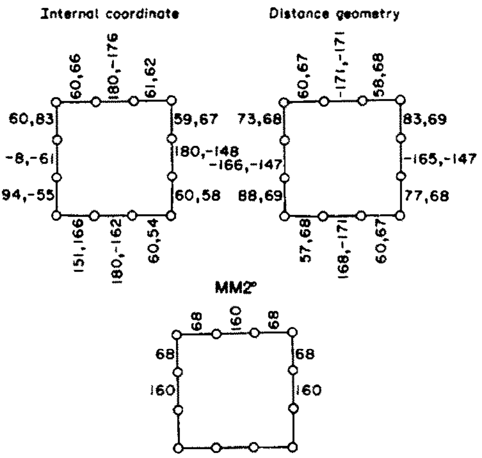
*Minimized using the MM2VAX[41] program.

112.3° for bond angles, and specified the twelve dihedrals using dihedrals of ±60° for gauche and 180° for *trans* with a "range" of ±20 degrees. We wish to emphasize that while we used an equal number of specified constraints in both techniques, only the distance geometry generated structure satisfied *all* the constraints.

In Table 3 and Fig. 2 we detail the initial and optimized structures for cyclododecane as generated by the internal coordinate and distance geometry techniques. Clearly, the internal coordinate technique fails to close the ring properly, as is shown by the extended C1–C12 distance (3.63 Å) and the associated bond and dihedral angles. While the bond angles of the initial distance geometry structure are somewhat irregular, varying from 106–120°, the dihedral constraints are clearly satisfied in a fashion superior to those in the internal coordinate method. The C_1 internal coordinate structure minimized to an energy only 1 kcal/mol above the ~D_4, while the distance geometry structure refined to an energy 0.1 kcal/mole *lower* than that previously found by MM2 in which D_4 symmetry had been imposed.[40] That this was a real result was confirmed by relaxing the symmetry constraints and driving the MM2 D_4 structure toward the distance geometry MM2 optimized structure and examining the energy along this path.

D. *Androstane*-3,17-*dione*

A final test case for the distance geometry technique was a more complex multi-ring system. We chose androstane-3,17-dione for several reasons, the most crucial of which were: (1) the crystal structure was known[42] and therefore the essential structural data were available to check our results; (2) this system represented a molecule of a size which is commonly studied in medicinal chemical problems; and (3) an accurate force field treatment could be done to evaluate the quality of the distance geometry generated structure.[40] Our goal was simple—generate a structure as close to that found in the crystal with as few constraints as possible.



*Assume D_4 Symmetry, These values are within 2° of the experimental values [32]

Fig. 2. A comparison of dihedral angle values for cyclododecane (initial, refined).

We decided initially that a minimum of 2 and a maximum of 5 non-bonded distances would be used together with default bond length and bond angle values. To these we would add dihedral constraints until a satisfactory (by visual examination) structure was found. While 50 or so schemes were used (with 10 sets of random numbers) the final sets of constraints listed in Table 4 were found to be required for successful generation of a visually acceptable structure for androstane-3,17-dione. Also, we wished to keep the necessary data for constraint choices as simple as possible, essentially confining it to readily measured (off a Dreiding or Feiser model) quantities, as these would be most accessible to the average chemist.[43]

From numerous trials three trends became apparent: (1) the A-ring is both floppy (hard to constrain) and highly coupled to all other distance and adjacent dihedral constraints; (2) dihedral (or chiral) constraints are necessary across the ring junctions to insure the *trans, trnas, trans* skeleton; (3) a *combination* of non-bonded distances and dihedrals are imperative for a reasonable structure. It also became clear that in a system with as many degrees of freedom as this steroid has it is possible, if not likely, that all the constraints can be met, yet the structure does not appear at all like one might expect. For example, if we employ all the constraints of Table 3 except the non-bonded distances, we generate the structure shown in Fig. 3. In the generation of this structure, even though we had specified the torsional angles H5–C5–C10–C19 and H14–C14–C13–C18 to be near *trans* (175 ± 5°), the program did not successfully meet these constraints, nor did it give chemically reasonable ring geometries. However, this fact demonstrates the range of possible structures which can easily be accessed by the user. By utilizing an appropriate sampling of random

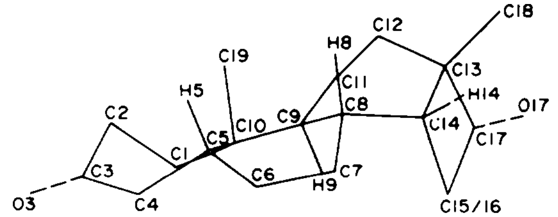


Fig. 3. A drawing of the distance geometry generated androstane-3, 17-dione structure using only the dihedral angle constraints listed in Table 4.

numbers and well-chosen constraints, *an unbiased* sample of several conformational and configurational spaces can be obtained.

