RINGS — a general program to build ring systems

C W v d Lieth, R E Carter*, Daniel P Dolata* and Tommy Liljefors*

Deutsches Krebsforschungszentrum, Zentrale Abteilung Spektroskopie, Im Neuenheimer Feld, Heidelberg, FRG

* University of Lund, Chemical Center, Molecular Graphics Laboratory for Organic Chemistry, Organic Chemistry 2 and 3, S-221 00 Lund, Sweden

RINGS is a computer program for the general and rapid construction of ring skeletons. Prestored molecules are connected together by various operations, such as joining along one bond, fusion of one or more bonds, spiro joining, and carbon-bridge building. RINGS uses full 3D molecular structure information from the outset and offers the user full control of all geometrical features at each step of construction. The program is user-friendly, with a graphical man–machine interface (PLOT10-compatible) and extensive checking for possible input errors.

Keywords: 3D molecule connection, RINGS, checking

received 19 December 1984

With the rapid development of digital computers and graphics facilities and their decreasing prices during recent years, computerized 3D molecular modelling has become an extremely useful tool in many branches of chemistry. The method can provide ready access to knowledge about 3D molecular structure. This is of great importance for our understanding of structure-reactivity and structure-biological activity relationships, in conformational analysis, and in the interpretation of spectroscopic data. The requisite atomic coordinates for molecular modelling are often extracted from collections of X-ray crystallographic data. However, such data are of course not always available for the molecules(s) of interest, and it is desirable to be able to easily create a specific conformer or stereoisomer, or an unorthodox structure as a transition-state model.

Programs capable of calculating Cartesian coordinates from internal coordinates for the atoms in a molecule are available from the Quantum Chemistry Program Exchange (QCPE)[1]. However, most of these programs were written by theoretical chemists in the early days of computational chemistry, and consequently they are mostly batch-oriented and have no graphics interface. In addition, detailed knowledge of bond lengths, bond angles and torsional angles is often required as input. For 'normal-sized' molecules of interest to organic chemists, such as natural products and/or relatively complex ring systems, using such programs is quite tedious and unattractive.

An interactive program that accepts 2D graphical input and then employs a crude forcefield calculation to obtain a 3D molecular geometry, was originally developed at Princeton University in the early 1970s, and was subsequently incorporated into the synthetic planning program SECS[2,3]. This approach is also used in the modelling system MOLY[4], and an improved modification of this technique has been achieved in SCRIPT[5]. For relatively simple molecules, the 2D to 3D conversion constructs reasonable 3D models, but as the size and complexity of the structure increase, it becomes increasingly difficult to predict the exact stereochemical outcome of the minimization procedure.

A useful approach which has been incorporated into a number of interactive modelling systems involves the use of the Cambridge Crystallographic Data Base[6–8] as a library of 3D molecular structures (or 'structural fragments'), which can be used as the basis for a desired model. The Merck Molecular Modelling System (MMMS)[9] and CAMSEQ-II/CHEMLAB[10–12] are examples of systems that use this approach. Both MMMS and CHEMLAB allow the user to start with a structural fragment, whether obtained from the Cambridge Data Base or elsewhere, and to generate a structure in the form of 'internal coordinates' (i.e. bond lengths, bond angles, and torsional angles). In addition, CHEMLAB offers the option of joining two molecules from a pre-generated library along one bond at any user-defined position. To construct a polycyclic system with a stereochemistry for which suitable fragments are not already present in the 'fragment library' is still not straightforward with either of these systems, however.

Other molecular modelling systems of interest here are CHEMGRAF[13] and MMS-X/SYBYL[14,15], both of which start with a prestored set of 3D molecular fragments, which can be joined together. With CHEMGRAF the resulting structure can be modified by deletion or addition of atoms and other fragments, making or breaking bonds, as well as by changing bond lengths, bond angles or torsional angles. It constructs rings by first building the corresponding alicyclic chain, then deleting two atoms at either end of the chain, and finally closing the ring. A useful ring-building option in the SYBYL system allows the user to fuse two molecular fragments along one bond, which permits the facile construction of many complex ring systems. However, fusion along more than one bond, bridge-building, and spiro fusion are options that to our knowledge are lacking in current modelling systems.

BACKGROUND FOR RINGS

The development of a program system called MIMIC (Methods for Interactive Modelling in Chemistry) was begun in 1979 at the Molecular Graphics Laboratory for Organic Chemistry at the Chemical Center, University of Lund/Lund Institute of Technology. The original purpose of this development was to create an interactive graphics interface for work with Allinger's molecular mechanics programs, MM2/MMP2[16–18], which would allow the user to maintain full control of the 3D geometry of the desired structure throughout the building process. A detailed description of the system's molecule builder (MOLBUILD) from the user's point of view has recently been published[19]. The construction of ring systems with MOLBUIILD is accomplished as described above for CHEMGRAF by an independently developed algorithm. One important limitation of this method is that although it is suitable for relatively simple ring systems, the creation of complex polycyclic skeletons (e.g. for work with natural products) places considerable demands on the time and patience of the novice or occasional user. As one of the cornerstones of the philosophy behind our Molecular Graphics Laboratory is that non-computer-oriented chemists should be able to use our computer graphics tools with a modicum of training and practice, it was of great interest for us to improve and facilitate the construction of complex ring systems. Thus, the main purpose of the present work was to create a general ring-building program to produce hydrocarbon skeletons, which could then be 'fitted' with functional groups and/or heteroatoms by means of MOLBUILD operations. The use of energy-minimized 'building blocks' can also reduce the computing time required by the molecular mechanics programs.

