MOLGEN

Version 3.0

Our philosophy: *"To provide you with the highest quality molecular modelling software available at a fair and reasonable price. To listen to your suggestions, and to add improvements on a continuing basis. Our customers are our most valuable asset."*

If you use and reference Molgen in a technical article, you may qualify for a FREE upgrade or program add-in. Mail a copy or reprint of your article to J. Eric Slone (address on following page).

Molgen is Copyright © 1993 - 1995 by Peter Baricic, Martin Mackov, and J. Eric Slone, All Rights Reserved. The Molgen software and User's Manual are protected under U.S. and International Copyright Law.

Molgen User's Manual translated from Slovak by Milan Hudecek. Maintained by Martin Mackov and J. Eric Slone.

REGISTRATION INFORMATION

In order to qualify for upgrades and program enhancements, you must complete and mail in this product registration.

Serial Number _____

Company _____

Name _____

Address _____

_____

City, State Zip _____, _____ _____

Country _____

Telephone (_____) _____ - _____

Fax Number (_____) _____ - _____

Email _____

Molgen was purchased for use primarily in:

[ ] Business [ ] Education [ ] Personal

Printer (Model/Manufacturer) _____

Computer (Processor/Speed/RAM) _____

Video (Display Driver/Memory/Monitor)_____

Comments _____

_____

_____

_____

_____

Please mail the completed form to:

J. Eric Slone
5500 Holmes Run Parkway, #501
Alexandria, Virginia 22304-2851
United States of America

Molgen is written by:

Peter Baricic
Komenskeho 15
SK-900 01 Modra
Slovakia

Martin Mackov
Nobelova 1/A
SK-831 02 Bratislava
Slovakia

J. Eric Slone
5500 Holmes Run Parkway, #501
Alexandria, Virginia 22304-2851
United States of America

Distributors:

J. Eric Slone
Scientific Consulting Services
5500 Holmes Run Parkway, #501
Alexandria, Virginia 22304-2851
United States of America

Telephone: (703) 461-7078
Fax: (703) 751-6639
Email: eslone@patriot.net

INSTALLATION INSTRUCTIONS

Molgen comes with an installation program which automates the process for you. To begin installation, insert **MOLGEN INSTALLATION DISK 1** into your disk drive (drive A: will be used for the examples, substitute your drive letter (i.e. drive B:) and change to that drive:

```
C:\> A: ↔
A:\> install ↔
```

The first step in the installation process is to enter the SERIAL NUMBER which is provided on the **MOLGEN INSTALLATION DISK 1** and your USER NAME.

You will now select a drive and directory to install the program into. For further information on using these screens, see Section ***17*** in this manual. The program will install itself into the drive and directory that you specify; if the directory does not exist, it will be created.

STARTING MOLGEN

To start *Molgen*, change to the drive and directory where you installed the program, for example, if you installed *Molgen* on drive D: in the subdirectory \MOLGEN:

```
C:\> d:
D:\> cd \molgen
D:\molgen> molgen
```

After successfully entering this information, you will be presented with the *Molgen* title screen. Press any key or click a mouse button to progress to the main screen.

1 BUILDER (Builder of molecules)

The ***BUILDER*** module provides the possibility to generate new chemical structures using either the atomic types implemented in *Molgen*, or the substructures that are supplied with the package or generated in advance by the user. The molecular structures can be modified by changing the atomic types, deleting or adding atoms or fragments, altering bond lengths, bond angles or torsion angles.

1.1 ALTER (alter the geometry of a molecule)

The ***ALTER*** module allows you to change the value of a bond length, bond or torsion angle, or to orient molecules in a given way (two bonds into line, three atoms into a plane, or a bond into the z-axis).

1.1.1 BOND LENGTH (alter a bond length)

The ***BOND LENGTH*** module allows you to alter a bond length or an intermolecular distance defined by marked atoms. To alter an interatomic distance, select two atoms. After the second atom has been clicked, the interatomic distance is displayed. If no change in the distance is required, press ESC or ENTER. Otherwise type in the new value. If another change is needed, mark another pair of atoms, otherwise, click on the function ***EXIT (16.1)***. It is not possible to alter a bond length within a ring.

1.1.2 BOND ANGLE (alter a bond angle)

To alter a bond angle, mark the three atoms defining the angle to be altered. After the third atom has been selected, the value of the angle is displayed on the screen. Enter the new value for the angle or press ESC or ENTER if no change is required. The selected atoms are marked until the first atom of a new set is selected, or the function ***EXIT (16.1)*** is clicked, or another function of the module ***ALTER (1.1)*** is called. The marked atoms must form a bond angle; i.e. the atom marked as second in the sequence must be bonded to the first and the third one. It is not possible to alter a bond angle within a ring.

1.1.3 TORSION ANGLE (alter a torsion angle)

To alter a torsion angle, select four atoms of a molecule. After the fourth atom has been marked, the value of the angle is displayed. To alter it, type in a new value or press ESC or ENTER if no change is required. The selected atoms are marked until the first atom of a new set is selected, or the function ***EXIT (16.1)*** is clicked, or another function of the module ***ALTER (1.1)*** is called. It is not

possible to alter a torsion angle in the case when all four atoms are part of the same ring and when three atoms are part of a ring and the

fourth one is part of another cycle that is connected to the first ring. The torsion angle can be altered in condensed rings in special cases that are described in the example of *phenanthrene*:

It is possible to alter torsion angles defined by the following atoms:

15-1-2-16 1-2-3-17

12-13-4-3 8-7-12-13

It is not possible to alter torsion angles defined by the following atoms:

11-12-13-14 7-12-13-14

1-2-3-4

1.1.4 COLLINEAR BONDS (bonds into line)

This function allows you to orient two molecules so that a pair of marked atoms from a molecule lies in a line with a pair of marked atoms of the other molecule. The distance between the second and the third marked atom is set to the default value of 3 Ångströms. The value can be then altered using the module ***BOND LENGTH (1.1.1)***.

1.1.5 INTO XY PLANE (atoms into xy plane)

A molecule can be oriented so that the plane defined by three marked atoms, which do not lie in a line, is identical with the plane of the screen, i.e. the z-coordinates of marked atoms are zero.

1.1.6 INTO Z-AXIS (bond into z-axis)

After two atoms have been selected, the molecule is oriented so that the joint of the two atoms lies in the direction perpendicular to the plane of the screen (in the direction of z-axis). The atom, which has been selected first, is the farther one.

1.1.7 MOVE ALONG Z-AXIS

This function allows to shift molecules in the direction of z-axis. If a negative value of the shift is entered, the molecule "submerges" to the screen, and *vice versa*.

1.2 MONITOR (rotation about a bond)

This module allows you to modify structures by rotation about up to eight selected rotation bonds with a given increment of rotation. At the same time, it is possible to follow up to eight interatomic distances during rotations.

1.2.1 ROTATION BOND (select a rotation bond)

Below the title "ROTATION BOND", there are eight rectangles of different colors displayed. After any of these has been clicked, the user is prompted to mark first atom of the pair forming the rotation bond. After the second atom is clicked, a test is run to check whether the pair of atoms is connected by a bond. If not, a green or white dotted line is displayed between the atoms. The color depends on whether the function ***ROTATION IN CYCLE (1.2.4)*** has been selected (green) or not (white). In the middle of the joint, the letter 'R' is displayed in the color of the corresponding rectangle. Clicking this 'R' will cause the rotation with the increment pre-defined using the function ***STEP (1.2.3)***. If the joint is white in color, the conformation of the structure remains unchanged but the structure rotates about the selected joint as a whole. If a green dotted joint has been selected, the rotation alters the molecule's conformation. In this way it is possible to alter, for instance, the boat conformation of a cyclohexane skeleton into the chair conformation and *vice versa*. If the selected joint is not a dotted line, then it is actually a bond in the molecule, and two cases may occur:

- The bond is part of a ring and then the whole molecule rotates about the bond, i.e. no change of its conformation occurs
- The bond is not part of any ring and then torsion angles as well as the conformation is altered

Clicking the other mouse button alters the sense of the rotation. Up to eight rotation bonds may be selected at a time. If the same rectangle is selected again, the following selection of an atom will remove the 'R' mark of the corresponding color from the original pair, and after the second atom is selected, the 'R' is placed at the new pair. Providing that the function ***ROTATION IN CYCLE (1.2.4)*** has not been selected, the value of the selected bond's torsion angle is displayed in the upper left corner of the screen. The angle is determined by four atoms where the second and the third one form the selected bond, the first one is the atom with the lowest sequential number of those connected to the second one, and the fourth atom is the one with the lowest sequential number connected to the third atom. To view the atomic sequential numbers, toggle the ***LABELS (16.13)*** switch to the ***NUMBER*** mode.

1.2.2 DISTANCE (select the distance to be monitored)

Below the title "DISTANCE" there are eight rectangles of different colors displayed. After any of these has been clicked, the user is

prompted to mark first atom of the pair of atoms, the distance of which is to be monitored. The selected pair is marked by '@' mark of the selected rectangle's color. If the same rectangle is selected again, the

following selection of an atom will remove the '@' marks of the corresponding color from the original pair, and after the second atom has been selected, the '@' marks of the corresponding color are placed at the new pair. After each selection of pair of atoms, the value of their distance is displayed in the color identical with the color of the corresponding rectangle and '@' mark.

1.2.3 STEP (the increment of rotation)

Below the title "STEP", the available values of the rotation's increment displayed. To select a value, click on it.

1.2.4 ROTATION IN CYCLE

This function allows you to specify a fictive rotation bond in a ring between atoms not connected by a bond, and to modify the structure by rotation about this bond. After this function has been selected and a color field has been clicked, the user is prompted to mark three atoms. The first two atoms define the rotation bond and the third one determines the part of the molecule to be rotated.

1.2.5 UNSELECT ALL

This function removes the markings from all of the atoms.

1.3 ATOM ADD

This function allows you to create new structures by adding atoms from the selected atomic types, or to modify already created structures. After this function has been selected, atomic symbols of available atoms are displayed. The ones marked with an asterisk '*' are those which are available in more types according to the bonding states in which they occur in compounds. After clicking an atomic type, an auxiliary window appears with the structure of the selected atomic type. Each of the types is assigned standard bond angles and types according to the corresponding electronic state. For example, the C3 carbon type has four single bonds pointing to the corners of regular tetrahedron which is characteristic for the sp^3 hybridization.

1.4 ATOM DEL (delete an atom)

This function allows you to remove atoms or unfilled valences. To delete an atom, select the function and then click the atom to be removed. The atom is removed together with all bonds emanating from it. At the atoms to which

the deleted atom were attached, unfilled valences appear marked with red tilde '~'. The orientation of the added valences are identical with the direction of the deleted bonds. Any univalent atoms connected to the deleted atom are removed as well. This function can delete also unfilled valences and lone electron pairs.

1.5 FRAGMENT PATH (specify the path for fragments)

This function defines the path to the subdirectory which contains fragments. A more detailed description can be found in Section ***17.1***.

1.6 FRAGMENT LOAD

This function allows you to read a fragment from disk. After this function has been selected, all files with the extension FRG from the selected subdirectory are displayed in window. These are text files of the standard *Molgen* format containing molecular structures with at least one unfilled valence. After any of the filenames displayed is clicked, the corresponding structure appears in the auxiliary window. A more detailed description is given in Section ***17.2***.

1.7 FRAGMENT SAVE

This function saves the selected structure in the main window to disk. The structure must have at least one unfilled valence. If there is more than one structure in the main window, it is necessary to select one. After the structure has been specified, the user is prompted to enter a filename of up to eight characters. The name of the file to which the structure is written, is appended with the extension FRG.

1.8 GROUP PATH

This function allows to specify the path to the subdirectory containing groups. A more detailed description is given in Section ***17.1***.

1.9 GROUP LOAD

This function allows you to load functional groups for the generation of new structures. After this function has been selected, all files with the extension GRP from the selected subdirectory are displayed. These are text files in the standard *Molgen* format containing molecular structures with at least one free valence. The difference between FRG and GRP files consists only of the fact that in general GRP files contain smaller, more frequently used functional groups. After any of the filenames have been clicked, the corresponding

structure appears in the auxiliary window. A detailed description can be found in Section ***17.2***.

1.10 GROUP SAVE

This function allows you to save the selected structure from the main window to disk. The structure must have at least one unfilled valence. If there is more than one structure in the main window, it is necessary to select one. After the structure has been specified, the user is prompted for a filename of up to eight characters. The name of the file to which the structure is written, is appended with the extension GRP.

1.11 BOND CREATE

This function allows you to form new bonds. To create a bond, select two unfilled valences that do not come from atoms already connected by a bond. In order to have the operation successfully completed, both valences must be of the same type and both atoms must belong to the same molecule or disjoint molecule. A disjoint molecule can be created from two or more structures using the function ***M-UNITE (16.15)***.

1.12 BOND BREAK

This function allows you to remove bonds. After a pair of bonded atoms has been marked, the bond is replaced by an unfilled valence marked with a red tilde '~' on each of the atoms. It is possible to attach an atomic type, fragment, or group, and the structure thus created can be saved to disk as fragment or group .

1.13 VALENCES ADD (add unfilled valences)

By clicking this function and a structure in the main window, free valences are generated on the selected structure. Unfilled valences are added to atoms where the displayed number of bonds is lower than the number corresponding to their electronic state. For example, the C3 atomic type is sp^3 hybridized carbon with four single bonds with tetrahedral angles of approximately 109.2°. If some of the four valences are missing, for example after deleting hydrogens, this function will add the unfilled valences. The geometry of the added valences corresponds to the atomic type.

1.14 CHANGE OF TYPE (change of atomic type)

This function allows you to modify the molecular structures by the exchange of atomic types. After the atom to be replaced is clicked, the user is prompted

to choose the required type of atom from the list of atoms. If the atomic types are interchangeable, the operation will be completed. Otherwise the user is notified that the atomic types are not interchangeable. An atomic type can be replaced with a type which meets the following conditions:

- An equal number, multiplicity and geometry of valences with the selected atomic type, for instance:
- C3 type is interchangeable with types Si, N3+
- O3 type is interchangeable with S3

Note: The C3 type is interchangeable with the N3 type only if there are unfilled valences on the C3 atom.

- The atom to be replaced has more valences, these being of the same type as the marked atom has, for example:
- The N3 type is interchangeable with the types C3, Si
- The O3 type is interchangeable with the types N3, C3, Si

Note: In this case, the unfilled valences on the new atoms will be missing. These can be added using the function ***VALENCES ADD (1.13)****.*

1.15 OPTIMIZATION (optimization of molecule's geometry)

This function provides for the optimization of the structure of a designed molecule. For the optimization, the molecular mechanics is used which employs the potentials from the molecular mechanics MM2. The influence of hydrogen bonding is included in the calculations. The electrostatic interactions are approximated by the dipole-dipole interaction. If there is more than one molecule displayed in the main window, the user is prompted to select the molecule to be optimized. Then the molecule is automatically assigned atomic types occurring in the molecular mechanics used. If all the types are defined, the calculation is commenced. If any of the atomic types is not known for the mechanics used, the optimization is not possible. The changes of structure are displayed during the optimization. The calculation can be interrupted by pressing ESC. The speed of the optimization depends considerably on the calculation method used and on the presence of a coprocessor which markedly speeds the calculation up. It is recommended that the generated molecule is built from its natural units (rings, chains, etc.) which have been optimized separately and stored to disk as fragments and only then built in the final molecule. It is also advisable to use the possibility to rotate an attached fragment about the new bond to a reasonable position. A molecule generated in this way provides a reliable and time-saving start for

the optimization. When generating new molecules it is useful to apply the molecular fragments supplied with the *Database Of Molecules*, the contents of which is designed to cover as large a spectrum of types of molecular structures as possible.

1.16 CHIRALITY

In molecular modeling, it is extremely important to distinguish between optical isomers. Frequently the isomers may have very different activities. Chiral centers are assigned on the basis of the well-known Cahn, Ingold and Prelog rules. Once the sequence of substituents around the chiral atom has been determined (based on atomic numbers and extended connectivity), a three-dimensional representation of that center is held so that the smallest substituent in the sequence is oriented away from the viewer. If the remaining substituents are arranged clockwise by priority, the center is designated R, otherwise it is designated S.

Molgen determines cis/trans isomerisation on double bonds not included into cycle. The priorities of the substituents are determined by the same rules as for R/S centers. If the two largest substituents are in a position trans to the double bond it is assigned an E, otherwise it is assigned a Z. *Molgen* determines the correct chirality for all tetrahedral and pyramidal atomic types and cis/trans isomerisation for all double bonds not included into cycles. If there are any chiral centers and/or double bonds cis/trans isomery in a structure on the screen, all they will be assigned R or S, and/or E or Z symbols, respectively. A sub-menu containing the commands ***INVERT MOLECULE (1.16.2)*** and ***INVERT ATOM (1.16.1)*** will be displayed and the message "CHIRAL" appears in the window for labeling. If there is no chiral center on the screen, the message "No chiral center" will appear.

1.16.1 INVERT ATOM

This function allows you to invert one chiral center in a molecule on the screen. After clicking any chiral center, the two smallest adjacent branches will exchange position and all other chiral centers are recalculated. *Molgen* can invert all chiral centers S/R, and E/Z not included in cycles. Chiral centers positioned at the spiro atoms can be inverted. The Invert Atoms command remains active until the ***EXIT (16.1)*** command is selected. Occasionally, the inversion of one chiral center may affect the chirality of another center in molecule. This occurs when the specified inversion affects the relative priority of its substituents and is not due to a repositioning of the atoms about the center. In some cases the inversion of one center cancels chirality of another center.

1.16.2 INVERT MOLECULE

This function allows you to invert all R/S chiral centers in a molecule by using a single command. You will be asked to select the molecule to be operated upon by this command. After selection, all chiral centers R/S in the molecule will be converted. This command remains active until the ***EXIT (16.1)*** command is selected. The E/Z isomerism is not affected by this function.