The best distance geometry structure for androstane-3,17-dione, shown in Fig. 4, was obtained using the constraints listed in Table 4. It was imperative to specify dihedral constraints "*trans*" across the ring junctions, as well as to specify 4 of the 7 dihedral angles determining the conformation of the A-ring. However, the total of 9 dihedrals represents less than half of the total number of necessary dihedrals using any internal coordinate method. In addition, we had to specify the C18–C19, 03–017, 03–C18 and 017–C19 non-bonded distances, in order to get reasonable conformations of both the A-ring and the correct orientation of the angular methyls C18 and C19. All non-bonded distance and dihedral constraints are satisfied, although they do impose a slight exaggeration of the curvature of the steroid skeleton. However, on the whole, the structure generated compares favorable with both the X-ray results and the MM2VAX[41] minimized geometry, as seen in Table 4. The principal deviations in the initial distance geometry structure occur in the D-ring, where we have allowed the ring to assume any pucker which satisfies the constraints used, none of which involve a specific pucker for this ring. However, upon energy refinement of the distance geometry structure using MM2VAX, the D-ring assumes the same pucker as found in the crystal structure. The discrepancies between the MM2 refined and X-ray values for the A-ring dihedrals and consequently the 03–017 distance may come from the fact that in the crystal environment the carbonyl groups are stacked antiparallel over one another and this appears to introduce significant further flattening of the A-ring (compare cyclohexanone where the dihedral angle corresponding to the 1-2-3-4 dihedral angle in androsterone is 54.4°).[43] The other calculated geometrical variables show very good agreement with experiment.[42] Our conclusion regarding the application of distance geometry to this problem is that the technique allows easy access to a variety of plausible structures and these structures can be energy refined quite easily.

Table 4. An evaluation of the distance geometry generated structure for andiostane-3,17-dione

Feature	Constraint value	Actual value
Interatomic distances (Å)		
03–017	10.7 ± 0.5[a]	10.63
C18–C19	4.7 ± 0.25	4.93
03–C18	9.0 ± 0.50	9.87
017–C19	7.7 ± 0.50	7.20
Bond lengths (Å)[b]		
Csp^3–Csp^3	1.540	1.550 (Ave)
Csp^3–Csp^2	1.515	1.525 (Ave)
$Csp^2 = Osp^2$	1.210	1.200 (Ave)
Csp^3–H	1.110	1.106 (Ave)
Bond angles((°)[b]		
x-Csp^3–y (x, y = C)	112.3	96–139
x-Csp^3–y (x or y = H)	109.5	104–128
x-Csp^2–y	120.0	110–124
Dihedral angles (°)		
H5–C5–C9–C19	175.0 ± 0.0	174.6
H8–C8–C9–H9	175.0 ± 0.0	−158.0
H14–C14–C13–C18	175.0 ± 0.0	175.4
C3–C4–C5–C6	180.0 ± 0.0	−179.9
C1–C2–C3–03	240.0 ± 0.0	237.5
C1–C2–C3–C4	60.0 ± 0.0	56.7
C2–C1–C10–C9	175.0 ± 0.0	175.3
C1–C10–C9–C8	180.0 ± 0.0	−179.6
C11–C12–C13–C17	175.0 ± 0.0	174.4

[a] These non-bonded distances derived from crystallographic data, see Ref. 42.

[b] These bond lengths and angles are implicit in the data base and were used unchanged.

CONCLUSIONS

We have used distance geometry to generate the structure of a number of simple and complex ring systems. Such generations have been carried out with or without aliphatic hydrogens and the structures refined using two different molecular mechanics programs, MM2 and AMBER. We should emphasize that the generation of reasonable structures is, of course, independent of method for subsequent refinement of such structures.

For a system like cyclooctane, we can use random numbers to consistently generate most of the reasonable conformations with no dihedral constraints, although some of the more congested conformers require such constraints.

We do require torsional angle constraints to consistently generate "reasonable" structures for a "floppy" system like 18-crown-6 or a more complex system like androstane-3,17-dione. However, with reasonably chosen constraints, it is quite simple for the experienced chemist

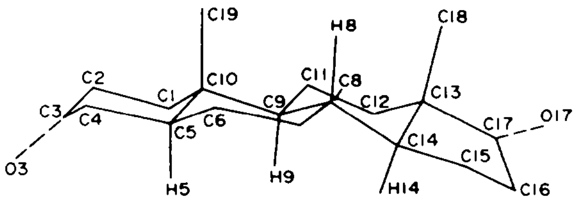


Fig. 4. A drawing of the initial distance geometry generated structure for androstane-3,17-dione using the "minimal" constraints in Table 4.