BASIC CONCEPT

The program RINGS follows the main philosophy of MOLBUILD: it is 3D from the very beginning and offers the user full control of all ring junctions at every step of construction. For building ring systems, this is accomplished by using prestored structures of appropriate conformations, which can be combined together by means of various joining and fusion operations. These structures may be derived from X-ray coordinates, from energy minimization, by construction on the diamond lattice, or from any other source of molecular geometries that the user wishes to consult.

The stereochemistry of ring junctions is established by indicating the direction(s) of the new bond(s). The operation of joining along one bond has already been shown to be very powerful in various modelling systems, as indicated above. A second necessary operation is the fusion of small rings alone one bond, since many organic and biochemical polycyclic molecules can be dissected into a number of smaller rings. For more complex systems, such as bridged or spiro-fused polycyclic skeletons, special operations are needed. It is possible to construct most bicyclic and bridged ring systems by joining two suitable molecules along more than one bond. A special option for spiro fusion is also necessary. Finally, it was found that a routine for building one- or two-atom bridges is quite useful in practice, although in principle not indispensable.

Ring fusion

All the processes of joining and fusion made available in the RINGS program are based on essentially the same fundamental steps, which are illustrated in Figure 1 for the case of fusing two structures along one bond.

In the first step, the two desired structures are chosen from a menu of molecules and are displayed (Figure 1(a)). The user then indicates the bonds to be fused by specifying appropriate atoms (Figure 1(b)) which are to be superimposed. Next, the user indicates the two H-atoms in the first molecule that should be deleted to indicate the direction of the new bonds (Figure 1(c)).

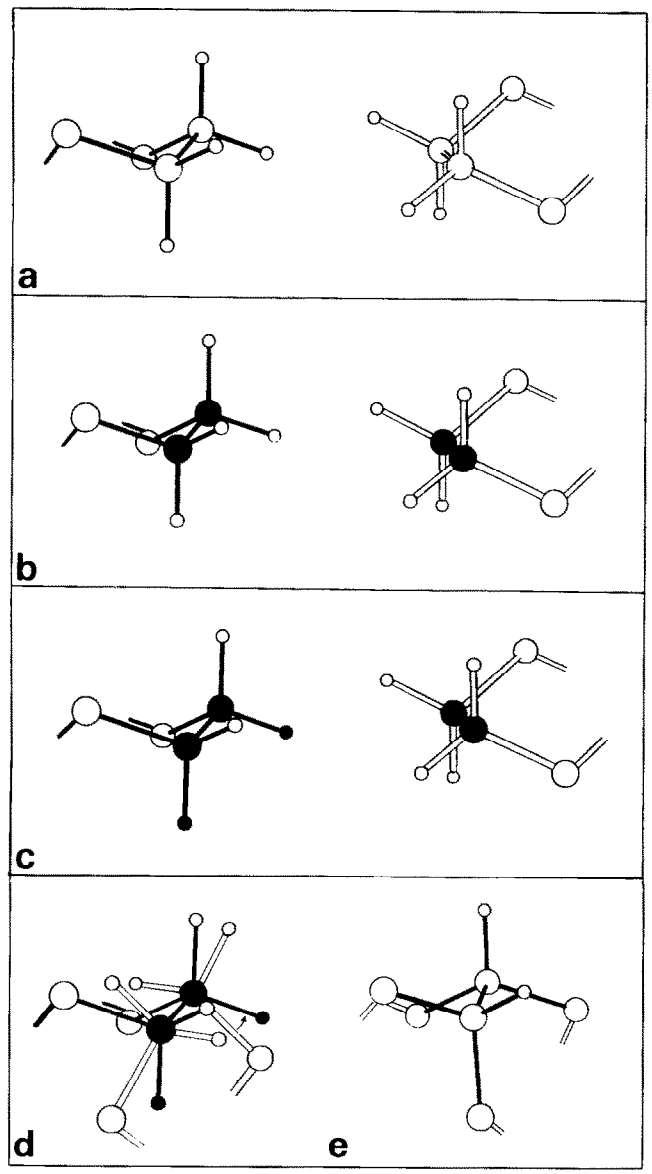

Figure 1. The Ring Fusion Concept (only the relevant parts of the rings are drawn). Figure 1(a) user input: choose two molecules from library; (b) user input: choose atoms defining the bonds to be fused (filled circles indicate user input); (c) user input: create 'free valences' (see text) to indicate the directions of the new bonds; (d) program: puts the indicated bonds together, adjusts the 'free valences' of the first molecules to the new bond in the second molecule, deletes superfluous atoms and renumbers the molecule; and (e) program: displays the fused ring system

This operation will be called 'creation of free valences' (for the new bonds) in the following text.