2 *INTERFACES (links with other systems)*

This module contains subroutines securing the transferability of structural and some other data between *Molgen* and selected programs such as the crystallographic database CSD (Cambridge Structural Database System), program packages for molecular graphics like Sybyl, and programs allowing for the input of the structures in free format by means of Cartesian or crystallographic coordinates.

2.1 MOLGEN (the default program)

This function allows you to read or to write to a disk, structures in the native *Molgen* format. Description of the format:

1st line:

number of atoms (N_at)number of bonds (N_b) total charge name of the structure

2 to N_at + 1st line:

sequential num. of atom (i) Molgen type[i] x[i] y[i] z[i] atomic charge[i] user's type[i]

N_at + 2 to (N-at + 2 + N_b)-th line:

sequential number of the bond (i) sequential number of the first atom of the i-th bond
sequential number of the second atom of the i-th bond type of the i-th bond

Molgen type: the description is in the file ATOMDEF.TAB

Bond type: SINGLE DOUBLE TRIPLE AMIDE AROMATIC

An example of this format is given in ***Appendix A.***

2.2 SYBYL (an external program for molecular graphics)

This function allows you to read and write structures in the format of the Sybyl program. Sybyl is an extensive program package, a product of the Tripos Company. It is aimed at complex study of organic and organometallic compounds. The package contains various methods of optimization for molecular complexes (*ab initio*, semiempirical methods, methods of molecular mechanics and molecular dynamics). It provides the possibility to study polymeric compounds, to process NMR spectra et cetera. An example of this format is given in ***Appendix A.***

2.3 CHEMWINDOW

This function allows you to read structures in the format of the ChemWindow program. ChemWindow is a program operating in Microsoft's Windows environment. It is designed for drawing two-dimensional chemical structures. To obtain an output that can be loaded by *Molgen*, it is necessary to

save the file in Molfile format (extension MOL; menu FILE EXPORT MOLFILE). The stereochemistry of this kind of file is accepted in the sense that the atoms that are bonded via a bond expressed as a full prolonged triangle (i.e. the bond sticking out of the drawing plane) are assigned a z-coordinate increased by 1.0, and the atoms bonded via bonds expressed as a dotted line are assigned a z-coordinate diminished by 1.0. The rest of the atoms lie in the plane with z=0.

2.4 CSD (database of optimized structures)

This function allows you to read structures in the CSD output format. For a description of the format, see the manual for the CSD database. CSD is a database containing verified structures of chemical compounds in the form of crystallographic coordinates. Every year it is supplemented with approximately 10,000 new compounds. At the end of 1991, it contained about 110,000 chemical structures. *Molgen* provides automatic transformation of crystallographic coordinates to Cartesian ones and automatic assignment of atomic types used in the program *Molgen*. Atomic types are determined on the basis of bond lengths and identified bond angles. An example of this format is given in ***Appendix A***.

2.5 XYZ FORMAT (Cartesian coordinates)

This function allows you to read and write on disk structures in the XYZ free format in Cartesian coordinates. Description of the format:

1st line:

number of atoms (N)

2-nd to (N+1)-st line:

x[i] y[i] z[i] atomic number [i]

Molgen allows for the automatic assignment of atomic types used in the program *Molgen*. Atomic types are determined from the bond lengths and identified bond angles. An example of the format is given in ***Appendix A***.

2.6 X-RAY FORMAT (Crystallographic coordinates)

This function allows you to read from structures written in the X-Ray free format in crystallographic coordinates. Description of the format:

1st line:

number of atoms (N)

2-nd line:

A, B, C, ALPHA, BETA, GAMMA

3-rd . to (N+2)-nd line:

a[i], b[i], c[i], atomic number [i]

A, B, C, ALPHA, BETA, GAMMA: the crystallographic parameters of the basic cell

a[i], b[i], c[i]: the crystallographic coordinates of atoms

An example of the format is given in ***Appendix A***.

2.7 AMPAC/MOPAC

Ampac and Mopac are an extensive programs packages containing the frequently used quantum chemistry semiempirical methods AM1, MNDO, MINDO/3 and PM3 (only in Mopac). Both systems are distributed by QCPE (Quantum Chemistry Program Exchange). The function ***AMPAC/MOPAC*** allows you to create inputs for quantum chemistry methods included in the Ampac and Mopac system, to read files generated before, and to read the output results.

2.7.1 BOND SHOW

To display the bonds selected for optimization or that of bonds optimized already by Ampac. After this function has been selected, the following bonds are displayed as thick green lines:

- Bonds defined by the module ***BOND CHOICE (2.7.2)***
- Bonds to be optimized, which have been read from the input file for Ampac (Mopac)
- Bonds which have been optimized and which have been read from an output file from Ampac (Mopac).

If this function is active, the text "BOND" is displayed in the window for atomic labels. After you quit the function ***AMPAC/MOPAC***, the information as to which bonds have been optimized is lost.

2.7.2 BOND CHOICE

Designate the bonds to be optimized by methods included in Ampac and Mopac. After this function has been selected, the user is prompted to click on two atoms. After the second atom of the bond forming pair has been clicked, the bond is displayed as a thick green line. The selected bond must not be part of a ring. The maximum number of selected bonds is 100. If the pair is marked one more times, the mark is removed. A bond is designated to be optimized only if the optimization by selection has been chosen by the ***CONFIGURATION (2.7.11)*** module. Otherwise, any marking is ignored. If the full optimization has been selected, then, regardless of which parameters have been chosen by this function, all geometry parameters will be optimized. If the file which is read by Ampac or Mopac, contains optimized bond lengths, these are read as well, and also marked after selection of the appropriate function.

2.7.3 ANGLE SHOW

To display the bond angles selected for optimization or that of bond angles optimized already by Ampac (Mopac). After this function has been selected, the following bond angles are highlighted by thick green lines:

- Bond angles defined by the module ***ANGLE CHOICE (2.7.4)***
- Bond angles to be optimized which have been read from the input file for Ampac (Mopac)
- Bond angles which have been optimized and which have been read from an output file from Ampac (Mopac)

If this function is active the text "ANGLE" is displayed in the window for atomic labels. After you quit the function ***AMPAC/MOPAC***, the information as to which bond angles have been optimized is lost.

2.7.4 ANGLE CHOICE

Used to designate bond angles to be optimized. Three atoms should be marked in the order that they form the bond angle to be optimized. After the third atom has been marked correctly, the bonds which form the bond angle are marked by a thick green line. The bond angles thus selected will be optimized only if the option **[Selected]** in ***CONFIGURATION (2.7.11)*** is chosen when defining the optimization. If the trinity of atoms is selected in the same or the opposite order for the second time, the marking by the thick green line is removed. If full

optimization has been selected, then regardless of which parameters have been chosen by this function, all geometry parameters will be optimized. If the file, which is read by Ampac contains optimized bond angles, those are read as well, and also marked after selection of the appropriate function.

2.7.5 ROTATION SHOW

To display the torsions selected for optimization or that of torsions optimized already by Ampac. After this function has been selected, the following torsions are displayed as thick green lines:

- Torsions defined by the module ***ROTATION CHOICE (2.7.6)***
- Torsions to be optimized which have been read from the input file for Ampac (Mopac)
- Torsions which have been optimized and which have been read from an output file from Ampac (Mopac)

If this function is active, the text "ROTATION" is displayed in the window for atomic labels. After you have quit the function ***AMPAC/MOPAC***, the information as to which torsions have been optimized is lost.

2.7.6 ROTATION CHOICE

Used to mark bonds with free rotation (torsions) for optimization. A pair of atoms forming this bond must be selected. After the second atom has been clicked, the bond is tested for the torsional rotation. If the marked atoms form the bond about which torsional rotations take place, the bond will be highlighted by thick green line. In the input file for Ampac or Mopac), all torsional rotations related to the bond will be included in optimization. The maximum number of selected bonds is 100.

2.7.7 RING SHOW

Used to display the rings included in optimization. After this function has been selected, the rings defined by the module ***RING CHOICE (2.7.8)*** are displayed as thick green lines. If this function is active, the text "RING" is displayed in the window for atomic labels. After you quit the function ***AMPAC/MOPAC***, the information concerning the rings is lost.

2.7.8 RING CHOICE

Used to select rings to be optimized. Particular atoms should be selected in the order which they are built in the cycle. If an atom is selected which is not bonded to the previously selected one, or it is not part of the ring under selection, an error message appears. If the most recently selected atom is clicked again, its mark is removed. After an atom connected to the first selected atom has been clicked, the marks in the cycle are removed and the bonds of the selected cycle are displayed as thick green lines. The maximum number of selected rings is 10, the maximum number of atoms in one cycle being 30. In the produced input file, all geometry parameters (bonds, bond angles, torsions related to the atoms included in the cycle) will be optimized. Upon reading a file with rings selected for optimization, the marking is read for bonds, and bond and torsion angles.

2.7.9 RING UNSELECT

Used to cancel the marking of rings selected for optimization. This function removes marking of all rings at the same time. Before carrying out the command, a message appears to make sure whether or not you really want to cancel the selection.

2.7.10 BOND ORDER

Used to display the value of bond order of a selected bond. The bond order is an important quantity which represents the strength or the stability of a bond. The matrix of bond orders may be included in the output file from Ampac if it has been defined by the keyword "BONDS" in the input file. When generating a input file for Ampac, it is possible to define the printing of bond orders into the output file by using the module ***CONFIGURATION (2.7.11)***. The print option **[Final bond-order matrix]** must be set to **[YES]**. As a default, this printing is suppressed. When an OUA file with the bond orders included it is read, a new file with the same name and the extension of DAT is written to the same directory. This file contains the matrix of bond orders and it is used by *Molgen* for the reading of particular values of bond orders. The file has the standard size of 40,000 bytes. After a pair of atoms has been clicked, the value of the bond order is displayed.

2.7.11 CONFIGURATION

Using this function, it is possible to define the operation parameters for file generation. Options included in configuration:

2.7.11.1 DIMINISH THE CONVERGENCE CRITERION

The default is **[NO]**.

2.7.11.2 SHELL OPEN/CLOSE

Selection of an electron state (default is **[CLOSE]**).

2.7.11.3 STATE (SINGLET, ...SEXTET)

Configuration of the valence electron orbitals (default is **[SINGLET]**).

2.7.11.4 OPTIMIZATION (NO, SELECTED, FULL)

If selected then Ampac optimizes geometry parameters chosen using functions ***BOND CHOICE (2.7.2)***, ***ROTATION CHOICE (2.7.6)*** and ***RING CHOICE (2.7.8)***. **[NO]** means one SCF calculation and **[FULL]** means that total optimization of the molecule will be done irrespective of for optimization selected geometry parameters (default is **[NO]**).

2.7.11.5 METHOD - (AM1, MNDO, MINDO/3)

Selection of semiempirical method (default is **[AM1]**).

2.7.11.6 CHOICE OF OUTPUT OPTIONS

Definition of the extent of output information.

- Final one-electron matrix **[YES]**/**[NO]**
- Final density matrix **[YES]**/**[NO]**
- Gradients to be minimized using NLLSQ
- Localized orbitals **[YES]**/**[NO]**
- Final UHF spin matrix **[YES]**/**[NO]**
- Final eigenvectors **[YES]**/**[NO]**
- Final bond-order matrix **[YES]**/**[NO]**

NOTE: The default is NO for every option. Bond orders contained in final bond-order matrix are transformed into particular file with the extension DAT. Using the ***BOND ORDER (2.7.10)*** function you can click two atoms and the related bond order will be displayed.

2.7.11.7 COMMENT

Used as description of the output file.

Upon reading an output from Ampac, it is possible to display the optimized geometry parameters using the functions ***BOND SHOW (2.7.1)***, ***ANGLE SHOW (2.7.3)*** and ***RING SHOW (2.7.7)***. The appropriate optimized parameter is displayed as a thick green line. When choosing the number of parameters for optimization, it is necessary to keep in mind the computing ability of the computer.

2.7.12 PATH (see Section 17.1)

2.7.13 IMPORT

List files with the AMP extension (Ampac or Mopac input files). The information on geometric parameters to be optimized is retrieved as well (see also Section ***17.3***).

2.7.14 EXPORT

Write a molecular structure into file with standard format of Ampac (Mopac) input files. The resulting file contains keywords specified by means of the function ***CONFIGURATION (2.7.11)***, and information on geometry parameters to be optimized (if they had been specified and the optimization had been required in configuration). Before this input file is generated, it is possible to check the geometry automatically. If there is more than one molecule displayed on the screen, a single input file will be generated and, as a result, also intermolecular interactions will be calculated (for further reference see also Section ***2.13***).

2.7.15 READ OUTPUT

To load an output file with the extension OUA from the Ampac (Mopac) program. The following data are read and displayed:

- The resulting coordinates of atoms and their atomic numbers. From the coordinates, *Molgen* types of the resulting structure are generated automatically.
- The information as to which geometry parameters have been optimized. After the structure has been displayed, the optimized parameters may be highlighted using the functions ***BOND SHOW (2.7.1), ANGLE SHOW (2.7.3), RING SHOW (2.7.7).***
- The atomic charges that may be displayed by toggling the atomic labels switch to the function ***CHARGES (12.3)***.
- The matrix of bond orders which is rewritten into a special compact file with the name of the main file differing in the extension DAT. The bond orders of a molecule can be viewed using the function ***BOND ORDER (2.7.10)***.

This file may be used later to view the resulting bond orders. The condition is that the file is in the same directory as the corresponding file in *Molgen* format and that the main name is the same. The *Molgen* system only allows you to load the standard formats of Ampac outputs. In the case of non-standard output modifications, an error may occur upon loading.

2.8 PCILO

PCILO is a method designed for the calculation of the properties of ground electron state of a molecular system within the range of semiempirical methods for all valence electrons, electron pairs being considered as localized.

2.8.1 FIRST THREE ATOMS

Used to select first three atoms for the generation of an input file for PCILO method. So called A and B numbering is used for PCILO method input files:

A Absolute numbering. First, all non-hydrogen atoms are numbered, then hydrogen atoms and last, electron pairs are numbered.

B Numbering by chains. In this case the character of atoms is not important.

By means of this function, it is possible to specify three non-hydrogen atoms to be numbered as first for A-numbering.

2.8.2 PATH

To define the directory for PCI files (see ***2.13***).

2.8.3 IMPORT

To browse and read files to be processed for PCILO (see ***2.14***).

2.8.4 EXPORT

To write an input for PCILO method PCI files (see ***2.15***).

2.8.5 READ OUTPUT

To list and read files processed by PCILO. They have the standard extension OUM.

2.9 GEOMO

Generation of input data files for quantum chemistry methods included in GEOMO (CNDO, INDO, MINDO and MINDO/3). The input data files for GEOMO have the extension GMO and are text files. If the output from GEOMO are to be read by *Molgen*, they must have the standard extension OUG. The parameters in the generated file correspond to a structure with complete electron shells in singlet state with no geometry parameters being optimized, and the geometry is written in internal coordinates. Eventual discrepancies concerning the number of electrons in the molecule, the selected electron state and the completeness of electron shells must be solved by

means of a text editor according to the manual for GEOMO system.

2.9.1 PATH

To define the directory for GMO files (see ***2.13***).

2.9.2 IMPORT

To browse and read files to be processed for GEOMO (see ***2.14***).

2.9.3 EXPORT

To write an input for GEOMO method (GMO files) (see ***2.15***).

2.9.4 READ OUTPUT

To list and read files processed by GEOMO. They have the standard extension OUG.

2.10 MM2

To generate input data files for the MM2 method. The generated data files have the extension MM2 and are text files. If the MM2 outputs are to be read by *Molgen*, they must have the standard extension OUM. The generation of file for MM2 is only active if a molecule has been selected by means of the function ***CHOICE OF MOL (2.10.2)***. The attributes of the file for MM2 may be modified using the function ***CONFIGURATION (2.10.5).***

2.10.1 CHANGE OF TYPES

To change atom types for molecular mechanics. Those atoms are interchangeable which have the same number and type of bonds. The new atom type must not be in conflict with the method for determination of MM2 types. For example, the type N.am can not be exchanged for the type N.pl3, because the amide nitrogen type N.am follows from the presence of carbonyl group.

2.10.2 CHOICE OF MOL

This function is used to select molecules for the external molecular mechanics (MM2, MM2PI) or for the molecular mechanics implemented in *Molgen*. For the selected molecule, MM2 types are generated. These may be altered by means of the function ***CHANGE OF TYPES (2.10.1)***. Only a molecule thus selected is applicable for fixed atoms determination (for optimization), conjugated atoms selection (for MM2PI), and file generation. After another molecule has been selected, the MM2 types of the original molecule are transformed back to *Molgen* types and MM2 types are assigned to the new molecule.

2.10.3 FIXED SHOW

To display the atoms, the x,y,z-coordinates of which have not been changed during the optimization by MM2 or MM2PI methods. After this function has been called, the fixed atoms of a selected molecule are marked by an asterisk, and in the window for displaying the function

of labeling atoms, the label "FIXED" appears. The atoms can be selected by means of the function ***FIXED CHOICE (2.10.4).***

2.10.4 FIXED CHOICE

To select the atoms, the x,y,z-coordinates of which will remain constant during the optimization by MM2 or MM2PI methods. The atoms are selected by clicking. If the labeling of atoms is in the this function, an asterisk '*' appears at the selected atom. Repeated selection will remove the mark. If the labeling is in an other function (e.g. ***USER'S (11)***), the atom will be considered to be fixed, but the mark does not appear. To display the marks, use the function ***FIXED SHOW (2.10.3)***. The selection of atoms is only active for a molecule selected by means of the function ***CHOICE OF MOL (2.10.2)***. The fixing of atoms or of whole parts of a molecule decreases the demands for computing upon the optimization. This procedure is especially suitable in the cases when it is not necessary to optimize a whole molecule. The fixed atoms should form a continual part of a molecule. If this requirement is not met, the geometry parameters, values of which depend upon the constellation of atoms in fixed parts of a molecule (bond lengths, angles and torsion angles), will not be optimized correctly.