Table 5. A comparison of selected dihedral angles and non-bonded distances in androstane-3,17-dione

Dihedral angle (°)	Distance geometry Value	MM2 optimized[41]	X-ray[42]
1-2-3-4	56.7	48.1	33.4
1-2-3-03	−123.5	−134.3	−152.6
1-10-5-4	−55.1	−56.4	−59.9
2-1-10-9	175.3	172.1	171.3
2-1-10-19	−70.0	−66.2	−66.3
3-4-5-6	−179.7	−177.7	176.3
7-8-9-10	52.7	55.4	58.2
8-9-10-5	−53.4	−55.6	−59.4
8-14-13-12	−64.8	−64.0	−64.5
11-9-8-14	−75.2	−50.8	−52.0
15-14-13-17	27.4	45.5	47.0
14-15-16-17	47.9	17.1	18.4
Non-bonded distances (Å)			
03-017	10.63	10.54	10.72
C18-C19	4.93	4.74	4.69
03-C18	8.97	9.04	9.02
017-C19	7.20	7.19	7.18

to generate coordinates for structures of any complex organic molecule. Our study of cyclododecane demonstrates that the distance geometry method is indeed superior to the internal cooordinate method for such structure generations. For complex molecules, it is clear that distance geometry requires much less effort than internal coordinate or model projection techniques.

Finally, we emphasize that the approach can be used effectively by the chemist to either generate structures with particular features built in (distances or angles between pharmacophores, dihedral angle relationships, stereochemistry), or to randomly generate structures with minimal constraints, in order to find new and unexpected conformational relationships.

Note added in proof: After the completion of this work another article came to our attention that also uses connectivity lists, atom types, and user specified constraints to generate a 3-D structure by distance geometry (J. C. Wenger and D. H. Smith, *J. Chem. Inf. Comp. Sci.* **22**, 34 (1982)). This paper presents a clearly written, detailed explanation of the basic method of distance geometry. However our paper explores in much greater detail the potential applications. It should be noted that the Wenger and Smith paper is incorrect in stating that we (referring to our work in Ref. 8) use two-dimensional screen coordinates as our starting point. The graphical sketch is only a convenient means of entering connectivity and constraints.

Acknowledgements—This work was carried out at the University of California San Francisco Computer Graphics Laboratory, R. Langridge, director, supported by RR-1081. PAK and IDK would like to acknowledge research support (GM-29072 to PAK and GM-19267 to IDK). G. Wipff thanks the NSF for a U.S.-French Binational Fellowship. S.P. is grateful for support by the National Resource for Computation in Chemistry under a grant from the National Science Foundation (No. CHE-7721305) and by the Basic Energy Division of the U.S. Dept. of Energy (Contract No. W-7405-ENG-48). T. Havel was supported by NIH Training Grant GM-07379 and NASA Training Grant 05-029-800.

REFERENCES AND NOTES

[1] N. C. Cohen, P. Colin and G. Lemoine, *Tetrahedron* **37**, 1711–1721 (1981).

[2] M. Dygent, N. Go and H. Scheraga, *Macromolecules* **8**, 750–761 (1975).

[3] G. M. Crippen, *Distance Geometry and Conformational Calculations*. Research Studies Press–Wiley, New York (1981).

[4] W. Braun, C. Bosch, L. Brown, N. Go and K. Wurtrich, *Biochem. Biophys. Acta* **667**, 377–396 (1981).

[5] T. F. Havel, G. M. Crippen, I. D. Kuntz, *Biopolymers* **18**, 73–81 (1979).

[6] G. M. Crippen, *J. Med. Chem.* **22**, 988–997 (1979).

[7] I. D. Kuntz and G. M. Crippen, *Biophys. J.* **32**, 677–696 (1980).

[8] P. K. Weiner and S. Profeta, Jr., *Abstracts of Meeting*, 181st American Chemical Society Meeting, Atlanta, GA, Mar.-Apr. 1981, Abstract No. Comp 15.

[9] TEMPLATE, a program written at UCSF Computer Graphics Laboratory by P. K. Weiner.

[10] Original version of EMBED was written by G. Crippen, I. D. Kuntz, T. Nordland and T. Havel at UCSF and modified by P. Weiner.

[11] N. L. Allinger, *J. Am. Chem. Soc.* **99**, 8127 (1977).

[12] P. K. Weiner and P. A. Kollman, *J. Comp. Chem.* **2**, 287 (1981).

[13] Program written by P. K. Weiner.