The program then translates and rotates the second molecule so as to superimpose the previously chosen atoms. Finally, the second structure is swung about the bond of fusion so as to minimize the difference between the directions of the 'free valences' and the appropriate C–C bonds of the second molecule (Figure 1(d)). After the second structure is properly oriented, the program deletes the superfluous atoms, new bonds are entered in the connection table, the atoms of the molecule are renumbered, and the results are displayed (Figure 1(e))[20].

Depending upon the specified H-atoms, the stereochemistry of the resulting fusion may be established as either *cis* or *trans*. The construction of *cis*- and *trans*-decalin by this method is shown in Figure 2, which is reproduced directly from the terminal screen.

AVAILABLE OPTIONS FOR JOINING AND FUSION

Joining along one bond

This option can be used for the joining of any two molecules and is thus not restricted to the joining of rings. The bonds to be joined are specified by indicating one H-atom in each structure, and the program then finds the attached C-atoms. The appropriate

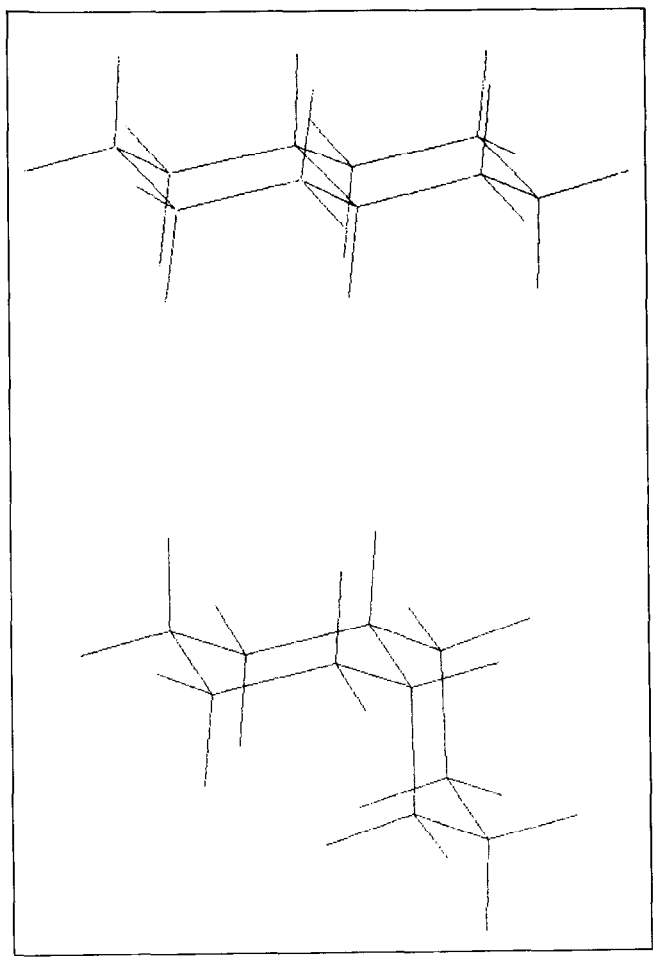

Figure 2. Construction of cis- *and* trans-*decalin by fusing two chair cyclohexane rings along one bond. The pictures are taken directly from the screen*

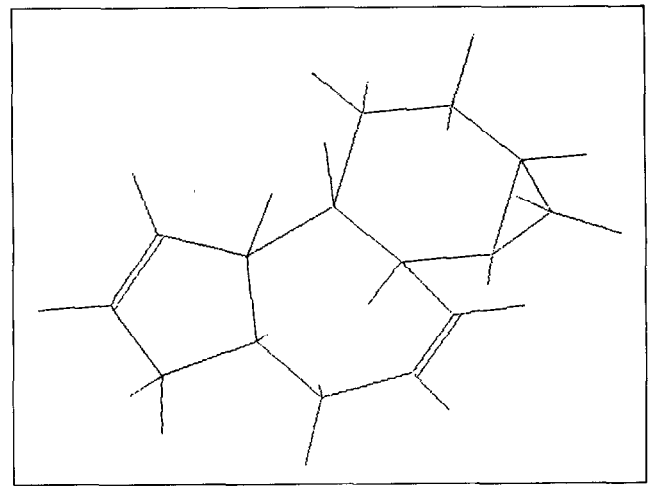

Figure 3. Fusion along one bond. The diterpene skeleton shown was constructed in four steps using the standard small ring molecule set and fusion along one bond (four times)

carbon atom in molecule 1 and the specified hydrogen atom in molecule 2 are superimposed, as are the hydrogen atom in molecule 1 and the appropriate carbon atom in molecule 2. The bond length of the resulting connection is adjusted by look-up in a bond-length matrix. If the new bond is between two sp^3 carbons, dihedral angles are adjusted to achieve a staggered conformation. If two sp^2 atoms are joined, the planarity of the pi-electron system is maintained. The conformation may be further adjusted at the discretion of the user; the second molecule may be rotated about the new bond by any desired angle.