2.10.5 CONFIGURATION

To set attributes for MM2 and MM2PI methods.

2.10.5.1 OPTIMIZATION <YES/NO>

For **[NO]**, only one calculation of internal energy is carried out (default is **[NO]**)

2.10.5.2 FIXATION OF ATOMS <YES/NO>

For **[YES]** only the position of unfixed atoms will change (default is **[NO]**).

2.10.5.3 DEF. OF PI-SYSTEM

If **[Manually]**, the user is prompted to click atoms included in the system (default is **[Automatically]**). This option is active only for MM2PI method.

2.10.5.4 TYPE OF SYSTEM

This option is active only for MM2PI method (default is **[Planar]**).

2.10.5.5 DIP-DIP INTERACTION

For the option **[Excluded 1,3]**, the interaction of dipoles on the same atom (so called 1,3-interaction) is not included. For the option **[Included 1,3]**, the

1,3-interaction is included. For the option **[Suppressed]**, no interaction is considered. For the option **[From charges]**, the calculation of electrostatic interactions is carried out on the basis of known partial atomic charges (this is the default).

2.10.5.6 VAN DER WAALS' ENERGY

If **[Approximated]** is selected, a simplified calculation is performed. If **[Calculated]** is selected, a complete calculation (this is the default).

2.11 MM2PI

To generate input data files for the MM2PI method. The generated data files have the extension MM2 and are text files. If the MM2PI outputs are to be read by *Molgen*, they must have the standard extension OUM. The generation of file for MM2PI is only active if a molecule has been selected by means of the function ***CHOICE OF MOL (2.11.2)***. The attributes of the file for MM2PI may be modified using the function ***CONFIGURATION (2.10.5).***

2.11.1 CHANGE OF TYPES

To change atom types for molecular mechanics. Those atoms are interchangeable which have the same number and type of bonds. The new atom type must not be in a conflict with the method for determination of MM2 types. For example, the type N.am can not be exchanged for the type N.pl3, because the amide nitrogen type N.am follows from the presence of carbonyl group.

2.11.2 CHOICE OF MOL

This function is used to select molecules for the external molecular mechanics (MM2, MM2PI) or for the molecular mechanics implemented in *Molgen*. For the selected molecule, MM2 types are

generated. These may be altered by means of the function ***CHANGE OF TYPES (2.11.1)***. Only a molecule thus selected is applicable for fixed atoms determination (for optimization), conjugated atoms selection (for MM2PI), and file generation. After another molecule has been selected, the MM2 types of the original molecule are transformed back to *Molgen* types and MM2 types are assigned to the new molecule.

2.11.3 FIXED SHOW

To display the atoms, the x,y,z coordinates of which have not been changed during the optimization by MM2 or MM2PI methods. After this function has been called, the fixed atoms of a selected molecule are marked by an asterisk '*' , and in the window for displaying the function of labeling atoms, the label "FIXED" appears. The atoms can be selected by means of the function ***FIXED CHOICE (2.11.4)***.

2.11.4 FIXED CHOICE

To select the atoms, the x,y,z-coordinates of which will remain constant during the optimization by MM2 or MM2PI methods. The atoms are selected by clicking. If the labeling of atoms is in the function ***FIXED***, an asterisk '*' appears at the selected atom. Repeated selection will remove the mark. If the labeling is in another function (e.g. ***USER'S***), the atom will be considered to be fixed, but the mark does not appear. To display the marks, use the function ***FIXED SHOW (2.11.3)***. The selection of atoms is only active for a molecule selected by means of the function ***CHOICE OF MOL (2.11.2)***. The fixing of atoms or of whole parts of a molecule decreases the demands for computing upon the optimization. This procedure is especially suitable in the cases when it is not necessary to optimize a whole molecule. The fixed atoms should form a continual part of a molecule. If this requirement is not met, the geometry parameters, values of which depend upon the constellation of atoms in fixed parts of a molecule (bond lengths, angles and torsion angles), will not be optimized correctly.

2.11.5 CONJUGATED SHOW

To provide the user with the possibility to display the Pi-electron system of a molecule which has been defined by means of the function ***CONJUGATED CHOICE (2.11.6)***. The atoms which are constituent parts of the -system are marked by an asterisk, and in the window for displaying the modes of atoms labeling, the title "CONJUG." appears. The definition of a Pi-system is important for the MM2PI method.

2.11.6 CONJUGATED CHOICE

To provide the user with the possibility to define those atoms in a molecule which will be considered as conjugated in optimization of the molecule by the MM2PI method. The atoms are selected by using the cursor. If the marking of atoms is in the ***CONJUG*** function, an asterisk appears at the selected atom. Repeated selection of the atom will remove the mark. If the labeling of atoms is in an other function (e.g. ***USER'S***), the atom is considered to be conjugated but the mark does not appear. To display the marks, use the function ***CONJUGATED SHOW (2.11.5)***. The selection of atoms is only active for a molecule selected by means of the function ***CHOICE OF MOL (2.11.2)***. Definition of a Pi-system, in general, improves the quality of the geometry calculation for molecules with conjugated systems. Using the function ***CONFIGURATION***, it is possible to define whether the Pi-system is to be considered as planar or non-planar.

2.12 CONFORMERS

Reading of special files which contain, besides the standard *Molgen* format, also information on the molecule conformers from conformational analysis. The files of this type contain the standard *Molgen* format, in which the basic structure is written. Then the word "CONFORMERS" follows labeling the beginning of information about the conformers. The following data are given in the file:

a) number of rotation bonds

b) numbers of atoms which form the optimized torsion angles

c) original values of the optimized torsion angles

d) step used at the conformation analysis

e) energetic limit

f) optimized torsion angles

g) values of energy determined by the method of molecular mechanics

The files of this type have the standard extension CON. After the file with information on conformers has been read, all the conformations are displayed, the particular conformers being differentiated by colors. The relative energies of the conformations may be followed by the function ***ENERGY***.

2.13 PATH

This function allows to specify the path to the subdirectory containing the files of a defined format. A description is given in Section ***17.1***.

2.14 IMPORT

This function allows you to read files containing molecular structures. The possibility to change the default directory is included. A description is given in Section ***17.3***.

2.15 EXPORT (save file)

This function allows to write a displayed molecule to a file in the directory defined by means of the function ***PATH (17.4)***. The molecule is saved in the format of the selected method. After this function has been called, the user is prompted to choose the molecule to be saved (if there is more than one structure) and to enter a name for the new molecular file.

3 *COLOR (color molecules and atoms)*

This function allows you to color particular atoms or whole molecules independently.

3.1 COLOR MOLECULE

After this function has been selected, the user is prompted to pick the molecule to be colored. The molecule will be colored with the color from the palette to which the white arrow points. Clicking another color filed in the palette will re-color the atom as required. Thirteen colors are available. Clicking ***DEF*** in the palette will color the atoms of the molecule according to the corresponding atomic types.

3.2 COLOR ATOM

After this function has been selected, the user is prompted to pick the atom to be colored. The atom will be colored with the color from the palette to which the white arrow points. The color may be changed arbitrarily by clicking another color. Thirteen colors are available. Clicking ***DEF*** in the palette will color the atom according to the corresponding atomic type.

4 *GEOMETRY VIEW (geometry parameters)*

This function allows you to determine geometry parameters such as interatomic distances, bond and torsion angles. It is also possible to check the geometry of a selected structure or that of all structures displayed on the screen at once. The check controls the reasonability of bond lengths, the values of bond angles and of non-bonding interatomic distances.

4.1 DISTANCE

After this function has been selected and two atoms have been marked, the corresponding distance is displayed. If the marked atoms come from different molecules and one of them is being moved, the changing distance is being displayed during the operation. After an interatomic distance has been displayed, it is possible to mark another pair of atoms immediately.

4.2 BOND ANGLE

After this function has been called and three atoms selected, the bond angle between the joint of the first and second atom and the joint between the second and third atom is displayed. The marked atoms do not have to form an actual bond angle and do not need to be from the same molecule.

4.3 TORSION ANGLE

After this function has been selected and four atoms marked, the angle between the planes is displayed. The first plane is given by the first, second and third atom marked in sequence and the other plane is determined by the second, third and fourth atom. The marked atoms do not have to form a torsion angle corresponding to a particular bond.

4.4 ONE MOLECULE (check the geometry of a selected molecule)

After this function has been selected, the user is prompted to select the molecule, the geometric parameters of which are to be controlled. In the chosen molecule, all interatomic distances (both bonding and non-bonding) and bond angles are checked. Should any chemically unacceptable values be found, an error message containing the number of wrong parameters appears. Then an interactive process is started processing all the faulty values in order interatomic distances - bond angles. The maximum number of registered erroneous parameters is 10. When such a value is found in the course of the control, the user is prompted to enter a new value. Then the geometry is checked again. In the case of false non-bonding distance, the function ***ALTER (1.1)*** is started, by means of which it is possible to change

interatomic distances, bond and torsion angles. After the corrections have been made, the geometry is checked again. The corrections may be interrupted using the function ***EXIT (16.1)*** or by pressing the right mouse button. When the wrong parameters are searched for, these are compared with intervals of chemically acceptable values determined on the basis of atomic type,

hybridization, bond type and other criteria. For example, a non-bonding distance cannot be smaller than the sum of the corresponding van der Waals' radii.

4.5 ALL MOLECULES (control of geometry for all displayed molecules)

For this function the same procedure applies as for the function ***ONE MOLECULE (4.4)***, the only difference being that all molecules displayed on the screen are controlled as well as the mutual distances of their atoms.

5 *FIT (superimposition of two structures)*

The function ***FIT*** is designed to compare the geometry of two molecular structures using the least squares method. The aim of the superimposition is to place the molecules in such a way that the sum of distances of the selected atomic pairs is minimal. After the superimposition has been finished, the molecules are displayed in the main window and the total sum of the mutual distances of atomic pairs is displayed. A lower value represents a closer superimposition and therefore higher similarity of geometry. To obtain a clearer picture, it is advisable to color the molecules using different colors. After the superimposition has been finished, the molecules are unified into a disjoint molecule.

5.1 SHOW MAIN WINDOW

This module allows you to display a reference molecule in the main window.

5.2 SHOW AUX WINDOW

This module allows you to display a comparison molecule in the auxiliary window.

5.3 READY TO FIT (start the superimposition)

Providing that at least three pairs of atoms have been selected, this function will place the compared structure over the reference one by minimizing the mutual distances between the selected atomic pairs using the least squares method. After the superimposition, the molecules are formally unified into a disjoint molecule. This is important for storing the result of the superimposition for further work. To separate the molecules, use the function ***M-BREAK (16.6)***. After the function ***NORMAL/QUALITY (16.12)*** is toggled to the ***[QUALITY]*** mode, the parts of the superimposed molecules occupying the same space are colored white. This coloring remains active also after the function ***FIT (5)*** has been exited.

6 MODELS

The ***MODELS*** module includes functions to allow the display structures in various ways. The output can be directed to the main window, the whole screen, the disk, or directly to a printer or plotter. The following kinds of models are available:

- ***Stick (6.1)***
- ***Space filled (6.2)***
- ***Ball and stick I (6.3)***
- ***Ball and stick II (6.4)***
- ***Dotted van der Waals' surface (6.5)***

All of the models can be moved and rotated on screen. However, the speed of the movement depends markedly on the model type used, on the size of the molecule and on the efficiency of the computer.

6.1 STICK (stick model)

The stick model is the basic and the most simple model that is used to display the structures during all operations of the program. Bonds in this model are represented by joints of atoms, each half of the joint being colored by the corresponding atomic type. By means of the function ***NORMAL/QUALITY (16.12)***, it is possible to alter the mode of the display. In the mode **[*QUALITY*]** , the bonds are displayed in thicker lines and the visibility of bonds is taken into consideration. The function, by means of which the labels can be switched to various modes (***LABEL+***, ***LABEL-***, ***NUMBER***, ***USER'S***, ***CHARGE***), is described in Section ***16.13.*** This kind of model can be directed to the working window ***(6.6)***, the whole screen ***(6.7)***, printer or plotter ***(6.8)***). The output port can be defined by means of the function ***MODEL SETUP (6.10)***.

6.2 SPACE FILL (space filled model)

In this kind of model, the atoms are represented as balls with van der Waals' radii, the values of which are given in Appendix B. The model can be directed to a ***WORKING WINDOW (6.6), WHOLE SCREEN (6.7), PRINT... (6.8)***. Before the model is displayed on the whole screen, a test for the presence of a VGA card is run and if it is present, the model is displayed in 256 color mode. The following resolutions can be used for display:

360x480 VGA resolution of 360x480 in 256 colors requires a minimum of 256 kB of videocard memory.

640x480 VGA resolution of 640x480 in 256 colors requires a minimum of 512 kB of videocard memory.

800x600 VGA resolution of 800x600 in 256 colors requires 1024 kB videocard memory.

1024x768 Super VGA resolution of 1024x768 in 256 colors requires 1024 kB videocard memory.

Using the function ***PRINTER SETUP (6.9),*** the output port, format and any switches for printer, plotter or output file can be specified. Additionally, it is possible to select the ***[ANIMATION]*** option which provides a series of images of the same molecule. After the module ***MODEL SETUP (6.11)*** is exited, the user is prompted to enter a filename for the images to be stored. Then the model is automatically rotated about the x-axis with a selected step. The sequence thus generated can be screened using the module ***VIEW PICTURE (6.10)***. The display of a model can be interrupted by pressing ESC. When an animation is being generated, first the current picture is finished and only then can the process be interrupted.

6.3 BALL & STICK I

In this kind of model the atoms are represented as balls and the bonds as thick lines. Using the function ***MODEL SETUP (6.11)***, the radius of the balls (Ball factor) can be changed. This model can be directed to the ***WORKING WINDOW (6.6), WHOLE SCREEN (6.7), PRINT...(6.8)***.

6.4 BALL & STICK II

This model represents atoms as balls and bonds as tubes. Perspective is included in this kind of depiction. Using the function ***MODEL SETUP (6.11)*** the thickness of bonds can be specified as well as the perspective, the method of coloring of molecules, and whether the picture should be saved to a file. The model can be directed to the ***WORKING WINDOW (6.6), WHOLE SCREEN (6.7), PRINT... (6.8)***.

6.5 DOT SURFACE (dotted van der Waals' surface)

In the dotted model, the atoms are represented by their van der Waals' radii regularly covered with colored dots. The values of van der Waals' radii are given in Appendix B. The dotted model can be directed to the ***WORKING WINDOW (6.6), WHOLE SCREEN (6.7), PRINT...(6.8)***. Filename for saving an animation, density of the dots, and the presence of stick model within the dotted one can be specified using the function ***MODEL SETUP (6.11).***

6.6 TO WINDOW (output to working window)

The model is displayed in the working window.

6.7 TO SCREEN (output to whole screen)

The model is displayed on the whole screen. This output is suitable for photographing from the screen. The program automatically finds out the maximum possible resolution for the graphic card used and uses it for the display. As a default, the program works in 640x350 resolution which corresponds to an EGA graphic card. If a VGA card is present, the graphic mode is set to 640x480, or for SVGA to 1024x768. When a mouse button or ESC is

pressed, the graphic mode is reset to 640x350 regardless of the type of card used.

6.8 PRINT...

Output to printer, plotter or file.

6.9 PRINTER SETUP

Molgen provides device drivers for man printers, plotters and files:

9-pin dot matrix printers - including Epson FX and MX series printers, IBM Graphics Printer and Proprinter, and Panasonic, OkiData, and other printers with Epson or IBM emulation up to 240x216 dpi.

24-pin dot matrix printers - including Epson LQ, NEC Pinwriter, Proprinter X24, and other printers with Epson LQ or Proprinter X24 emulation-at up to 360x180 dpi.

Laser and ink jet printers - including LaserJet, DeskJet, DeskJet 500C, DeskJet 550C, PaintJet, and Canon LBP 8.

PostScript printers
Hewlett-Packard pen plotters (HPGL) - including HP7475

File formats:

PCX PaintBrush
BMP Bitmap
IMG Gem IMG
TIF (compressed and uncompressed)
CGM Computer Graphics Metafile (ANSI)
DXF AutoCad
WMF Windows Metafile

WPG Word Perfect Graphics (vector form)
PCT Color QuickDraw (PICT)
VSHO General Parametrics Video Show (ANSI NAPLPS)
AI Adobe Illustrator PostScript

The graphics device may be either parallel (LPT1-4) or serial (COM1-4). You may also send output to a file and print it later by using the program GCOPY.

No Reset

This option applies only to the Laser and DeskJet drivers. Normally these drivers reset the printer with an ESC-E each time you initiate printing. This will eject a partially printed page. This option suppresses the initial ESC-E. Do not use this option if you are also using printer fonts.

Fine Patterns

In medium and high resolution modes, the device drivers magnifies fill patterns by a factor of 2 or 4 times to make the hatch visible. This option suppresses the pattern expansion. It has no effect in low resolution modes, on plotters, or in PostScript mode.

Rel Advance

This option controls how the top margin is measured when using a "Custom Page Size" on PCL printers. Normally is measured from the top edge of the paper. When this option is set it is measured from the current cursor position (like dot matrix printers). The following restrictions apply:

- You may also need to use "Form feed - Never" and/or "No Reset" to suppress page ejects.

- This option won't work with LJ, LJ3R, or LJ4 in landscape mode.

If your graph doesn't fit on the paper the bottom will be cut off.