[14] T. Havel, Ph.D. thesis, U.C. Berkeley, 1982.

[15] F. A. L. Anet, *Top. Curr. Chem.* **45**, 169 (1974).

[16] H. E. Bellis and E. J. Slowinski, Jr., *Spectrochim. Acta* **15**, 1103 (1959).

[17] D. E. Sands and V. W. Day, *Acta Cryst.* **19**, 278 (1965).

[18] A. Almenninger, O. Bastiansen and H. Jensen, *Acta Chem. Scand.* **20**, 2689 (1966).

[19] F. A. L. Anet and M. St. Jacques, *J. Am. Chem. Soc.* **88**, 2585 (1966).

[20] H. E. Strauss, personal communication; *J. Phys. Chem.* in press.

[21] P. Groth, *Acta Chem. Scand.* **A35**, 117 (1981).

[22] J. B. Hendrickson, *J. Am. Chem. Soc.* **86**, 4584 (1964).

[23] J. B. Hendrickson, *Ibid.* **89**, 736 (1967).

[24] J. B. Hendrickson, *Ibid.* **89**, 7043 (1967).

[25] J. B. Hendrickson, *Ibid.* **89**, 7047 (1967).

[26] K. B. Wiberg, *Ibid.* **87**, 1070 (1965).

[27] M. Bixon and S. Lifson, *Tetrahedron* **23**, 769 (1967).

[28] N. L. Allinger, J. A. Hirsch, M. A. Miller, I. J. Tyminski, F. A. VanCatledge, *J. Am. Chem. Soc.* **90**, 1199 (1968).

[29] S. Profeta, Jr., *QCPE Bulletin* 1, 57 (1981); Program No. 423.

[30] We have applied this technique in this particular case in the following way: (1) carbon atom coordinates were projected onto paper by using a Dreiding model; (2) to these coordinates we added hydrogens using MM2CDC and minimized the energy using the C_{2h} symmetry constraints necessary to maintain the

chair geometry; (3) the optimized carbon atom coordinates corresponding to an exact chair symmetry were extracted from the optimized coordinates; (4) to these coordinates we again added hydrogens (at 109.5° angles and 1.095 Å bond lengths) using MM2CDC, used *no* symmetry constraints, and allowed energy refinement to proceed. We did this to test the usefulness of using partially optimized structures and to compare the efficiency of this coordinate transfer technique to distance geometry.

[31] J. D. Dunitz, M. Dobler, P. Seiler and R. P. Phizacherly, *Acta Cryst.* **B30**, 2733 (1974) and J. D. Dunitz and P. Seiler, *Ibid.* **B30**, 2739 (1974).

[32] P. Seiler, M. Dobler and J. D. Dunitz, *Ibid.* **B30**, 2744 (1974) and P. Groth, *Acta Chem. Scand.* **25**, 3189 (1971); M. Dobler, R. Phizacherly and J. D. Dunitz, *Acta Cryst.* **30**, 2746, 2748 (1974).

[33] M. Dobler, J. D. Dunitz and P. Seiler, *Ibid.* **30**, 2741 (1974).

[34] O. Nakano, A. Kobayashi and Y. Sasaki, *Bull. Chem. Soc. Japan* **51**, 790 (1978).

[35] G. Wipff, P. K. Weiner and P. A. Kollman, *A Molecular Mechanical Study of 18-crown-6 and Its Alkali Complexes: An Analysis of Structural Flexibility, Ligand Specificity and the Macrocyclic Effect, J. Am. Chem. Soc.* **104**, 3249 (1982).

[36] L. Caswell and D. Savannut, *Solvent and Temperature Effects on the Experimental Dipole Movement of Crown Ethers*, presented at the Southeast/Southwest Regional Meeting of the ACS, Dec. (1980).

[37] J. D. Dunitz and H. M. M. Shearer, *Helv. Chem. Acta* **43**, 18 (1960).

[38] F. A. L. Anet and T. N. Rawdah, *J. Am. Chem. Soc.* **100**, 7166 (1978).

[39] P. Groth, *Acta Chem. Scand.* **A33**, 203 (1979).

[40] N. L. Allinger and S. Profeta, Jr., unpublished results; see also, N. L. Allinger, *J. Am. Chem. Soc.* **99**, 8127 (1977) for other MM2 alkane results.

[41] S. Profeta, Jr. and D. Pensak, VAX 11/780 version of MM2.

[42] V. M. Coiro, E. Giglio, A. Lucano and R. Puliti, *Acta Cryst.* **B29**, 1404 (1973).

[43] See N. L. Allinger, M. T. Tribble and Y. H. Yuh, *Steroids* **16**, 398 (1975) for an earlier discussion of a similar treatment of androsterone.