Fusion of two molecules along one bond

One of the available default molecule sets contains energy-minimized conformations for 3-, 4-, 5-, 6- and 7-membered rings. The use of this structure set and the 'fusion along one bond' option allows the construction of a variety of polycyclic natural product classes, such as steroids, terpenes, alkaloids, etc, with full control of the ring fusion geometry. For 'normal-sized' molecules (less than 100 atoms), the construction of the ring system may be accomplished in less than five minutes of the user's time.

An example of the use of this option is the construction of a tetracyclic diterpene skeleton shown in Figure 3 (ring system of phorbol[21]), which was easily accomplished in four steps. Difficulties which may arise in achieving a 'perfect' adjustment of one molecule to the other in the final step of a bond-to-bond fusion are resolved as described in detail in the Appendix (see below).

If the atoms in the bonds to be fused are sp^2 hybridized, it is not necessary to ask the user to create free valences, since there is only one possibility.

One limitation of the fusion option is that if both chosen atoms in molecule 1 are sp^2 hybridized, both C atoms to be fused in molecule 2 must also be sp^2, since both the bond lengths and bond angles of sp^2 and sp^3 systems are of course quite different. However, the fusion of a single bond next to a double bond is possible, as for example in 1-bicyclo[4.4.0]decene or testosterone.

FUSION OF MORE THAN ONE BOND

In many polycyclic and bridged systems more than two carbon atoms may belong to more than one ring. Such systems can be readily constructed by the simultaneous fusion of two, three or four bonds. Only the fusion of structures with nearly equal bond geometries at the bonds to be fused is considered reasonable. After the program has superimposed the appropriate atoms, the distance between the first and last atoms defining the bonds of fusion is calculated. If this distance is greater than 0.2 Å, the program does not attempt the requested fusion and outputs an error message. Otherwise, the fitting of the 'free valences' and the C–C bonds is done as described above for the fusion of one bond.

As a first step in a two-step procedure for building the adamantane skeleton, two chair cyclohexane rings are fused along two bonds (see Figure 4). The final step is to make a bridge, which will be described as the next program option.

Bridge building

RINGS only allows the construction of bridges with one or two sp^3 carbon atoms, since it should be possible to create most simple bridged hydrocarbon skeletons by fusing several bonds simultaneously. In some cases, however, and especially when larger skeletons are being constructed, it is quite convenient to use the bridge-building routines.

The bridge-building option differs from all of the other program options in that new bonds and atoms are created and added to the chosen molecule. After the user has specified the number of bridge atoms (1 or 2), two hydrogen atoms must be chosen for deletion to make free valences for the new bonds. The program then adds a carbon atom along the first chosen C-H bond, and the distance to the second specified carbon atom is calculated. If this distance is about the length of a C–C single bond (+ − 0.2 Å), a bond will be made.

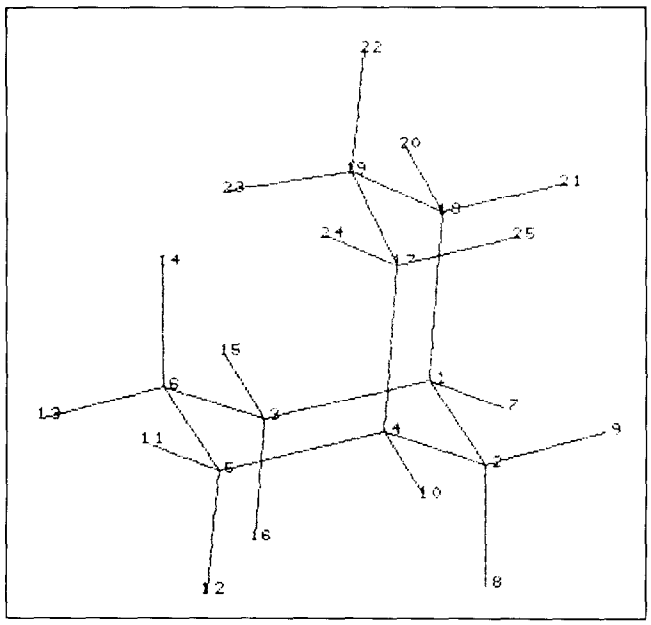

Figure 4. Fusion along more than one bond. Two chair cyclohexane rings are fused along 2 bonds

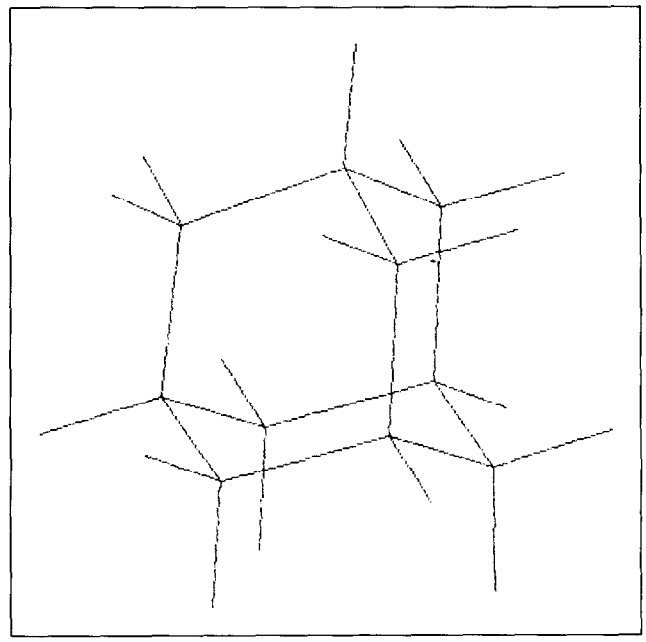

Figure 5(a). Build a one carbon atom bridge. Adamantane is constructed from the molecule of Figure 4 by building a one C-atom bridge