Enhance DJ Color

This option improves the output quality from DJC and DJC550 at the expense of slower print times. You should always set this option when plotting on transparency film or glossy paper.

No Background Paint

This option has no effect for printers or plotters. For vector DTP files, it instructs the driver not to draw the background color.

Invert

This option has no effect for printers or plotters. It inverts the background color for some of the file formats on your DTP driver disk.

Custom Page Size

Use this option to override the graph's default size or location on the page. Width, Height, Left Margin and Top Margin are also required.

Width, Height

These fields are used only if "Custom Page Size" option is set. Width is the length of the x-axis and Height is the length of the y-axis. Units are inches x 1000. The plotter drivers assume 1000 dots per inch, so 1/1000 inch corresponds to a physical plotter unit. The precise value for HP plotters is 40 dots per millimeter (a little more than 1000 dpi).

Left Margin, Right Margin

These fields are used only if the "Custom Page Size" option is set. They control the graph's location on the page. Units are inches x 1000. There is no bottom margin when you use a custom page size. If you want white space below the graph you must create it yourself by printing blank lines (dot matrix printers) or issuing a cursor control escape sequence (PCL printers).

Plotters

On plotters the point (LeftMargin, TopMargin) is the Pl position, or lower left corner of the plot. In other words, what is called 'top margin' is really the bottom margin. In both rotated (portrait) and normal (landscape) modes, LeftMargin is the distance from the edge of the paper to the left edge of the graph.

Printers in Portrait Mode

Dot matrix and laser printers have an unprintable region at the top of the page. The driver will increase your top margin if it is too small. The unprintable area for dot matrix printers is not an exact value, since it depends on how the user loads the paper in his printer. The program assumes that the paper is at top of form when you print, and that the unprintable area is 1/3 inch. If the "Rel advance" option is ON, the top margin is measured from the printhead position at printing, with no adjustment for an unprintable area. PostScript's natural coordinate system is with the y value increasing upward. This option sets the top margin assuming an 11 inch page length. It calculates the PostScript y-coordinate for the top of the graph from:

$$y0 = 11000 - \text{TopMargin} \quad \text{(inches x 1000)}$$

With long paper you may need to use a negative TopMargin to get your graph in the right place.

Printers in Landscape Mode

Printers in landscape mode are more confusing, because the margins are measured from the upper left corner of the physical page (the way the paper emerges from the printer) rather than the logical page (the orientation of your graph). The "top margin" will never be above your graph. It will be either on the left or the right, depending on what printer you have.

Form Feed

Normally the only drivers that issue a form feed after printing a graph are the LaserJet, DeskJet, and PostScript printers. The dot matrix drivers advance the paper exactly 11 inches (full page mode) by counting lines. The form feed options let you suppress or force a form feed for all devices expect PostScript. They occupy two bits, if neither bit is set you will get the driver default.

On the LaserJet, the page doesn't print until the printer is issued a form feed or a reset instruction. If you use "Never", your program must send the form feed.

On a plotter, "Always" will place the HPGL instruction AF in the output. Only sheet-fed plotters implement this instruction.

If you use "Always" on a dot matrix printer in standard-size full page mode you will get an extra blank page after the graph.

GCOPY.EXE

Stand-alone program distributed with *Molgen* for the printing of graphics images. The program is a proprietary product of ***FLEMING SOFTWARE.*** When you send graphics to DiskFile, the file captures the data stream that would otherwise be sent direct to the device. In theory, you can print this file with the COPY command. In practical, COPY may send data faster than the device can process it. GCOPY is a substitute for COPY that slows transmission to the device speed.

Syntax:

GCOPY FileList Port[:Parms] [/X=#] [/T=n] [/HP] [/HI] [/F=n]

Parameters:

FileList:	List of files to copy (wildcards accepted)
Port:	LPT1-4 or COM1-4 (not PRN or AUX)
Parms:	baud, parity, data, stop for the serial port:

baud = 110, 150, 300, 600, 1200, 2400, 4800 or 9600
parity = N, E or O
data = number of data bits (7 or 8)
stop = number of stop bits (1 or 2)

If you omit the Parms option, GCOPY will not initialize the serial port. If you enter some but not all parameters the omitted ones will default to :

parity = N
data bits = 8
stop bits = 1

You must enter Parms if you use /X option.

/X Use this option for XON/XOFF handshake on the serial port. # is the port's IRQ number. It is required for COM3-4 and optional for COM1-2. When you use /X for an HP plotter you must also use /HP.

/T=n Change the device time-out to n seconds (default is 30)

/HP Indicates that the device is an HP plotter. GCOPY will send the HP ON/XOFF setup string if you use the /X option. It will also stop at the beginning of each page and ask you to insert paper (on the /F action)

/HI GCOPY will detect DPM/L begin page instructions and prompt you to insert paper.

/F=n Use this option only to print to an HPGL plotter equipped with a sheet feeder. Set n to P to prompt the user to insert paper at the beginning of each page (default), Y to issue an AF instruction, or N to just print continuously.

PRGUIDE.EXE

Stand-alone program distributed with *Molgen*. The program is a proprietary product of ***FLEMING SOFTWARE***. PRGUIDE lets you browse through a list of about 900 printers to see what driver is appropriate. The data file PRGUIDE.TXT must be in the some directory as PRGUIDE.EXE. In addition to scrolling with the cursor keys, you can use the alphabetic keys to jump around the manufacturer name.

6.10 VIEW PICTURE

All models (stick, space fill, ball & stick) can be saved on disk. By clicking the 'S', the user is prompted to define a filename where the picture or set of pictures used in the animation will be stored. The animation is available for the space filled model for 16 colors and dot surface model. The set of animations is stored in pictures created by the rotation around the x-axis with the selected step (15, 30, 45 or 60 degrees).

The choice of a model and setting up of its attributes must be done before clicking "S" in PROPERTIES.

The option for saving the picture into a file is included in the menus of all previously mentioned functions. The animation is available for the space filled model for 16 colors and dot surface model.

6.11 MODEL SETUP

After this function has been selected, a configuration window for the given model appears in the main window, showing the options for its configuration. This function allows you to change the parameters of a models. Each model has its own configuration module.

6.11.1 SPACE FILL

Animation

If Animation is ***Yes***, a series of pictures is generated automatically to the whole screen by sequential rotation about the x-axis with the step selected from the scale of predefined sizes. The process can be interrupted by pressing ESC. The series can be viewed by means of the function ***VIEW PICTURE (6.10).***

Type of model

16 colors	640x350
	640x480
256 colors	360x480
	640x480
	800x600
	1042x768

6.11.2 BALL & STICK I

Ball factor

Sets the size of the balls representing atoms.

Bond factor

Sets the size of the bonds between atoms.

Bond color

Yes - the bonds are green

No - the bonds are black

Atom color

Nothing - atoms are black

Hetero - hetero atoms are colored

All - all atoms are colored

6.11.3 BALL & STICK II

Bond factor

Sets of the size of the bonds between atoms.

Perspective

Controls the perspective for viewing the molecule.

Colored

Atoms - All atoms are colored separately

Molecules - Molecules are colored separately

6.11.4 DOTTED SURFACE

Stick drawing

If ***Yes,*** a stick model is displayed within the dotted one.

Density of points

Allows you to control the density of points for the model and thus also to influence the speed of display.

Size of points

Small, Medium, or Large

Animation

If Animation is ***Yes***, a series of pictures is generated automatically to the whole screen by sequential rotation about the x-axis with the step selected from the scale of predefined sizes. The process can be interrupted by pressing ESC. The series can be viewed by means of the function ***VIEW PICTURE (6.10).***

7 HYDROGENS ADD

The function allows you to add hydrogen atoms to the unfilled or missing valences of a molecule. The hydrogen atoms are added with respect to the standard bond lengths and angles.

8 HYDROGENS REMOVE

This function provides the ability to remove all hydrogens using a single command. The function is useful when the picture on the screen gets unclear because of a great number of displayed atoms.

9 EL PAIRS ADD (add lone electron pairs)

This function allows you to add free electron pairs to nitrogen, oxygen, sulfur and halogen atoms, using a single command. The electron pairs are added with respect to the standard bond angles. The standard size of a lone electron pair is one Ångström.

10 EL PAIRS REMOVE

This function allows you to remove all free electron pairs using a single command.

11 USER'S TYPE (user defined atomic labels)

This function allows you to assign the atoms different labels. After this function has been selected, the user is prompted to mark the atom to which the user's label is to be assigned. The function can be exited by using the function ***EXIT (16.1)***. User defined types can be displayed by toggling the atomic label switch to the function ***USER'S (15.13)***.

12 PROPERTIES

The **PROPERTIES** module includes the calculation and graphic depiction of lipophilicity represented by the logarithm of partition coefficient (log(p), $K_{o/w}$, or K_p) in the system of n-octanol/water; the calculation of polarizability of molecules represented by molar refraction; calculation of the surface and volume of molecules; calculation of Henry's Constant, and graphic depiction of electrostatic and lipophilicity potential. In addition, the module includes procedures providing a comparison of mentioned potentials and molecular shapes.

12.1 LIPOPHILICITY

This module provides for the calculation of the logarithm of the partition coefficient in the system of n-octanol/water and spatial depiction of the lipophilicity employing the defined lipophilicity potential. In the log(p) calculation, the fragmentation method according to Crippen[1] is used. Two methods are used: a topological approach based exclusively on Crippen's fragment contributions and a method that considers the solvent accessible surface into account. The last method is applicable only in molecules composed from C,H,N,O atoms. The list of used fragments and their contributions to the total value of log(p) are given in the text file CRIPPFRG.TAB. The fragmentation method provides the estimate of log(p) with the standard error of 0.4%. Before the calculation is commenced, the user is prompted to specify whether the calculation should be conformationally dependent or independent. In the former case, the actual conformation of the molecule on the screen is taken into account. The latter option provides a calculation based on standard average values of particular fragments. In the calculation of the total value of log(p), only the contributions of marked atoms will be considered. The marked atoms can be made visible by switching the window for atom symbols to the function ***<MARK>***. An asterisk (*) appears on the marked atoms.

The log(p) values are computed using equation:

where is the atomic contribution to the total log(p), is the actual solvent accessible van der Waals' surface of the i*th*-atom, is a maximum solvent accessible van der Waals' surface of the i*th*-atom and is the stepwise function that relates the number of possible external hydrogen bonds to accessible area of the i*th*-atom.

12.1.1 COMPAR. OF POTEN

This function allows you to compare cuts of potentials for two molecules. The comparison consists of subtracting the value of the second molecule's potential from that of the first one. The comparison starts from the actual Cartesian coordinates of the structures to be compared. This

means that if the comparison is to be of any use, the structures should be placed one on the other (e.g., using the function ***FIT (5)***). Before the function is started, both molecules must be united (using the function ***M-UNITE (16.15)***) into a single disjoint molecule and a plane must be generated. It is advisable to set the scale of potentials manually so that the maximum and minimum values differ only in the sign (e.g. ±0.1).

12.1.2 DISTRIBUTION NUM.

The function provides for the display of atomic contributions used in log(p) calculation for all molecules at the same time.

12.1.3 CALCULATION

Calculation of the logarithm of partition coefficient (log(p)). If there is more than one molecule displayed on the screen, the user is prompted to select one of them for which the calculation is to be carried out. This molecule will be displayed in the form of stick model with thicker lines. The calculation can be carried out for a molecule consisting of at most 100 atoms. The algorithm used provides for the accuracy of the calculation with the error of 0.4%. Should there be a fragment in the molecule that is not present in the CRIPPFRG.TAB file, the error message "Fragmentation failed" is displayed. Upon calculation of the total log(p) value, only the contributions of marked atoms are considered. The marking of atoms can be displayed by toggling the atomic symbol switch into the ***MARK*** function whereby an asterisk appears at the marked atoms. The atoms can be marked or their marks removed using the functions ***MARK ATOM BY ATOM (12.6), MARK ALL ATOMS (12.7), DELETE ALL MARKS (12.8)***. Before the calculation is commenced, all atoms are checked for marks. If there are no marked atoms present, all atoms are marked.

12.1.4 POTEN. ON SURFACE

If the conformationally dependence of lipophilicity is followed, the lipophilicity points are calculated using equation:

where the meaning of the variables is the same as in the previous equation and in equations used for log(p) calculations (see ***12.1***).

Depiction of the lipophilicity potential[2]. Lipophilicity potential provides a picture as to the extent of lipophilicity for individual parts of a molecule or a system of molecules. It is useful as one of the criteria for assessing the similarity of molecules, e.g., upon a study of biological activity of materials. Unlike the total log(p), the

lipophilicity potential describes also differences between molecules with equal log(p) but different distribution of lipophilicity due to, for instance, a different constellation of polar groups. After this function has been triggered, the van der Waals' surface is regularly covered with points where the lipophilicity potential values are calculated. To display the results, the standard scale of 13 colors is used. Blue shades represent hydrophilic parts of the molecule; red ones represent the lipophilic parts. The mode of display can be modified using the module ***CONFIGURATION (12.1.10)***.

12.1.5 POTEN. IN PLANE

This function allows you to depict the lipophilicity potential in a cut parallel with the screen plane. The cut is represented as a rectangle generated by means of the function ***GENER. PLANE***. If the molecules have been united using the function ***M-UNITE (16.15)***, only one rectangle is generated for all molecules. After a rectangle has been generated, its z-coordinate is equal to zero and its center is identical with the center of the molecule. To define the position of the cut plane exactly as required, use the function ***ATOMS INTO PLANE (12.1.8)*** to specify a plane given by any three atoms, which will then be parallel with the screen plane. Then it is possible to set the distance of this plane from the screen plane in the range of ±100 Ångströms using the function ***Z (12.1.9)***. The size of the rectangle with respect to the molecule can be interactively adjusted using the functions ***PX+***, ***PX-***, ***PY+***, ***PY-***.

12.1.6 GENER. PLANES

This function generates a rectangle representing a cut plane for each displayed molecule (as well as for a disjoint molecule which can be obtained using the function ***M-UNITE (16.15)***). The rectangles are generated so that their centers are identical with the centers of the molecules and that there is enough space for display of the potential. The size of a rectangle can be adjusted using the functions ***PX+***, ***PX-***, ***PY+***, and ***PY-***. In the function ***GLOBAL***, the corresponding dimensions of all displayed rectangles are altered, while in the function ***RELATIVE***, only the selected rectangle is affected. The molecules in the rectangles can be moved or rotated by means of the standard functions whereby the rectangles move with the structures. Also, upon

enlarging or shrinking of the displayed structures, the size of the corresponding rectangles follows the changes.

12.1.7 DELETE PLANES

This function allows you to remove all rectangles representing cut planes.

12.1.8 ATOMS INTO PLANE

After any three atoms of a molecule have been selected, the structure is oriented so that the marked atoms are within the screen plane. This function is important when defining an exact position of molecules upon calculations of potential cuts.

12.1.9 DIST. IN Z-AXIS

This function allows you to shift a selected molecule along the z-axis by a defined distance. The function is active only if the selected molecule has been oriented by means of the function ***ATOMS INTO PLANE (12.1.8)*** so that three of its atoms lay in the screen plane (xy-plane). After the function ***DIST. IN Z-AXIS*** has been selected, the distance of the plane defined by three selected atoms from the xy-plane is displayed. Entering a new value for the distance will change this distance appropriately. Positive values cause "submerging" of the structure into the screen and negative values cause its "emerging" from the screen. By means of this function it is possible to create a series of parallel cuts of a molecule. The distance of the cut plane from xy-plane will be displayed under the resulting picture.

12.1.10 CONFIGURATION

This function allows you to modify the graphic output of potentials (electrostatic, lipophilicity) on van der Waals' surface or in a cut plane.

12.1.10.1 DRAWING OF THE MODEL <YES/NO>

If **[YES]**, the image produced will include the stick drawing of the molecule with atoms beyond the cut plane being colored white.

12.1.10.2 DENSITY OF POINTS <LOW/HIGH>

Adjust the density of dots used to represent the van der Waals' surface.

12.1.10.3 SIZE OF POINTS <SMALL/NORMAL/LARGE>

Definition of the size of dots used for the dotted surface model or the raster for the display of cut of potentials.

12.1.10.4 TYPE OF MODEL <DOT SURF/SPACE 16/SPACE 256>

Define the type of model to display the potential on van der Waals' surface. For the **[DOT SURF]** model, potential is calculated in regularly distributed points. It can be combined with stick model. For the **[SPACE 16]** or **[SPACE 256]** space filled model, the potential is represented by a coherent van der Waals' surface in the function of 16 or

256 colors. The models **[DOT SURF]** and **[SPACE 16]** are applicable for EGA, VGA or SVGA graphic cards in the raster of 640x350 or 640x480 points, respectively. **[SPACE 256]** is only available with VGA and SVGA (see ***6.2***).

12.1.10.5 DEFINE SCALE <BY PROGRAM/BY USER>

Set the range for values of potential to be displayed. If **[BY PROGRAM]**, the scale is defined so that all the available color palette is used. In **[SPACE 256]** model, the scale of 13 basic colors is used, the rest being used to represent shades. If **[BY USER]**, the scale can be changed manually using the windows **[RANGE FROM]** and **[RANGE TO]**.

12.1.10.6 INTERACTION <YES/NO>

This option allows you to suppress interactions between molecules. This is important when a visual comparison of a series of materials is required. If you want to examine the influence of interactions on the shape of the potential, the molecules must be unified into a disjoint molecule using the function ***M-UNITE (16.15)***. Then, only one rectangle is generated.

12.1.10.7 PICTURE <SAVE/ANIMATION/NOT SAVE>

If **[SAVE]** or **[ANIMATION]**, the resulting images will be stored on disk in files that can later be loaded using the

module ***VIEW PICTURE (6.10)***. If **[ANIMATION]**, the structure will be rotated about the x-axis with the selected step. The process can be interrupted by pressing ESC. The pictures generated before interruption remain stored on disk. For **[SPACE 256]** model, the animation is not available. If **[NOT SAVE]**, the image will only be generated on screen.