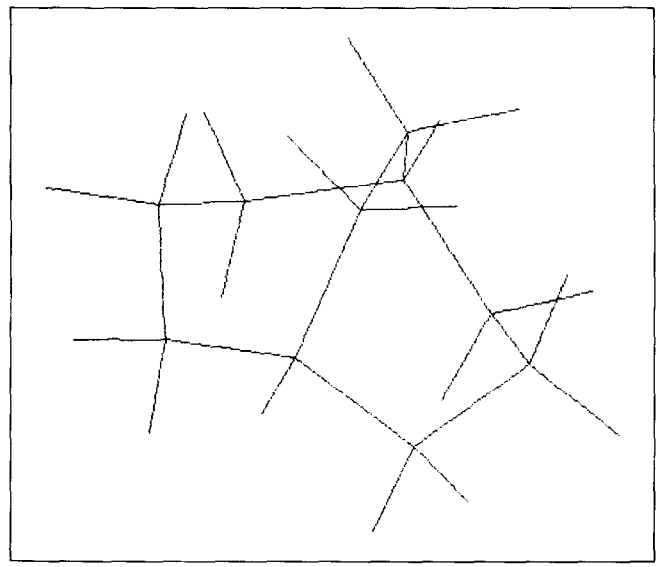

Figure 5(b). Build a two carbon atom bridge. A 2C atom bridge is constructed across the crown conformation of cyclooctane

Otherwise, the new carbon atom is swung to the midpoint between the two indicated ring carbon atoms, maintaining a constant length of the bond to the new atom. This two-step procedure enables the construction of bridges which do not lie on a direct line between two carbon atoms, but which have bond angles close to tetrahedral.

The construction of the adamantane skeleton was completed by making a one-carbon bridge between atoms 14 and 23 in Figure 4 (see Figure 5(a)). In a second example of the use of the bridge-building option, Figure 5(b) shows the result of placing a two-carbon bridge across the crown conformation of cyclooctane.

Spiro fusion

The spiro fusion of any of two sp^3 carbon atoms of any two molecules may be accomplished with this option. After the user has specified two hydrogen atoms in each molecule to create the appropriate free valences at the spiro centre, the program constructs a 'dummy atom' at the midpoint of each specified pair of hydrogens. The carbon atoms to be fused are then placed at the origin, with the 'dummy atoms' on the positive and negative *X*-axes. The actual process of fusion is then similar to that described above for fusion along one bond. An unsubstituted ring system inspired by the structures of agarospirol and hinesol[22] is shown in Figure 6.

Sugar joining

An interest in the possibility of using molecular mechanics calculations as a tool in the conformational analysis of carbohydrates led to the inclusion of a sugar fusion option in RINGS. Sugar fusion is a special case of the fusion of two molecules along one bond.

The molecule names and their numbering are in accord with the abbreviated IUPAC nomenclature for oligosaccharides[23]. Ring C-atoms are specified by the user, and the corresponding O- and H-atoms are found by the program and joined. The two molecules are then mutually oriented, taking into account the exo anomeric effect of the C–O–C bond, and ensuring a staggered position of the remaining H-atom of the second molecule with respect to the O–C bond of the first molecule. On demand, a rotation of the added part of the molecule by any desired angle about the new bond is possible. Figure 7 shows an example of a tetrasaccharide, which was constructed in four steps within a few minutes.