12.1.10.8 DESCRIPTION

Display of the scale on x- and y-coordinates.

12.1.10.9 INTERVAL

Set the interval for the graph.

12.1.10.10 COMMENT

Input text to be displayed above the picture.

12.1.10.11 PALETTE

Allows you to select from the list of palettes or to define a new palette using the standard set of colors. Here, the red shades represent lipophilic parts of the molecule and the blue shades represent hydrophilic areas. This module has its own menu which includes:

[ACCEPT] - Update the current palette selected from the list and return to the previous menu. The default palette is highlighted in a red frame. If there is no palette in the default directory, the standard palette is used.

[VIEW] - Display and edit colors of the current palette. To browse the colors, use the left and right arrows. The color can be set using the up and down arrows. To save the created palette into a new file, press |'. To overwrite the original palette with a newly defined one, press ENTER. Pressing ESC returns to the previous menu without accepting any changes.

[NEW] - Allows you to enter a filename for a new palette and the number of colors. Then palette editor, described at the function **[VIEW]**, is started.

[DELETE] - Allows to delete existing palettes.

[STANDARD] - Install the standard palette and return to the previous menu.

[VIEW ST] - Edit the standard palette (see the description under the function **[VIEW]**).

12.1.10.12 SURFACE <YES/NO>

If **[YES]**, the calculation will include the computation of surface of particular isopotential areas. The resulting values will be displayed by the corresponding intervals in the scale used.

12.2 MOLAR REFRACTIVITY

This module provides the calculation of molar refractivity of a selected molecule. For the calculation of log(p), the Crippen fragmentation method is used. A list of used fragments can be found in the text file CRIPPFRG.TAB. The fragmentation method provides estimates of molar refractivity with the standard error of 1.1%. Only the contributions of marked atoms are considered.

The marking of atoms can be made visible by toggling the atomic label switch into the function ***MARK***. Then, an asterisk '*' appears at the marked atoms.

12.2.1 DISTRIBUTION NUM.

This function returns atomic contributions used in the calculation of molar refractivity.

12.2.2 CALCULATION

This function triggers the calculation of molar refractivity. If there is more than one molecule displayed on the screen, the user is prompted to specify for which molecule the molar refractivity is to be calculated. The selected molecule is then displayed using the stick model with thicker bonds. The calculation is available only for molecules consisting of up to 100 atoms. The algorithm used provides an accuracy with error of 1.1%. If there is a fragment in the molecule that is not present in the file of fragments (CRIPPFRG.TAB), the calculation will not commence and the error message "Fragmentation failed" appears. Only the contributions of marked atoms are considered. The marking of atoms can be made visible by toggling the atomic label switch into the function ***MARK***. Then, an asterisk '*' appears at the marked atoms. The atoms can be marked or their marks removed using the functions ***MARK ATOM BY ATOM (12.6), MARK ALL ATOMS (12.7), and DELETE ALL MARKS (12.8)***.

12.3 CHARGE

Depiction of partial atomic charges, electrostatic potential and gradient on van der Waals' surface or in a cut plane. Reliable values of partial atomic charges can be obtained from semiempirical quantum chemistry methods included in the program packages Ampac, Mopac or Geomo. *Molgen* can read output files from these programs and use values thus obtained for depiction of electrostatic properties of the corresponding molecules. Upon saving, the partial atomic charges are written to all files MOL and XYZ. Besides the atomic charges from quantum chemistry calculations the atomic charges can be calculated by an original empirical method based on iterative partial equalization of orbital electronegativity[3]. Whenever the user switches the labels to this function, before the values are displayed, a prompt appears to specify which values should be used: original - i.e. those from the MOL file (if not calculated by quantum chemistry methods previously then zeros would appear) or whether they should be calculated (the above-mentioned empirical method is used).

12.3.1 COMPAR OF POTEN

This function provides a visual comparison of cuts of two molecules. The comparison consists of subtracting the value of the second molecule's potential from that of the first one. The comparison starts from the actual Cartesian coordinates of the structures to be compared. This means that if the comparison is to be of any use, the structures should be placed one on the other (e.g., using the function ***FIT (5)***). Before the function is started, both molecules must be united (using the function ***M-UNITE (16.15)***) into a single disjoint molecule and a plane must be generated. It is advisable to set the scale of potentials manually so that the maximum and minimum values differ only in the sign (e.g. ±0.1).

12.3.2 COMPAR OF GRAD

This function provides a visual comparison of cuts of gradients of two molecules (see also Section ***12.3.1***).

12.3.3 NEW OR CHANGE

This function allows you to change partial atomic charges. After it has been started, all atomic charges of the displayed molecule are shown. Clicking an atom is followed by a prompt to enter a new value for the atomic charge. When changes are complete, click ***EXIT (16.1)*** to get a menu consisting of ***[NO CHANGE]***, ***[UPDATE]***, and ***[CANCEL]***. ***[NO CHANGE]*** restores the original values of atomic charges, ***[UPDATE]*** replaces the original values with new ones, and ***[CANCEL]*** cancels the most recent selection of an atom and prompts to select another one.

12.3.4 DISTRIBUTION NUM.

This function returns partial atomic charges for all molecules. An equal effect can be achieved by toggling the atomic label switch into the function ***CHARGE (12.3)***.

12.3.5 CALCULATION

To calculate the atomic charges in sigma-bonded and nonconjugated Pi-systems. Rapid empirical calculation of atomic charges that can be used in the description of the electrostatic potentials. The "Partial Equalization of Orbital Electronegativity" PEOE method is used for atomic calculation introduced by Gasteiger and Marsili[8] .

12.3.6 POTEN. ON SURFACE

Depiction of electrostatic potential on van der Waals' surface. The permitivity of the medium is considered to be constant and therefore it is not included explicitly in the relationship for the calculation of potential. After this function has been started, van der Waals' surface of a molecule is covered regularly with points where the values of electrostatic potential are calculated. For display, a scale of 14 colors is used. Red colors represent positive values of electrostatic potential (i.e.

sites where the positive unit charge is repulsed) and blue shades mean negative values of electrostatic potential (i.e. sites where the positive unit charge is attracted). The mode of display of potential can be modified by means of the module ***CONFIGURATION (12.1.10)***.

12.3.7 GRAD ON SURFACE

Depiction of the gradient of electrostatic potential on van der Waals' surface. The permitivity of the medium is considered to be constant and therefore it is not included explicitly in the relationship for the calculation of gradient of electrostatic potential. After this function has been started, van der Waals' surface is covered regularly with points where the gradient of electrostatic potential values are calculated. A scale of 14 colors is used for display of the results. Red shades represent areas where the positive unit charge is repulsed and blue shades mean sites where the unit charge is attracted. The mode of display can be modified by means of the module ***CONFIGURATION (12.1.10)***.

12.3.8 POTEN. IN PLANE

This function provides the display of electrostatic potential in a cut plane defined by the screen plane. To obtain that, a rectangle must be defined using the function ***GENER. PLANES (12.1.6)***. It is recommended to orient the molecule prior to the calculation (by means of the function ***ATOMS INTO PLANE (12.1.8)***) so that the position of the molecule with respect to the cut plane will be precisely defined. It is possible to set further attributes of the display using the function ***CONFIGURATION (12.1.10)***.

12.3.9 GRAD IN PLANE

The function provides display of gradient of electrostatic potential in a cut plane defined by the screen plane (see ***POTEN. IN PLANE (12.3.8)***).

12.3.10 GENER. PLANES

See Section ***12.1.6.***

12.3.11 REMOVE ALL PLANES

See Section ***12.1.7***.

12.3.12 ATOMS INTO PLANE

See Section ***12.1.8***.

12.3.13 DIST. IN Z-AXIS

See Section ***12.1.9***.

12.3.14 CONFIGURATION

See Section ***12.1.10***.

12.4 VOLUME AND SURFACE

This module allows you to calculate volume and surface of all or part of a molecule. If there is more than one molecule displayed on the screen, the user is prompted to select the molecule for which the values are to be calculated. Only marked atoms are included in the calculation. (Atoms can be marked or their marks removed using the functions ***MARK ATOM BY ATOM (12.6), MARK ALL ATOMS (12.7), DELETE ALL MARKS (12.8)***.) If there has been no marked atom when the function has been triggered, all atoms are marked automatically. The marked atoms can be made visible by toggling the atomic label switch into the function ***MARK***. The calculation can be interrupted by pressing ESC. The result is displayed in the upper right corner of the screen. The progress of the calculation is represented graphically. For the calculation, van der Waals' radii according to Gavezotti[4] are used. The algorithm works with an error of 0.7%. After this function has been called, the following possibilities of calculation are available: **[SOLVENT ACCESSIBLE]** and **[WITHOUT SOLVENT]**. The former option allows you to calculate the solvent accessible surface when the solvent is approximated by a sphere with a defined radius. For water, the standard solvent, the value of 1.4 Å is used. The size of solvent radius can be set up by user. The last option makes it possible to calculate the volume and surface of the described molecules (formally the solvent radius is zero).

12.5 3D COMPARISON

This module provides for the comparison of the shape of molecules and that of their three-dimensional electrostatic and lipophilicity potentials. The aim of this method is to find a constellation of the two molecules in which the differences of their studied properties are minimal and thus obtain the measure of their similarity. The examined properties can be the shape of van der Waals' surface, electrostatic, and lipophilicity potential. In this module, a

modification of an algorithm[5] is used which is based upon fast comparison of spherically distributed points around both molecules under comparison. The points represent tips of an icosahedron and dodecahedron with spherical symmetry. The values in these points represent, as required, shape, electrostatic or lipophilicity potential of a molecule. The symmetry of mentioned polyhedrons allows for a very fast comparison of polyhedrons of two molecules using relatively small number of operations. Thus it is possible to register all maxima of similarity as well as the global one that is difficult to find by means of other methods. The output of this method is a series of rotational shifts about the x, y, and z Cartesian axes that define a constellation of the molecules where the similarity is maximal. The resulting position of the molecules is displayed by means of a combination of ***BALL & STICK II*** model and ***DOT SURFACE*** model. Atoms with identical coordinates are highlighted white. After the procedure of ***3D COMPARISON*** has been started, all disjoint molecules are separated into individual structures. If the number of present structures then exceeds two, the calculation is terminated.

12.5.1 SHAPE

Distance of the tip from the van der Waals' surface of the molecule.

12.5.2 ELECTROSTATIC POT.

Value of the electrostatic potential in the point where the joint of the tip and the gravity center intersects the van der Waals' surface.

12.5.3 LIPOPHILICITY POT.

Value of the lipophilicity potential in the point where the joint of the tip and the gravity center intersects the van der Waals' surface.

12.6 MOLECULAR ORBITAL

12.6.1 GENERATION

This option provides the possibility of calculation of shape or electron density of an orbital. This module consists of two parts:

1. Selection of molecular structures with calculated canonical (CMO) or localized (LMO) molecular orbitals (max. 10 structures)

2. Depiction of a molecular orbital for particular molecules in the form of horizontal lines representing energy levels for particular a particular molecular orbital.

- Highlighting of HOMO and LUMO orbitals.

- Display of energy of selected orbital and its degenerations.

- Selection of molecular orbitals for calculation of shape or electron density of an orbital.

3. Specification of the space for orbital to be calculated, by means of setting of the size of a box.

- Basic functions for handling the molecule and box (rotation, translation, resizing, orientation, centering of a molecule or an atom within the box, various labeling of atoms).

- Three types of models to display molecular structure.

- Setting of the basic parameters for the calculation of orbitals (distance between points where the values are calculated, minimal development coefficient of atomic orbitals, surface constant of the orbital or the border value of electron density).

After a name is specified for the file in which the information is to be written, the calculation begins. It can be interrupted by pressing ESC. Values obtained up to that point are written in the specified file and even incomplete data like this can be used to display the orbitals.

Following is a sample Mopac input file for calculation of orbital data:

```
VECTORS LOCAL

1  0.00000  0  0.00000    0  0.00000    0  0  0  0
8  0.94209  0  0.00000    0  0.00000    0  1  0  0
6  1.42534  0  107.34589  0  0.00000    0  2  1  0
1  1.11119  0  107.98388  0  60.09530   0  3  2  1
1  1.11093  0  107.76638  0  180.00021  0  3  2  1
1  1.11118  0  107.98393  0  299.90439  0  3  2  1
```

12.6.2 LOAD ORBITALS

The orbitals to be loaded by this module are calculated in the module ***GENERATION (12.6.1)***. Orbitals of up to four molecules can be displayed at a time. The files containing the molecular structure have the standard extension ORB. A yellow color represents orbital electron densities, and the combination of blue and red colors is used to display the orbital shapes whereby negative part of the orbital is blue and the

positive one is red. The basic information on a calculated orbital can be obtained using the function ***INFO-MOL***.

12.6.3 DISPLAY AND SELECTION OF ORBITALS

This module serves for the graphic display of molecular orbitals depicted as horizontal lines representing energy levels of the molecular orbital. The HOMO and LUMO orbitals are distinguished by different colors; HOMO being displayed green, and LUMO being displayed red.

12.6.3.1 LOAD AMPAC (*.OUA)

This is a text file containing calculation results from the quantum chemistry package Ampac. To have the information on canonical (CMO) and localized (LMO) orbitals contained, the input for calculation must include the keywords VECTORS (write matrix of vectors) and LOCAL (write matrix of localized orbitals). These two matrices are identified, loaded, and written to special files with the extension CO1, CO2 (for CMO) and LO1, LO2 (for LMO) and the molecular structure is written into a file with the extension SFO.

12.6.3.2 LOAD SPECIAL (*.SFO)

In this case the orbital energies and the molecular structure are read directly from the files CO1, CO2, or LO1, LO2 and SFO.

12.6.3.3 DELETE

This function allows you to cancel the graphic display of energy levels of molecular orbitals.

12.6.3.4 CALCULATION

This function determines the type of the calculation (orbital electron density, or orbital shape) for those orbital types that are displayed at the moment. Then the orbital to be calculated must be specified.

- Canonical orbitals
  - Orbital electron density
    - Total density
    - HOMO
    - Selected orbital
  - Shape of the orbital

        - HOMO
        - LUMO
        - Selected orbital
    - Localized orbitals
        - Orbital electron density
            - Total density
            - Selected orbital
        - Shape of orbitals
            - Selected orbital

12.6.3.5 ENERGY

Displays the energy of orbitals in the window in the lower right corner of the screen. The displayed value refers to the orbital marked with the white arrow. The arrow can be moved by clicking another orbital or using the arrows placed above and below orbitals of the selected molecule.

12.6.3.6 STRUCTURE

Display the selected molecule.

12.6.3.7 VIEW

Specify the orbitals to be displayed (CMO, LMO, or HOMO and LUMO canonical orbitals only).

12.6.3.8 ZOOM

Scale the image of orbitals. After this function has been selected, two buttons with arrows appear. By means of these buttons, it is possible to specify a subregion of orbitals to be displayed in the whole window.

12.6.3.9 RESET

Reset to the original scale of display.

12.6.3.10 INFO

Information about orbitals for the selected molecules.

12.6.3.11 HELP

Description of functions of this module.

12.6.4 COLOR

(see ***3***)

12.6.5 GEOMETRY VIEW

(see ***4***)

12.6.6 FIT

(see ***5***)

12.6.7 INTO XY PLANE

(see ***1.1.5***)

12.6.8 INTO Z-AXIS

(see ***1.1.6***)

12.6.9 MOVE ALONG Z-AXIS

(see ***1.1.7***)

12.6.10 PRINTER SETUP

(see ***6.9***)

12.6.11 S

Printing of the drawn orbitals.

12.7 HENRY'S CONSTANT

Calculation of the logarithm Henry's law constant (HLC) (air-to-water partition coefficient). The HLC is expressed as ratio VP/WS where VP is the vapor pressure
in [mm Hg] and WS is solubility in [mg/l] as calculated in the program.

Two methods are used:

First method[9] is based on the bond contribution method. The method uses 59 bond contribution values and 15 correction factors. The contributions were calculated by least-squares method using a 345-chemical data set. The mean error (average difference between the known log(HLC) and the bond estimated log(HLC)) for the least-squares analysis was 0.30 log units, standard deviation was 0.45 log units and correlation coefficient was 0.94.

In the ***Second method***[10] the log(HLC) is calculated from the equation found using linear regression:

$$\log(HLC) = -0.34\ X + \text{sum}(Gi) + 0.40$$

$$(n = 229,\ r = 0.994,\ s = 0.20\ )$$

where X is the first-order connectivity index, and sum(Gi) is the sum of atom and group contributions. The multifunctional compounds were omitted from this study. This method should not be used for the compounds where distant polar interactions are recognized.

12.8 MARK ATOM BY ATOM

This function allows you to mark atoms that will be included in calculations in modules ***12.1, 12.2, 12.3, 12.4***, and ***12.5***. To mark an atom, click on it using the mouse. Repeated clicking will remove the mark. The marking can be made visible by switching the atomic label switch into the function ***MARK***. This function is useful when processing of a part of a molecule or that of a system of molecules is required. When the function ***PROPERTIES (12)*** is started, all atoms are taken as not marked.

12.9 MARK ALL ATOMS

The function allows you to mark all atoms on the screen using a single command.

12.10 DELETE ALL MARKS

The function allows you to remove marks from all atoms using a single command.

13 *MM2 CALCULATIONS*

This module provides for geometry optimization and conformation analysis using molecular mechanics. The MM2 force field potential functions are employed. The corresponding force constants are stored in the text file MM2-PAR.TAB. If needed, the original values of force constants can be changed or supplemented by means of any text editor. The following restrictions apply to eventual supplementing of MM2-PAR.TAB:

- max. number of bond force constants . 300
- max. number of angle force constants 500
- max. number of angle force constants for 4-membered cycles 50
- max. number of force constants for torsion 900
- max. number of force constants for torsion in 4-membered cycles 50
- max. number of out of plane force constants 100

The format, which is apparent from the structure of the MM2-PAR.TAB file, must be observed. The MM2 atom types included in the MM2-PAR.TAB must not be altered or added. This is connected to the fact that for each new MM2 type, there must be the possibility of its transformation to *Molgen* types provided in the program.

Warning: Exceeding the limits given above, or an incorrect format of newly written force constants may cause problems. Therefore, it is advisable to make a backup copy of MM2-PAR.TAB file before editing.

13.1 MONITORING-E

This function allows you to optimize the geometry of a molecule and to search for low-energy conformations.

13.1.1 CHOICE OF MOL.

See Section ***2.10.2***

13.1.2 FIXED PARAMETERS

To select the atoms, the coordinates of which will remain constant during the optimization. Fixing of atoms brings certain simplification into the calculation as the number of degrees of freedom is decreased and thus the speed of the calculation is increased. It is recommended to fix only those parts of molecules that do not contain free torsional

rotations and the geometry of which has been calculated or experimentally determined by reliable methods.

13.1.2.1 FIX ATOM BY ATOM

See Section ***2.10.4***

13.1.2.2 FIX ALL ATOMS

This function allows you to fix all atoms of an actual molecule using a single command. The function is useful when the fixation of a larger part of a molecule is required. Then it is feasible to fix all atoms by a single command and to subsequently release those atoms, the coordinates of which are to be optimized, by the function ***FIX ATOM BY ATOM (13.1.2.1)***. The fixed atoms may be highlighted by toggling the switch of atom marks into the function **FIXED**. Fixing of atoms or whole parts of a molecule decreases the demands for computing in an optimization. This procedure is especially suitable in the cases when it is not necessary to optimize a whole molecule.

13.1.2.3 UNFIX ALL ATOMS

This function provides the user with the possibility to release all the fixed atoms using a single command.

13.1.3 ROTATION BOND

Monitoring of the molecule's internal energy connected with a rotation about a bond or about a joint of two atoms in a ring.

13.1.3.1 ROTATION BOND

Selection of atoms forming a bond with the possibility of rotation for conformation analysis or optimization. After you have clicked a color field, you are prompted to mark a pair of atoms forming a bond with possible rotation. That means that the selected bond must not be part of a ring. The selected bond will be marked by an 'R' mark of the color of the selected field. The torsion angle value is displayed simultaneously with the rotation. The number is of the color of the selected field and of that of the 'R' mark. If you select

the same color field again and you mark another pair of atoms, the marking of the original pair will be canceled. The maximum number of pairs monitored at a time is eight. The atoms forming the monitored torsion angle can be found as follows; the second and the third atoms are those marked by yourself, and the first and the fourth atoms are the adjacent atoms with the lowest sequential number. To display the atomic numbers, switch atom labeling into the mode **[NUMBER]**.

13.1.3.2 DISTANCE

Selection of atoms, of which the distance is to be monitored for conformation analysis or optimization. After you have clicked a color field, you are prompted to mark a pair of atoms, the distance of which you want to monitor. The selected atoms will be marked by an '@' mark of the color of the selected field. The value of the distance is displayed simultaneously with the rotation. The number is of the color of the selected field and of that of the 'R' mark. If you select the same color field again and you mark an other pair of atoms, the marking of the original pair will be canceled. The maximum number of pairs monitored at a time is eight.

13.1.3.3 STEP

Selection of an increment (in degrees) of rotation for conformation analysis. Click any of the displayed rotation increment values. The selected value will be marked by a white triangle.

13.1.3.4 ROTATION IN CYCLE

This function allows you to modify the geometry of rings by means of flipping about joints of atoms. If this function has been selected before choosing rotations, the user is prompted to mark three atoms by a color of a color field. The joint of the first and the second atom represents the axis of rotation and the third atom specifies the part of the molecule that is to be rotated. Clicking the 'R'-mark will rotate the defined part of the molecule. Left and right mouse buttons will rotate the part in opposite senses. After each step, the total energy of the molecule is calculated and displayed. If there are any rotations in cycle selected for a molecule, the optimization of rotation bonds by ***OPT. ROTATION***

BONDS is not available. Then it is necessary to cancel the selection.

13.1.3.5 UNSELECT ALL

Cancellation of all defined rotation torsions or rotations in cycles.

13.1.4 CONFORM ANALYSIS

This function provides the conformational analysis using the molecular mechanics method with the MM2 force field. The system automatically carries out an exhaustive conformational search about the identified chemical bonds. The energies of conformers are calculated in the approximation of solid rotor. Before starting this module, the chemical bonds about which torsional rotations can

take place, and rotation increment must be defined.

13.1.4.1 NAME OF OUTPUT FILE

Name of the file with a standard extension CON where the results from conformational analysis will be written. The directory can be chosen by the ***PATH*** option. The default name is CONFORM.CON.

13.1.4.2 ENERGY LIMIT

A value of the energetic limit up to which the conformers are considered. As basic energy, the energy of the conformation with the lowest energy is taken. Default is 5 kcal/mol.

13.1.4.3 ROTATION ON SCREEN YES/NO

If **[YES]**, the rotations will be animated on the screen.

13.1.4.4 GRAPH TO WINDOW <YES/NO>

If **[YES]**, the conformational maps of one or two rotation bonds will be displayed after conformational analysis.

13.1.4.5 PATH

The choice of the directory where the CON file will be saved. For details see Section ***17.1***.

13.1.4.6 CANCEL

Return to the previous menu.

13.1.4.7 STANDARD

Accept the default options.

13.1.4.8 ACCEPT

Accept the actual options.

The conformation analysis can be interrupted by pressing ESC. In that case, the input information is stored in the file designated for the analysis results. The data included are the required rotation bonds, the increment of the rotation, the energetic limit and the number representing the method of dipole-dipole interaction calculation. The file thus formed can be processed by the module ***BATCH PROCESSING (13.2)***. The calculated conformers can be

displayed using the function ***CONFORMERS (2.12)***. If no rotation bond is selected, the calculation is not commenced. The system allows exhaustive conformational search for up to five rotation bonds.

13.1.5 FULL ENERGY

This function calculates the total energy of a chosen molecule using the built in molecular mechanics. The level of electrostatic interactions taken into account in calculations can be set using the function ***DIP-DIP INTERACT***. If more than one molecule is displayed on the screen, the user is prompted to select a molecule to be optimized. The molecule can be selected by means of the function ***CHOICE OF MOL (2.11.2)***. If electrostatic interactions are defined by ***DIP-DIP INTERACT*** the dipole moment will be calculated and displayed.

13.1.6 OPT. FULL ENERGY

This function carries on the optimization of geometry of a selected structure using the molecular mechanics with MM2 force field. If more than one molecule is displayed on the screen, the user is prompted to select a molecule to be optimized. The molecule can be selected by means of the function ***CHOICE OF MOL (2.11.2)***. Only those molecules are accepted where for all atoms the MM2 atom types are defined, otherwise an error message is displayed. Electrostatic

interactions in the molecular mechanics used are approximated by the mutual interaction of bond dipoles within a molecule. If partial atomic charges are known (these may be obtained, for example, by quantum chemistry calculations), it is possible to use these values for electrostatic interactions' calculation. The level of the calculation of electrostatic interactions can be set using the function ***DIP-DIP INTERACT***.

13.1.6.1 1,3 NOT INCLUDED

The dipole moments of bonds at the same atom are not taken into consideration.

13.1.6.2 1,3 INCLUDED

The dipole moments of all bonds in a molecule are taken into consideration.

13.1.6.3 NOT CALCULATED

The dipole-dipole interaction is not taken into consideration.

13.1.6.4 FROM CHARGES

Partial atomic charges are used to calculate electrostatic interactions.

A part of a structure can be fixed by means of the function ***FIX ATOM BY ATOM (13.1.2.1)***; the whole structure can be fixed by using ***FIX ALL ATOMS (13.1.2.2).*** All atoms can then be released by means of the function ***UNFIX ALL ATOMS (13.1.2.3)***. The changes to the geometry made during the optimization process are displayed. The calculation can be interrupted by pressing ESC.

13.1.7 OPT. ROTATION BONDS

This function carries out the minimization of internal energy of a molecule with respect to the selected rotation bonds. The energy of a molecule, in approximation of a solid rotor, is calculated by molecular mechanics using the form of potential functions implemented in MM2 program (Allinger). The simplex method is employed for energy minimization. This method allows for 8 rotation bonds to be included in the operation. The calculation can be interrupted by pressing ESC. In searching for the global energetic minimum, it is advisable to alternate the manual modus with the automatic one.

13.1.8 CHANGE OF TYPES

To change atom types for molecular mechanics. Those atoms are interchangeable which have the same number and type of bonds. The new atom type must not be in conflict with the method for determination of MM2 types. For example, the type N.am can not be exchanged for the type N.pl3, because the amide nitrogen type N.am follows from the presence of carbonyl group.

13.1.9 DIP-DIP. INTERACT

The function sets the type of dipole-dipole interaction for MM2 calculations. The electrostatic interactions in the applied molecular mechanics are approximated by the mutual interaction of bond dipoles in a molecule. If partial charges at atoms are known (those may be obtained from quantum chemistry calculations), it is then possible to calculate the electrostatic interaction employing these charges. The level and the modus of the calculation of electrostatic interactions can be set by means of the function ***DIP-DIP INTERACT***.

13.1.9.1 1,3 NOT INCLUDED

The dipole moments of bonds at the same atom are not taken into consideration.

13.1.9.2 1,3 INCLUDED

The dipole moments of all bonds in a molecule are taken into consideration.

13.1.9.3 NOT CALCULATED

The dipole-dipole interaction is not taken into consideration.

13.1.9.4 FROM CHARGES

Partial atomic charges are used to calculate electrostatic interactions.

13.2 BATCH PROCESSING

This function allows you to process files prepared in advance for conformation analysis and MM2 optimization (see Section ***2.10***) without any user intervention. Generally conformational analysis and optimization of large molecules represent time demanding computing; therefore it is convenient to prepare the related data files in advance and then work them up by batch processing. The files for conformational analysis can be generated by triggering and interrupting the module ***CONFORM ANALYSIS (13.1.4)***. In this case, the information regarding the selected rotation bonds, the step of rotation, and energetic limit, is written into a CON file. If the module ***BATCH PROCESSING*** finds this file on the disk, it will carry out the conformation analysis and the results (the calculated conformers and the corresponding energies) will be appended to this file. This will be repeated

until all files with extension CON in the actual directory are processed. The files for MM2 optimization can be generated by the function ***MM2 (2.10)***.

13.3 CONFORM. MAPS

This function is designed to provide a graphic output of conformation analysis of one and two torsional rotations. After this function has been started, a window appears containing basic functions for conformation analysis. All files from the directory for conformers with the extension CON are displayed. These files are generated during ***CONFORMATIONAL ANALYSIS (13.1.4)***. The path for this directory can be defined in the module ***CONFORMERS (12.2)*** using the function ***PATH (17.1)***. After a file has been selected, the corresponding graph is displayed.

13.3.1 CONFORMATIONAL ANALYSIS OF ONE TORSIONAL ROTATION

For this case, it is possible to display up to ten conformation analysis' results at a time. Particular graphs are distinguished by various colors. The scale is automatically adjusted after each loading of a new dependence so that complete graphs are always displayed, and it can be then manually set by means of numeric values typed or by using a cut-out activated by the function marked as 'I'.

13.3.1.1 QUIT

Quit the most recently used function. If there is no function in use at the time, this key terminates ***CONFORM. MAPS*** function.

13.3.1.2 COMMENT

A line of any text can be written to serve as a title for output to printer or plotter.

13.3.1.3 NEW GRAPH

Select a file to be loaded. Up to ten graphs can be displayed in a window. Particular graphs are distinguished by colors and in the left lower corner of the screen, color fields are displayed gradually, with each new graph a field of the same color appears. These fields serve as identifiers for individual graphs. If the function is in the automatic function, the scale on x and y axes is always adjusted so that entire graphs are visible. If the scale has been altered interactively **[I]** or by keyboard **[RANGE-MANUALLY]**, then the selected scale is set until the automatic function is selected again **[RANGE-AUTOMATICALLY]**.

13.3.1.4 DELETE

This function removes a selected graph from the window (not from the disk). If there are more graphs displayed, the program waits until you have specified the graph to be deleted (clicking its identifier). After the last graph has been deleted, the module for selecting and loading of files **[READ]** is activated automatically.

13.3.1.5 RANGE

This function allows you to change the scale of a graph.

13.3.1.5.1 CANCEL

Quit the function without accepting any changes.

13.3.1.5.2 MANUALLY

Set the scale according to values listed under Energy and/or Angle.

13.3.1.5.3 AUTOMATICALLY

The automatic function of display is activated. The scale will always be set so that all graphs are displayed in their entirety.

13.3.1.5.4 ENERGY

If the display is in the automatic function, the window "Range from" displays the lowest energy found in all displayed graphs and in the window "Range to", the highest value is shown. It is possible to enter other values here and to alter the scale of y-axis of displayed graphs.

13.3.1.5.5 ANGLE

The same functionality as for **ENERGY**, the only exception being that this function concerns torsion angles and the x-axis of the graph.

13.3.1.6 ENERGY

This function allows you to find out the energy and torsion angle for any displayed point on the graph. If there is more than one graph displayed, the key of this function remains depressed and the program waits until one of the graphs is selected by means of its identifier. Then the point with the lowest energy is highlighted automatically as a white circle with white auxiliary axes. At this point, a window is displayed with the corresponding energy and torsion angle. Clicking another point cancels the original values and the highlighting moves to the newly selected point. If there is more than one graph displayed, only the active one (its identifier is highlighted) can be worked with, the rest are inactive. Within this function, it is possible to alter the scales for graphs interactively and to enter the manual function. To return to the automatic function, use the function **[R]**.

13.3.1.7 CONFORMER

The conformer corresponding to the selected point in the displayed graph will be displayed in the window.

13.3.1.7.1 TO FILE

The displayed structure is saved on disk in *Molgen* format. The file can be then loaded using the function **[READ]** in the module ***INTERFACES (2).***

13.3.1.7.2 L+

Molgen atomic types are displayed for all atoms of the displayed structures.

13.3.1.7.3 L-

No atomic labels are displayed.

13.3.1.7.4 N

Atomic numbers are displayed.

13.3.1.7.5 +

Increases the size of the structure in the window.

13.3.1.7.6 -

Decreases the size of the structure.

13.3.1.7.7 ARROWS **<πθτυ>**

Rotate the structure in the window.

13.3.1.7.8 QUIT

This function returns to the function **ENERGY**. The next usage of the function **QUIT** terminates the function **ENERGY** in the case when only one graph has been loaded. Otherwise, only selection of the chosen graph is canceled and a new one can be

selected, or again, **QUIT** can be clicked and the function **ENERGY** terminated.

13.3.1.8 R. BARRIER

In general, this function returns the difference of energies corresponding to two selected points of a graph. It is possible to find the values of rotation barriers. If there is more than one graph loaded, the program waits until a graph is selected by means of its identifier. The first selected point is highlighted analogously as in the function **ENERGY**. After the second point has been chosen, the energy difference is depicted graphically. The values of both

points are displayed in windows and in a window between them, the value of the difference is shown. To leave the function, use **QUIT**.

13.3.1.8.1 R

Returns the display to ***AUTOMATIC*** function.

13.3.1.8.2 I

This function allows you to alter the display scale interactively. In the middle of the widow, a white rectangle appears. This rectangle specifies the part of the graph which will be displayed in the whole window. The position of the parallelogram can be adjusted simply by dragging and dropping it. Its size can be set either using the functions **[+]** and **[-]** or by dragging any of its corners and dropping. After defining the cut-out, it can be displayed using the function **[A]**. **[Q]** reverts to the original form of the graph.

13.3.1.8.3 P

This function activates or suppresses the display of individual points in all loaded graphs.

13.3.1.9 IDENTIFIERS

In conformational analysis of one rotation bond it is possible to load up to ten different graphs in a single window. The graphs are differentiated by colors. Certain functions (***DELETE***, ***ENERGY***, ***R. BARRIER***) require to select one of the loaded graphs. For this purpose, the color fields (identifiers) beneath the graphs are designed. Each graph has its identifier (of the same color). To specify a graph, click its

identifier, which will then be highlighted and from that point on, the graph will be active for the corresponding functions.

13.3.2 CONFORMATIONAL ANALYSIS OF TWO TORSIONAL ROTATIONS

An output of this kind of analysis is a map of isoenergetic lines, the values on x and y axes representing the torsion angles. The scale is set automatically so that the maps are displayed entire.

The energies are relative to the energy of the most stable conformer with zero energy.

13.3.2.1 COMMENT

A line of any text can be written to serve as a title for output to printer or plotter.

13.3.2.2 DELETE

This function removes a selected graph from the window (not from the disk). If there are more graphs displayed, the program waits until you have specified the graph to be deleted (clicking its identifier). After the last graph has been deleted, the module for selecting and loading of files (Read) is activated automatically.

13.3.2.3 RANGE

This function allows you to change the scale of a graph for both torsion angles, for maximum and minimum value of energy, and for the difference of energy of adjacent isoenergetic lines.

13.3.2.3.1 CANCEL

Quit the function without accepting any changes.

13.3.2.3.2 MANUALLY

Set the scale according to values listed under ***ENERGY*** and/or ***ANGLE***.

13.3.2.3.3 AUTOMATICALLY

The automatic function of display is activated. The scale will always be set so that all graphs are displayed in their entirety.