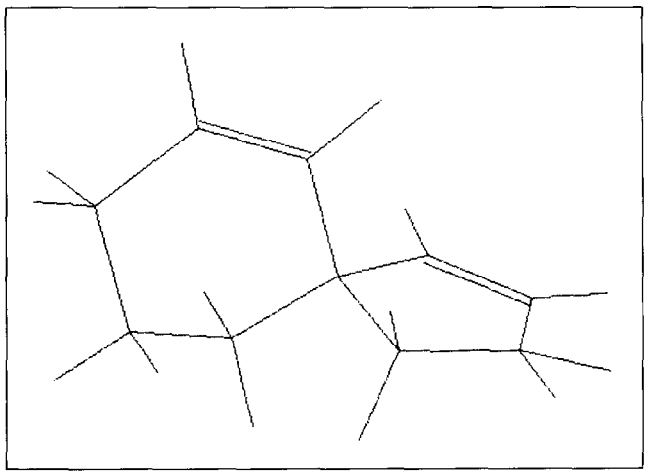

Figure 6. Spiro fusion: the ring skeleton of agarospirol/ hinesol

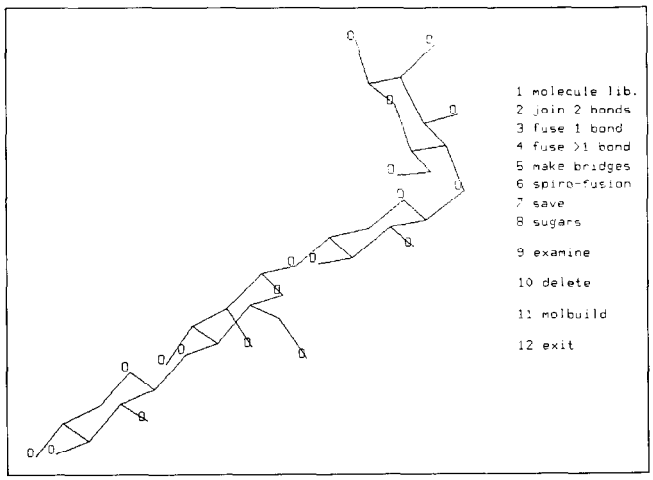

Figure 7. Sugar joining — construction of the tetrasaccharide β-D-Xylp(1-4)-β-D-Manp-(1-4)-β-D-Xylp-(1-4)-β-D-Manp. The main menu of RINGS is shown at the right

DATA MANAGEMENT AND THE MAN–MACHINE INTERFACE

A program to be used by non-computer-oriented chemists should of course be 'user-friendly', and should include routines that check the validity of all user input, to enable a return of control to the user if an input error is found.

To enable the user to escape from a self-created 'error condition', two interrupt options have been added to the program. The first option allows the user to jump back to the main control routine from any input prompt. The second interrupt permits manipulation of the individual molecule (or the nascent ring system) to obtain a more suitable projection of the structure(s) on the screen, and to allow the user to request geometrical information (distances, bond angles, and dihedral angles). All user input is checked for 'reasonableness' and for data type appropriateness. If the proper data type (integer, real, character) is not input for a given prompt, it is rejected and the user is reprompted. During ring construction, the program checks when necessary that user specified carbon atoms are adjacent, that the free valences to be created are connected to the indicated carbon atoms, and so on.

The entire program is menu-driven and is relatively simple to use, so that essentially no further information is necessary to enable the user to start creating ring skeletons. The main menu (see Figure 7) contains the various options for joining or fusion, and for making bridges, as well as options for saving and deletion (see below). In addition, the main menu allows access to submenus for choosing available structures (MOLECULE LIB.) or for manipulating and examining the ring system under construction (EXAMINE; an enhanced MOLDISP option from MIMIC[19]). The MOLBUILD menu item allows access to all of the available MOLBUILD operations[19].

With the SAVE option, the constructed molecule may be stored either as a new data file or as a new molecule in a previously user-defined molecule set. Since the directory for a new molecule set is created in connection with a read of the set, the user always has access to complete information about currently available molecules.

The format of the output is an ASCII file, which is easy to use in connection with other programs (such as MM2/MMP2 and DRACO[24]). The MOLECULE LIB. option displays all currently available molecule sets, and on demand allows the user to load a new one. For practical reasons (available space on the screen and time required to read the molecule file), the number of structures in a molecule set is currently limited to 30.

The upper part of Figure 8 shows the appearance of the screen after the user has selected two molecules for fusion from the 'small-ring library'. The contents of the library are listed on the right of the screen.

The DELETE option offers either the possibility to delete the entire molecule under construction, or the last molecule that was added to the structure. The latter possibility is quite useful if a complex skeleton has already been constructed, and the attempted ring-building operation has given an undesired result.

Hardware and implementation

RINGS has been developed as part of the MIMIC system (Methods for Interactive Modelling In Chemistry) of the Molecular Graphics Laboratory for Organic Chemistry at the Chemical Center, University of Lund/ Lund Institute of Technology. The graphics part of MIMIC includes MOLBUILD[19], the molecular display program MOLDISP, and a program for the visual