13.3.2.3.4 ANGLE Y

If the display is in the automatic function, then in the line **[RANGE FROM]**, the starting value of the torsion angle is displayed and in **[RANGE TO]** the final value. These values concern the rotation bond that was selected first.

13.3.2.3.5 ANGLE X

If the display is in the automatic function, then in the line **[RANGE FROM]**, the starting value of the torsion angle is displayed and in **[RANGE TO]** the final value. These values concern the rotation bond that was selected second.

13.3.2.3.6 ENERGY

If the display is in the automatic mode, then the value in **[RANGE FROM]** is the lowest value of energy of displayed isoenergetic lines and the value in **[RANGE TO]** the highest one. By changing these values, it is possible to select various regions of energies; for instance, the region of local minima that is usually considered up to 20 kcal/mol relative to the most stable conformer. After this region is displayed, it is substantially easier to find all these minima on the conformation map.

13.3.2.3.7 STEP

The value displayed in this window represents the difference between adjacent isoenergetic lines. That means that displayed are only those isoenergetic lines with energy being a multiple of this value ranging from the value in **[RANGE FROM]** to the value in **[RANGE TO]**.

13.3.2.4 ENERGY

This function allows you to find out the values of energy and both torsion angles for any point of the graph. Automatically, the point with the lowest energy is highlighted and the

values are displayed in a window. Selection of another point moves the window and the highlighting to the newly selected point (the values being updated). To make the points clearly visible, use the function **[P]**. In this mode it is possible to change the scale interactively using **[I]** that, at the same time, switches to the manual function. To restore the automatic function, use **[R]**.

13.3.2.5 CONFORMER

The conformer corresponding to the selected point of a graph can be displayed using the function.

13.3.2.5.1 TO FILE

The displayed structure is saved on disk in *Molgen* format. The file can be then loaded using the function **[READ]** in the module ***INTERFACES (2)***.

13.3.2.5.2 L+

Molgen atomic types are displayed for all atoms of the displayed structures.

13.3.2.5.3 L-

No atomic labels are displayed.

13.3.2.5.4 N

Atomic numbers are displayed.

13.3.2.5.5 +

Increases the size of the structure in the window.

13.3.2.5.6 -

Decreases the size of the structure.

13.3.2.5.7 ARROWS <πθτυ>

Rotate the structure in the window.

13.3.2.5.8 QUIT

This function returns to the function **[ENERGY]**. The next usage of the function **[QUIT]** terminates the function **[ENERGY]** in the case when only one graph has been loaded. Otherwise, only selection of the chosen graph is canceled and a new one can be selected, or again, **[QUIT]** can be clicked and the function **[ENERGY]** terminated.

13.3.2.6 ENERGY OF CONT.

This function allows you to find the energy of any displayed isoenergetic line. After this function has been activated, the program waits until a line is clicked. Then lines with the specified energy are highlighted green and in the lower part of the screen, and the following message appears: ”Energy of contour: (energy) kcal/mol“.

13.3.2.6.1 R

Returns the display to automatic function.

13.3.2.6.2 I

This function allows you to alter the display scale interactively. In the middle of the widow, a white rectangle appears. This rectangle specifies the part of the graph which will be displayed in the whole window. The position of the parallelogram can be adjusted simply by dragging and dropping it. Its size can be set either using the functions **[+]** and **[-]** or by dragging any of its corners and dropping. After defining the cut-out, it can be displayed using the function **[A]**. **[Q]** reverts to the original form of the graph.

13.3.2.6.3 P

This function activates or suppresses the display of individual points in all loaded graphs.

14 HELIX

This function generates polymer structures[6] in the form of a single or double helix from a monomer. The helix generation can be activated only if there is a single molecule displayed on the screen, representing the monomer of the polymer to be built. The method is designated, above all, for generation of monosaccharide helixes.

14.1 HELIX BUILDER

This function starts the generation of a polymer in the form of a single or double helix from a monomer. The geometry parameters necessary for the formation of the required shape are entered by means of the function ***CONFIGURATION (14.3)*** that is included in the same menu as ***HELIX BUILDER***. The parameters required for the generation of the spiral:

1) Total length of a helix

2) Number of monomer units distributed along this length

3) Number of windings of the helix

4) Angle (slope) of the plane given by three atoms in the monomer unit defined by the user, and the axis of the helix

5) Number of threads in the helix

The function is active only in the case when a single molecule is displayed with three atoms marked. The first and the second atom represent the sites by means of which the monomer units will be connected. The third atom completes definition of the plane of the molecule. The join of the first and the second atom determines the so called virtual bond, around which the helix is gradually winded while the required slope, with respect to the plane parallel with the helix axis, is being achieved. The atoms can be marked by means of the function ***CHOICE OF ATOMS (14.2)*** which is included in the same menu as ***HELIX BUILDER***..

14.2 CHOICE OF ATOMS

This function provides marking of three atoms to be used in the helix generation. The first and the second atoms marked by a circle represent the joints by which the monomers will be mutually connected. The third atom marked completes the definition of the plane of the molecule. The join of the first and the second atoms determines the so called virtual bond, around which the monomer is consecutively rotated with respect to the required slope of the plane of the monomer with the plane parallel to the helix axis. The atoms, by means of which the monomer is going to be polymerized, can have at most two valences. The third selected atom can be any atom from the molecule providing it is not collinear with the first two atoms. After selecting all three atoms and setting the required parameters, it is possible to generate the helix using the function ***HELIX BUILDER (14.1)***.

14.3 CONFIGURATION

This function provides interactive setting of parameters required for generation of a polymer in the shape of a helix. The menu contains the following options:

1) Total length of the thread in Ångströms
2) Number of monomers in the thread
3) Number of windings in the thread
4) Initial slope
5) Concluding slope (inactive)
6) Increment (inactive)
7) Helix type
8) Sense of rotation of the helix
9) Color atoms

The slope of the helix is defined as the angle between the monomer plane and the helix axis. The monomer plane is given by three atoms selected in ***HELIX BUILDER (14.1)***. The color spectrum used to distinguish monomers is predefined.

15 *DATABASE*

This module provides the possibility of communication with the *Molgen* database through structural or substructural search. It is not included as a part of the student version of the program. The *Molgen* database contains information on coded molecules. It includes optimized 3D structures with partial atomic charges mostly taken from AM1 calculations, experimental log(p) (logarithm of partition coefficient of n-octanol - water) that represent the lipophilic properties and molar refractivities that represent the polarizability of molecules. The database serves as source of different 3D fragments needed in building of the new molecule or as knowledge base of known log(p) and molar refractivities. The data in the database are accessible as 3D structures generated by *Molgen* (1).

15.1 STRUCTURE

This option provides the possibility of structural search through the *Molgen* database. If there is more than one molecule on the screen you will be asked to select the molecule to be searched in the database. Then you will be asked to define the manner of coding the selected molecule. There are two possibilities: **[WITH CHIRALITY]** or **[WITHOUT CHIRALITY]**. The former possibility causes generation of the full molecular code including chiral properties. The latter ignores molecular chirality during the structural search. The generated molecular code is compared with the codes in the database. If the specified search requirements are met the message ”Molecule is in the database“ will appear. After clicking **[CONTINUE]** the following information is displayed:

1. Name of the molecule
2. Literature
3. Note (conditions used in the log(p) measurement)
4. log(p) (logarithm of the partition coefficient n-octanol/water)
5. Molar refractivity
6. Hansch database (literature used as a primary source of the log(p) values)

15.1.1 STRUCTURE

Appears only if the 3D structure in the database exists. After clicking this option the stored structure is displayed in the screen and highlighted by a green color.

15.1.1.1 COLOR

Coloring of the displayed molecules.

15.1.1.2 GEOMETRY VIEW

Calculation of the geometry parameters.

15.1.1.3 FIT

Comparison of two structures by superposition.

15.1.1.4 EXIT

Return to previous menu. You will be asked which structure will be left on the screen. Click **[ORIGINAL]** to leave the original molecule or click **[FROM DATABASE]** to leave the molecule from database.

15.2 2D-QUERY

This option allows you to formulate the query for database searching and thus to find the molecular structures that fulfill the required conditions (e.g. contain the defined substructure). Up to 8 conditions can be included into query with the logical operator AND being automatically inserted. The operators OR and NOT are not available. Any field of the selected database can be used in the query (substructure, formula, etc.) and they can be combined arbitrarily.

15.2.1 SUBSTRUCTURE

If a substructure has not been included into query up to this point, this function will do it. In the opposite case, the substructure specified previously is excluded from the query.

15.2.2 SUMMARY FORMULA

This function allows you to include a summary formula. It can be entered in arbitrary order of atoms and atoms can be repeated (e.g. $C_6H_5CH_2COOH$). The function will test and sort the formula (the formula from the example will appear as $C_8H_8O_2$).

15.2.3 FIELDS OF DATABASE

After this function has been selected, a new window appears with all the fields of the database displayed as buttons. By pressing the corresponding buttons, the fields are included into query after further specification.

15.2.4 DETAILED QUERY

This function provides display of the query with particular values and formulae.

15.2.5 SPECIFIED QUERY

This window provides the description of the query with filed names and operators only.

15.2.6 CANCEL

Quit the query specification without a search.

15.2.7 HELP

Get information on particular functions.

15.2.8 SEARCH

Finish query specification and begin the search.

The substructure window displays the structure acting as substructure in substructural search (if one has been included into query). During the search, hydrogen atoms are ignored and in their places any atoms can appear in the structures resulting from the search.

15.2.9 LABEL

To change the labels of atoms on the screen.

15.2.10 ROTATION

To activate rotation.

15.2.11 TRANSLATION

To activate translation.

15.2.12 XY

Rotation or translation within the screen plane.

15.2.13 Z

Rotation around the Z-axis (perpendicular to the screen plane), translation in the Z-axis direction.

15.2.14 ZOOM

To activate zooming.

15.2.15 CENTER

Position the molecule into the center of the window.

15.3 QUERY FIELD

Selection of the database fields for query. If a numeric type database field is selected, there is a possibility to specify an interval so that the retrieved compounds will have the corresponding characteristics within that interval. If a string type database field is selected, the following possibilities are available:

15.3.1 AS SUBSTRING

The query is a substring of a string in the database field.

15.3.2 EXACTLY

The query matches a string in the database field exactly.

15.3.3 CANCEL

To exclude a database field from query.

15.3.4 TO INCLUDE INTO QUERY

The database field will be included into query.

15.4 SUBSTRUCTURAL SEARCH

This function allows you to edit the hits from a substructural search.

15.4.1 EXIT

To abandon this function.

15.4.2 LABEL

To change the labels of atoms on the screen.

15.4.3 SAVE AS

To save the molecule displayed in the main window.

15.4.4 INFO

Output of a complete set of the information regarding the described molecule.

15.4.5 INFO ALL

Text file that contains information from all hits (can be used in statistics).

15.4.6 ROTATION

To activate rotation

15.4.7 TRANSLATION

To activate translation.

15.4.8 XY

Rotation or translation within the screen plane.

15.4.9 Z

Rotation around the x-axis (perpendicular to the screen plane), translation in the z-axis direction.

15.4.10 ZOOM

To activate zooming.

15.4.11 CENTER

Molecule will be positioned in the center of the window.

15.4.12 SPECIFIED QUERY

To display the current query.

The window in the left lower corner displays the count of hits. The molecule can be displayed by clicking any number in this window. The movement of the molecules is carried out by clicking any point in the windows with displayed structures. The type of movement is toggled by options **[ROTATION]**, **[TRANSLATION]**, **[XY]** and **[Z]**.

15.5 OUTPUT TO…

Selection of the database fields to be included in output. The first click activates the desired field to be included into the output set. The second click of the same field will remove the field from output set. Considered fields will be arranged into columns.

Example of an output set:

Specified query

Summary formula: C_1 H_1 N_1 O_1 (Atoms exclusively)

log(p): 1.0000 <= x <= 1.5000

No.	Summary formula	log(p)	Name of molecule
1.	$C_{10}H_{13}N_1O_2$	1.000	m-Methoxy-N,N-dimethylbenzamide
2.	$C_6H_7N_1O_1$	1.000	4-Methoxypyridine
3.	$C_4H_9N_1O_2$	1.010	2-Methyl-2-nitropropane
4.	$C_{11}H_{14}N_4O_1$	1.010	2-(4-Carbamylbutyl)benzotriazole
5.	$C_{10}H_{11}N_1O_3$	1.010	N-Methyl-4-acetylphenylcarbamate
6.	$C_8H_8N_2O_3$	1.020	N-Methyl-2-nitrophenylcarbamate
7.	$C_9H_{12}N_2O_1$	1.020	1,3-Dimethyl phenyl urea

8. $C_8H_9N_1O_2$ 1.030 p-Methoxyformanilide

15.6 AUTO-INPUT

This function allows you to feed selected MOL files to a database. Only users' databases can be fed as no changes are allowed to the standard *Molgen* database. To create a new database, use the function ***NEW DATABASE (15.10)***. To fill a database with structures from a selected directory, do the following: using the option **[CHOOSE]** database, select one of user defined databases. Then activate the function auto-input. A window with the directory tree appears. Select the directory containing the MOL files you want to be fed into the database. Click **[ACCEPT]** and the files will be loaded. During this process, the structures are checked for the presence of features not acceptable for a database (e.g. the presence of a dummy atom).

15.7 NOTES

This function allows you to enter new entries and edit old ones in the database of notes. If the selected database contains a field for notes, it is possible to fill it exclusively by means of the database of notes. A note can consist of up to 5 lines by 50 characters. The notes in the standard database (MOLGENDB) refer mainly to the conditions of measurements of the partition coefficient (log(p)). It is not possible to make changes to this database.

15.8 LITERATURE

Adding of new entries to the literature database and editing of old ones. If the selected database contains a field for literature, this can be filled exclusively by means of the database of literature. An entry of a database of literature can consist of up to 2 lines by 50 characters. The literature for the standard database (MOLGENDB) represents a reference to measurement of partition coefficient (log(p)). It is not possible to make any changes to this database.

15.9 CHOOSE DATABASE

Choice of the database to be processed.

15.9.1 CANCEL

Quit the function.

15.9.2 USER DATABASE

Select one of databases generated using the function ***NEW DATABASE (15.10).***

15.9.3 MAIN DATABASE

Selection of the main database supplied with *Molgen*. Items in this database are updated by the authors exclusively within the regular overall program updates. The data is stored in the file MOLGENDB.MDA.

15.10 NEW DATABASE

This function allows you to create new database. After it has been activated, a window appears with a prompt to specify a name for the database. Using the button **[PATH]** the user can specify a directory to contain the new database. After the button **[ACCEPT]** has been clicked, the new database is created. To fill it, use first the function ***CHOOSE DATABASE (15.9)*** (select the database you have just created) and then the function ***AUTO-INPUT (15.6)***.

The general functions are those which are accessible from each module of *Molgen*. These functions provide for movements of structures, labeling of atoms, and removing of structures.

16.1 EXIT (quitting a function or the program)

This function exits any function of *Molgen* and ends the program, if used from the main menu.

16.2 HELP

This function provides more detailed information on any function of *Molgen*. Clicking this function will highlight its box and the help is provided on the next function clicked. The corresponding texts are stored in HLP files.

16.3 ROTATE/TRANSLATE

The rotation or translation movement can be accomplished by clicking inside the grid box composed of concentric circles in the lower left corner of the screen. To achieve a continual movement, keep the mouse button pressed. To the right from the grid box, there is a scaled box. This device allows you to rotate a structure about the z-axis (the axis perpendicular to the plane of the screen). The boxes surrounding the grid box allow you to rotate a structure by 90 degrees. The position of the cursor in the grid box defines the movement in the following way. In the function ***ROTATION***, the joint of the cursors position with the center of the grid box defines the axis of the rotation. Rotation about z-axis is available using the box to the right from the grid box. Depending on whether the function ***GLOBAL*** or ***RELATIVE (16.14)*** has been selected, the molecules rotate as a whole about the center of the main window (***GLOBAL***), or a selected molecule rotates about its center of gravity (***RELATIVE***). If the function ***SPECIAL*** has been selected, all displayed structures rotate about their own centers independently. The distance of the cursor from the center of the grid box determines the speed of the rotation. The boxes surrounding the grid box control the rotation in the following manner:

top rectangle	rotation about y-axis to the right
bottom rectangle	rotation about y-axis to the left
left rectangle	rotation about x-axis to the right
right rectangle	rotation about x-axis to the left

The rectangle placed above the scaled box for the rotation about the z-axis rotates a structure about the z-axis to the right, the rectangle under the scaled box rotates a molecule in the opposite direction. In the function ***TRANSLATE***, the joint of the cursors position and the center of the grid box defines the direction of the translation movement. Translation in the direction of the z-axis is provided in the scaled box to the right from

the grid box. In the function ***GLOBAL (16.14)***, all structures are moved; in the ***RELATIVE*** function, only the selected molecule is translated. The distance of

the cursor from the center of the grid box determines the speed of the translation. The boxes for rotation by 90 degrees are inactive in this function.

16.4 DEL-M (delete molecule)

After this function has been selected, its window is highlighted. To delete a molecule, click any of its atoms. The function is active until ***EXIT (16.1)*** is clicked. If there are several molecules displayed on the screen and all of them are to be deleted, it is advisable to use the function ***M-UNITE (16.15)*** to unify the structures into a disjoint molecule and then the function ***M-DEL (16.4)*** will remove all the structures at once.

16.5 CEN-M (center molecule)

This function places the formal gravity center of a molecule (or all molecules) into the center of the coordinate system (the center of the main window). The formal gravity center is the point, the coordinates of which are obtained as the average value of x and y coordinates of all atoms of the centered object regardless of the atomic mass. After this function has been selected, its box is highlighted prompting the user to pick the molecule to be centered. To quit the function, click ***EXIT (16.1)***.