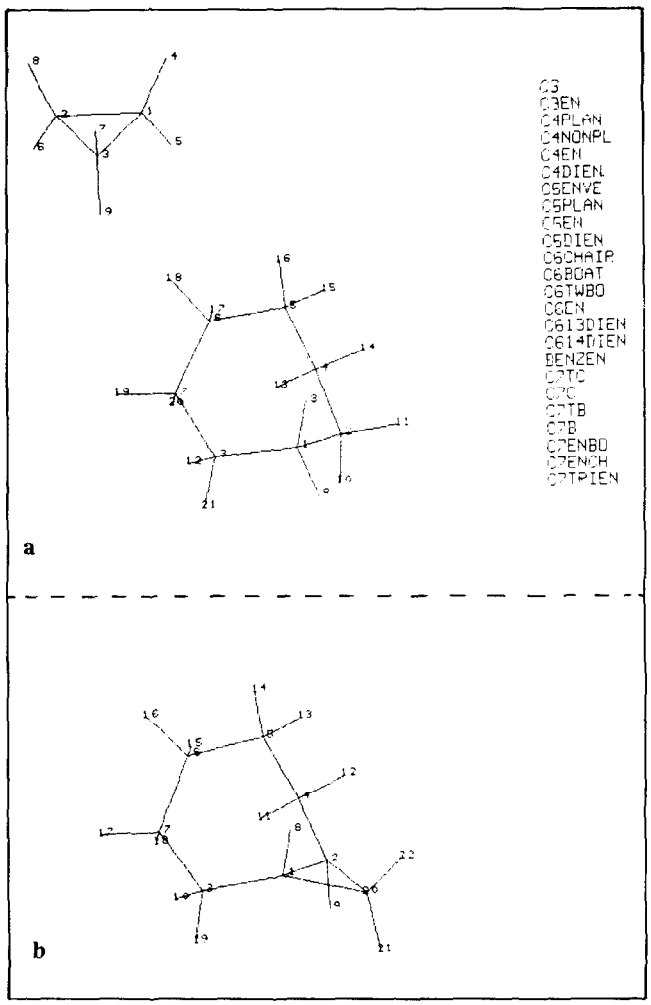

Figure 8(a). Organization of the screen. Molecule 1 is displayed in the middle of the screen, molecule 2 at the top left. The actual menu is shown on the right (menu region). The user input is underlined. In this example the standard molecule set of small rings is displayed in the menu region. The first two characters of the molecule abbreviations indicate the ring size, followed by obvious designations of configuration and conformation. Figure 8(b). Display of the molecule constructed from Figure 8(a)

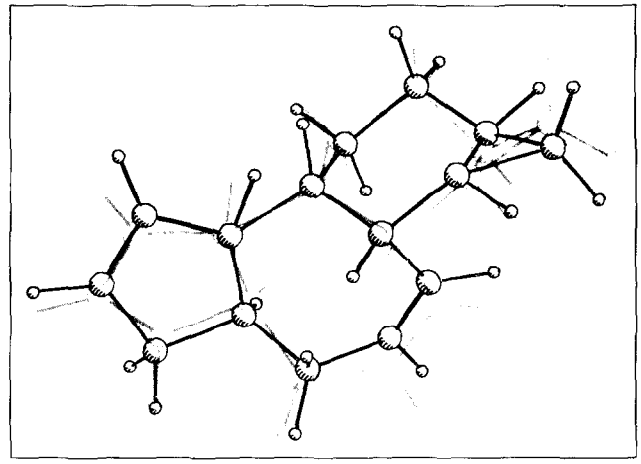

Figure 9. Energy minimization of the diterpene skeleton in Figure 4

and/or least-squares comparison of two structures (MOLCOMP), and it runs on an IMLAC Series II Interactive Graphics Terminal and on a PLOT-10 compatible graphics terminal (VISUAL 500 series[25]). The original version of MIMIC[19] runs on a DEC GT40 series graphics terminal. RINGS is written in FORTRAN-77 on a VAX 11/780, and runs on the VISUAL terminals. The keyboard is the only necessary 'input device'. A version of RINGS that runs on the IMLAC, with lightpen or graphics tablet input, is currently under development.

CONCLUSION

RINGS, together with MOLBUILD[19], provides a very flexible 'front-end' for molecular mechanics calculations, or indeed for any type of calculation that requires 3D molecular coordinates and connectivity as input. The use of previously minimized molecules to create the desired ring skeleton may also reduce the computing time necessary for minimization of the final structure, since the input molecule should already be close to an energy minimum. This is illustrated in Figure 9 with a display from MOLCOMP for the diterpene (phorbol) skeleton from Figure 4.

Since RINGS/MOLBUILD can either use predefined molecules or start 'from scratch' (i.e. from a fragment as small as a C–C bond), the program system can in principle be used to create almost any type of molecular structure.

ACKNOWLEDGEMENTS

C W von der Lieth would like to thank the German Cancer Research Center, Heidelberg, for providing funds for a stay in Lund, during which RINGS was written. We are indebted to Dr Thomas Olsson, Chalmers University of Technology, Goteborg, for providing us with useful graphics subroutines. We wish to thank Mr Anders Sundin and Mr Peter Fransson for creating the sugar database. Financial support for the operation of the Molecular Graphics Laboratory in Lund is provided by the National Swedish Board for Technical Development.

REFERENCES

1 **QCPE** Indiana University, Chemistry Department Bloomington, IN, USA
2 **Wipke, W T** in **Wipke, W T, Heller, S R, Feldman, R J, Hyde, E, (eds)** *Computer representation and manipulation of chemical information* Wiley, USA (1974) pp 147–174
3 **Wipke, W T, Braun, H, Smith, G, Choplin, F and Sieber, W** in **Wipke, W T and Howe, W J (eds)** *Computer assisted organic synthesis* ACS Symposium Series 61 (1977) pp 97–127
4 **Dyott, T M, Stuper, A J and Zander, G S** *J. Chem. Inf. Comput. Sci.* No 20 (1980) pp 28–35
5 **Cohen, N C, Colin, P and Lemoine, G** *Tetrahedron* No 37 (1991) pp 1711–1721
6 **Allen, F H, Kennard, O, Motherwell, W D S and Watson, D G** *J. Chem. Doc.* No 13 (1973), p 119
7 **Allen, F H, and Kennard, O** *Perspectives in computing* IBM No 3 (1983) pp 28–43
8 **Allen, F H, Bellard, S, Brice, M D, Cartwright, B A, Doubleday, A, Higgs, H, Hummelink, T, Hummelink, B G, Kennard, O, Motherwell, W D S, Rodgers, J R and Watson, D G** *Acta Crystallogr.* B35 (1979) p 2331
9 **Gund, P, Andose, J D, Rhodes, J B and Smith, G M** *Science* No 208 (1980) pp 1425–1431
10 **Potenzone, R P Jr, Cavicchi, E, Weintraub, H J R and Hopfinger, A J** *Comput. Chem.* No 1 (1977) p 187
11 **Potenzone, R P Jr** *Ph.D. Thesis* Department of Macromolecular Science, Case Western Reverse University, Cleveland, OH, USA (1978)
12 The CHEMLAB program was originally developed by A J Hopfinger and coworkers at Case Western Reserve. It is now available from Molecular Design Limited, 1122 B Street, Hayward, CA 94541, USA
13 The CHEMGRAF program is available from The Lodge, Botley Works, Oxford OX2 0NN, UK
14 **Humblet, C, Marshall, G R** *Drug Devel. Res.* No 1 (1981) pp 409–434
15 The SYBYL program was developed by G R Marshall and coworkers. It is now available from TRIPOS Associates Inc, 6548 Clayton Road, St Louis MS 63117, USA
16 **Allinger, N L, Eliel, E L and Wilen, S H (eds)** *Topics Stereochem.* John Wiley, USA, Vol 13 (1982) pp 117–193
17 **Allinger, N L** *Adv. Phys. Org. Chem.* Vol 13 (1976) pp 1–82
18 **Burkert, U, Allinger, N L** *Molecular mechanics* ACS Monograph 117, Washington DC, USA (1982)
19 **Liljefors, T** *J. Mol. Graphics* Vol 1 No 4 (1983) pp 111–117
20 Appropriate geometrical transformations are described by **Foley, J D and van Dam, A** *Fundamentals of interactive computer graphics* Addison-Wesley (1982) Chapter 7
21 See, for example, **Seip, E H, Ott, H H and Hecker, E** *Planta medica* Vol 49 (1983) pp 199–203
22 **Mongrain, M, Lafontaine, J, Belanger, A** *Can. J. Chem.* Vol 48 (1970) pp 3273–3274
23 **IUB-IUPAC Joint Commission on Biochemical Nomenclature** *Pure Appl. Chem.* Vol 54 (1982) pp 1517–1526
24 **Marsili, M, Floersheim P, and Dreiding A S** *Comput. Chem.* Vol 7 (1983) pp 175–181
25 Visual Technology Inc, 540 Main Street, Tewksbury, MA, USA

APPENDIX A

In connection with the bond-to-bond fusion operation, it is necessary to find the most suitable orientation of the second molecule with respect to the chosen 'free valences' of the first one. In the initial step of this procedure, the first chosen 'free valence' and the first indicated C–C bond of the second molecule are superimposed (Figure 10, top), and the residual angle (θ_1) between the second 'free valence' and the second new C–C bond is calculated. If θ_1 is less than 5 degrees, the orientation is considered to be 'perfect. If not, the entire procedure is redone for the angle θ_2 of the second 'free valence' and the second C–C bond (Figure 10, bottom). If θ_2 is less than 5 degrees, the program proceeds with the requested fusion. If a 'perfect' adjustment of one molecule to the other can not be made, we have found it useful to implement two different options:

- 1 if both angles between the free valences of the first molecule and the appropriate bonds of the second one is less than 30 degrees, the possibility which gives the smallest angular deviation is chosen.
- 2 If any such angle is greater than 30 degrees, the user is warned, 'This fusion gives unusual bond angles!' and is asked to confirm that the fusion is desired. This may indeed be the case if the user wishes to construct ring systems with small rings. When 'normal' rings are involved, this message may indicate that unsuitable free valences have been generated. If the question is answered in the affirmative, the program constructs the fused molecule by placing the bond of the second molecule at the bisection of the angle between the free valences. Otherwise, the user has to specify the free valences again.

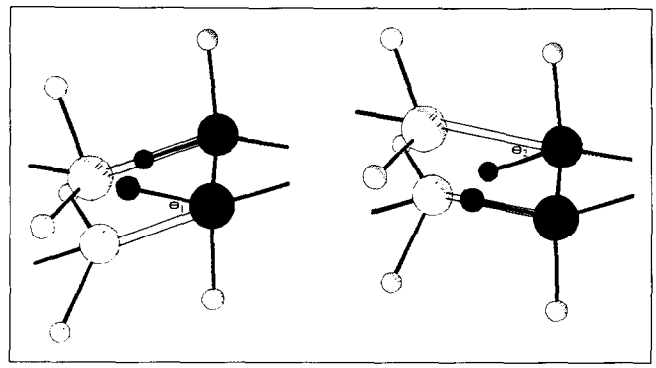

Figure 10. Steps in the bond-to-bond fusion operation