16.6 CEN-A (center atom)

This function places the selected atom of a structure into the center of the coordinate system (the center of the main window). It is used in the same way as function ***CEN-M (16.5)***. In the function ***GLOBAL (16.14),*** upon rotation, all molecules rotate about this atom.

16.7 RESET (restore the original coordinates)

The function reverts the values of the displayed structures' coordinates to the original ones. If the function ***RELATIVE (16.14)*** is active, only the coordinates of the selected molecule are reset. This function is useful when the original geometry of a structure is required after unwanted changes. The function ***M-UNITE (16.15)*** transforms the coordinates of all displayed molecules to the default.

16.8 INFO-M/INFO-SYS (information on a molecule/program environment)

After this function has been selected, its box is highlighted prompting you to specify the molecule on which the information is required. The following data will appear after the molecule has been clicked:

- Number of atoms
- Number of bonds
- Total charge of the molecule
- Molecular formula
- Name of the file from which the structure has been read

- Molecular mass
- Cartesian coordinates, bonds, atomic and bond types
- Atomic charges (if present in the source file)
- User defined types (if defined by the function USER'S TYPE)

If the data exceed the page, the display continues after clicking a mouse button. It can be interrupted by the function ***EXIT (16.1)***. If an INF file with the name of the source MOL file is present in the default directory, then its content is displayed as well. These files, which contain basic information on molecules, are included in the *Database Of Molecules* which is supplied with *Molgen*.

16.9 W+ (W-) (enlarge/shrink auxiliary window)

Providing that the auxiliary window is displayed (see the functions ***FIT (5), ATOM ADD (1.3), FRAGMENT LOAD (1.6), GROUP LOAD (1.9)***). Clicking inside the ***[W+]*** or ***[W-]*** box will enlarge or decrease the size of the auxiliary window, respectively.

16.10 S+ (S-) (enlarge/shrink structure)

This function allows to control the size of displayed structures in the main or auxiliary window. Clicking inside the ***[S+]*** or ***[S-]*** box will increase or decrease the size of the displayed structures, respectively.

16.11 SPREAD

This function distributes the displayed structures within the screen regularly.

16.12 NORMAL/QUALITY (mode of display)

Two regimes of display are available; ***NORMAL (16.12)***, which is a stick model with bonds being represented as joints of atoms; or ***QUALITY (16.12)*** which is a stick model with bonds being displayed as thicker lines whereby their visibility is taken into consideration. To toggle between the modes, click inside the appropriate box. The function ***QUALITY (16.12)*** is useful when a better orientation in overlapping structures is required.

16.13 LABEL+ (LABEL-, NUMBER, CHARGE, USER'S, HETERO)

The function allows to toggle between various modes of atomic labeling. The following exclusive functions are available:

- **LABEL+** *Molgen* types
- **LABEL-** No labels

- **USER'S** User defined types
- **NUMBER** Sequential atomic numbers
- **CHARGE** Charge if calculated, otherwise zeros will appear
- **HETERO** *Molgen* types labels on hetero atoms (not carbon and hydrogen)
- **H-BOND** Display hydrogen bonds
- **(DO NOT) DRAW H** (Do not) display hydrogen atoms
- **PERSPECTIVE** Drawing of structures by perspective

To switch between particular modes, click inside the appropriate box. In the window that has appeared afterwards, select the required mode, eventually switch on the display of hydrogen bonding.

16.14 GLOBAL/RELATIVE/SPECIAL

This function allows you to toggle between global and relative movements on the screen by clicking inside the appropriate box. In the ***GLOBAL (16.14)*** mode, any movement concerns all displayed structures. If ***RELATIVE (16.14)*** has been selected, the user is prompted to select the structure to be moved. This selection can be interrupted using **[EXIT]**. In the function ***ROTATE*** with the option ***SPECIAL*** all displayed structures rotate about their own centers.

16.15 M-UNITE (create a disjoint molecule)

This function allows you to unify all displayed molecules into a disjoint molecule. The disjoint molecule represents an object composed of several molecules that can be dealt with as with a single structure (coloring, display, deleting, saving, etc.). The function is useful when, e.g., the user wants to store the result of a ***SUPERIMPOSITION (5)*** to disk.

16.16 M-BREAK (disconnect a disjoint molecule)

This function acts in the opposite way as the function ***M-UNITE (16.15).*** It breaks joint molecules into individual molecules. It can be used when creating two molecules from one structure.

16.17 S (save)

This function (active only in the module ***MODELS (6)***) allows you to save a displayed image on disk. After this function has been clicked, the user is prompted for a filename. The image will be re-drawn to the whole screen and

then written to the file. It can be then screened using the module ***VIEW PICTURE (6.10)***.

16.19 EDITOR

The built-in editor provides standard possibilities of creating or editing text files. The commands/data for the editor are:

Max. number of columns per row	159
Max. number of rows	2000

CURSOR MOVEMENT COMMANDS

Character left	CTRL-S or τ
Character right	CTRL-D or υ
Line up	CTRL-E or π
Line down	CTRL-X or θ
Scroll up	CTRL-W
Scroll down	CTRL-Z
Page up	CTRL-R or PGUP
Page down	CTRL-C or PGDN
Go to line number	ALT-S G or ALT-R or choice by mouse
Go to col number	ALT-C or choice by mouse
On 1st row of window	CTRL-HOME
On last row of window	CTRL-END
On 1st row of file	CTRL-PGUP
On last row of file	CTRL-PGDN
On 1st col of current row	HOME
On last col of current row	END

INSERT & DELETE COMMANDS

Insert mode on/off	CTRL-V or INS
Insert line	CTRL-N

Delete line	CTRL-Y
Delete to end of line	CTRL-Q Y
Delete character left	CTRL-H or BACKSPACE
Delete character	CTRL-G or DEL

MISCELLANEOUS EDITOR COMMANDS

Exit editing	ESC or ALT-X or function Exit
Open an old file	F3 or function Open
Create a new file	function New
Save current file	F2 or function Save
Save as new filename	function Save as
Find specified text	CTRL-Q F or function Find
Change of path for files	function Path
Select keyboard	function Keyboard
Output to printer	CTRL-P or function Print
Help text for editor	F1 or function Help

17.1 SELECT DIRECTORY, CHANGE DIRECTORY

The following operations are available: select path for default subdirectory, browse directory tree, make directory, remove directory, view contents of directory. After this function has been called, a window appears displaying the directory tree of the specified drive. The default directory is highlighted in red and is placed in the middle of the window. Besides the menu, the information on the selected method or function (*Molgen*, Sybyl, Fragment and the like) is displayed. The menu contains the following:

17.1.1 CANCEL

Return to *Molgen* without accepting any changes.

17.1.2 ACCEPT

Confirm changes and return to *Molgen*. To change the default directory, click on a new one, which will then be highlighted in red, and then use this option.

17.1.3 REVERT

Change to the original directory.

17.1.4 MAKE/FIND

Left mouse button: Create a new directory as a subdirectory of default directory.

Right mouse button: Find the directory.

17.1.5 DELETE

Delete default directory. Only empty directories can be deleted.

17.1.6 FILES

View contents of default directory. The extension corresponds to the selected type of files. After this option has been selected, another menu appears:

17.1.6.1 CANCEL

Return to *Molgen* without accepting changes.

17.1.6.2 PATH

Return to previous menu.

17.1.6.3 DELETE

Delete selected file from disk.

17.1.6.4 ZOOM

Change the font size.

17.1.6.5 TO WINDOW

Change the size of the window.

17.1.6.6 DRIVE

Left mouse button: select drive; right mouse button: re-scan the directory tree (necessary after changing the directory

structure outside the *Molgen* environment).

17.1.7 ZOOM

Change the size of the font.

17.1.8 TO WINDOW

Change the size of the configuration window.

17.1.9 PATH/HISTORY

Left mouse button: returns complete DOS path. It is possible to edit the path and thus to change the directory or drive. The right mouse button lists up to 15 previously employed directories.

17.1.10 ARROWS <πθτυ>

Clicking the arrows will move the tree in the direction of the particular arrow.

17.2 CREATING STRUCTURES

The molecular structures can be generated from atomic types provided by *Molgen* and from structures created previously and stored on disk. These can be fragments (files with the extension FRG) and functional groups (GRP files). After an ***ATOMIC TYPE (1.3)*** has been selected, its structure appears in the auxiliary window; i.e. the geometry of its unfilled valences through which it can be attached to the structure being built. The number of the valences and the values of angles between them correspond to the hybridization of the corresponding atomic type. If a ***FRAGMENT (1.6)*** or a ***GROUP (1.9)*** has been opted for, the structure appears in the auxiliary window like in the case of atomic types. These structures may have been created and saved before. The attachable valences are marked with a red tilde '~'. The green asterisk '*' denotes the selected bond *via* which the structure will be attached to the one in the main window. This bond can be changed by clicking any of the unfilled valences marked with tilde '~'. A structure in the auxiliary window can be rotated, translated in all directions within the xy plane and in z axis. It can be positioned to the center of the window using the function ***[CEN-M] (16.5)***. The size of the auxiliary window can be adjusted by means of the function ***[W+],***

[W-] (16.9). A structure can be transferred from the auxiliary window to the main one in two principal situations:

1) *There is no molecule in the main window.* Clicking anywhere in the main window will place the structure in the center of the window.

2) *There is a structure in the main window already.* If the structure is to be attached, the structure in the main window must have at least one unfilled valence. The unfilled valence can be created, e.g. by ***DELETE AN ATOM (1.4)***). To attach the structure, click an unfilled valence of the structure in the main window. Both valences must be of the same

type. After the substructure has been connected, the auxiliary window disappears, but the substructure remains in the memory ready for another attachment, until another substructure is selected or the function ***[EXIT] (16.1)*** is used. Thus, by subsequent clicking, it is possible to build a structure by adding a sub-unit until the maximum number of displayable atoms is reached. After a substructure has been attached, an 'R' mark appears over the newly formed bond. Clicking this symbol will rotate the new part of the molecule. Left and right mouse buttons rotate the unit in opposite directions. The newly formed bond has standard length. If there is a substructure in the auxiliary window, but instead of an unfilled valence another place in the main window is clicked, the auxiliary window is removed as well, leaving the substructure remaining in the memory.

17.2.1 LOAD FILE

The files of *Molgen* format can be loaded by pressing ***ALT-R*** directly without having to activate the function ***INTERFACE (2.14)***. In main menu, module ***INTERFACES*** and ***BUILDER*** it is also possible to load *Molgen*-type file by pressing the right mouse button twice. The following menu is available for this function:

17.2.1.1 CANCEL

Return to *Molgen* without loading a file.

17.2.1.2 ACCEPT

Load the selected file. Alternatively, double-click on the filename.

17.2.1.3 PATH

Change the path.

17.2.1.4 DELETE

Delete the selected file.

17.2.1.5 ZOOM

Change the font size for name of the directory or name of the file.

17.2.1.6 TO WINDOW/TO SCREEN

Change the size of the configuration window.

17.2.1.7 EDIT

See Section ***16.18***.

17.2.1.8 ARROWS <πθτυ>

Pressing an arrow will move the list of files in the corresponding direction.

17.2.2 SAVE FILE

After this function has been selected, the user is prompted for a filename. Then the structure is written to the specified file. If a file with the given filename exists already, the user is notified. Then the file can be overwritten or the operation canceled. A structure can be saved in *Molgen* format by pressing ***ALT-W*** directly without having to activate the function ***12.15***.

APPENDIX A

An example of a structure written in Molgen format

```
24 ATOMS, 26 BONDS, 0 CHARGES, FOZKIO
1     O3          0.9702      -4.7509     2.3809      0.0000
2     C3          0.0674      -4.3910     3.4494      0.0000
3     C3          -0.9654     -3.3695     2.9933      0.0000
4     C3          -0.4622     -1.9584     2.7033      0.0000
5     C3          0.1464      -1.7920     1.3077      0.0000
6     NAM         1.2002       -2.7543    1.0978      0.0000
7     C2          1.9438      -3.0816     0.0226      0.0000
8     C3          2.7286      -4.1203     0.8085      0.0000
9     C3          1.8360      -3.7201     2.0058      0.0000
10     I          -2.3788     -3.2520     4.6345      0.0000
11    O           21.9255      -2.6576    -1.1145     0.0000
12    C2          0.7128      -0.3887     1.2090      0.0000
13    O2          1.6883      -0.0147     1.8021      0.0000
14    O3          -0.0042     0.3923      0.4107      0.0000
15    C3          0.4374      1.7754      0.3518      0.0000
16    CAR         -0.1860     2.4804      -0.7997     0.0000
17    CAR         -0.7760     1.8256      -1.8749     0.0000
18    CAR         -1.2908     2.5554      -2.9433     0.0000
19    CAR         -1.2068     3.9416      -2.8920     0.0000
20    CAR         -0.6249     4.6120      -1.8399     0.0000
21    CAR         -0.1249     3.8667      -0.8011     0.0000
22    NPL3        -1.7316     4.7289      -4.0043     0.0000
23    O2          -1.7418     5.9397      -3.9199     0.0000
24    O2          -2.1632     4.1330      -4.9877     0.0000

1     1     2     SINGLE
2     1     9     SINGLE
3     2     3     SINGLE
4     3     4     SINGLE
5     3     10    SINGLE
6     4     5     SINGLE
7     5     6     SINGLE
8     5     12    SINGLE
9     6     7     SINGLE
10    6     9     SINGLE
11    7      8    SINGLE
```

12	7	11	DOUBLE
13	8	9	SINGLE
14	12	13	DOUBLE
15	12	14	SINGLE
16	14	15	SINGLE
17	15	16	SINGLE
18	16	17	AROMATIC
19	16	21	AROMATIC
20	17	18	AROMATIC
21	18	19	AROMATIC
22	19	20	AROMATIC
23	19	22	SINGLE
24	20	21	AROMATIC
25	22	23	DOUBLE
26	22	24	DOUBLE

An example of a structure written in Sybyl format

```
24 MOL      FOZKIO      0

1     8     0.9702          -4.7509          2.3809
2     1     0.0674          -4.3910          3.4494
3     1     -0.9654         -3.3695          2.9933
4     1     -0.4622         -1.9584          2.7033
5     1     0.1464          -1.7920          1.3077
6     28    1.2002          -2.7543          1.0978
7     2     1.9438          -3.0816          0.0226
8     1     2.7286          -4.1203          0.8085
9     1     1.8360          -3.7201          2.0058
10    17    -2.3788         -3.2520          4.6345
11    9     1.9255          -2.6576          -1.1145
12    2     0.7128          -0.3887          1.2090
13    9     1.6883          -0.0147          1.8021
14    8     -0.0042         0.3923           0.4107
15    1     0.4374          1.7754           0.3518
16    3     -0.1860         2.4804           -0.7997
17    3     -0.7760         1.8256           -1.8749
18    3     -1.2908         2.5554           -2.9433
19    3     -1.2068         3.9416           -2.8920
20    3     -0.6249         4.6120           -1.8399
21    3     -0.1249         3.8667           -0.8011
22    19    -1.7316         4.7289           -4.0043
23    9     -1.7418         5.9397           -3.9199
24    9     -2.1632         4.1330           -4.9877
26 MOL
1     1     2     1
2     1     9     1
3     2     3     1
4     3     4     1
5     3     10    1
6     4     5     1
7     5     6     1
8     5     12    1
9     6     7     1
10    6     9     1
```

```
11     7      8      1
12     7      11     2
13     8      9      1
14     12     13     2
15     12     14     1
16     14     15     1
17     15     16     1
18     16     17     5
19     16     21     5
20     17     18     5
21     18     19     5
22     19     20     5
23     19     22     1
24     20     21     5
25     22     23     2
26     22     24     2
0 MOL
```

An example of a structure written in CSD format

```
#FOZKIO  1388071814  9  0  0  0  1  4 24  0  0 28132100000020000000000088
7944 15908  6860  9034  8850 10263333222 2 6 1 2 2 3  0175  2P-1240
R=0.0680
211 0121 0112 0
C  68O  68N  68I 140
O1  72290  41101 06930   C2  61590  64301  22511 C3  51620  13010  115861
C4  62030  22100 111631 C5  70620  23170  91280  N6  81240  16970  88220
C7  90030  14860  72541  C8  96721  81708  4001    C9  86220  10750 101461
I1   033619 1376  9139793 O1191370  17590  55960 C1281740  32210  89841
O1394880  34620  98490 O1475180  37240  78200 C1584660  46150  77341
C1679180  50690  60550 C1770260  46470  44871 C1866190  51170  29291
C1971140  60100  30040 C2080010  64420  45381 C2183860  59620  60530
N2267120  65170  13820 O2370380  72970  15050 O2460330  61330-520
2 3 4 5 6 7 8 9 1 3 7 5121214151617181916192222 6 92021
```

An example of a molecule written in Cartesian coordinates

```
0.97020              -4.75090        2.380908
0.06740              -4.39100        3.449406
-0.96540             -3.36950        2.993306
-0.46220             -1.95840        2.703306
```

APPENDIX B

List of atomic types used in Molgen

LABEL	ATN	GEOM	COLOR	LN	VDWRA	CHEMMOD	SYBYL	A_WEIGHT
1. C3	6	TH	light_grey	2.5	1.52	1	1	12.01100
2. O3	8	BI4	red	3.5	1.36	8	10	15.99400
3. N3	7	PYR	sky_blue	3.0	1.45	5	6	14.00670
4. H	1	L1	dark_cyan	2.1	1.08	13	15	1.00797
5. Car	6	TG	light_grey	2.5	1.53	3	3	12.01100
6. C2	6	TG2	light_grey	2.5	1.53	2	2	12.01100
7. C1	6	L23	light_grey	2.5	1.54	4	4	12.01100
8 O2	8	L12	red	3.5	1.36	9	11	15.99400
9. N2	7	B23	sky_blue	3.0	1.48	6	20	14.00670
10. N1	7	L13	sky_blue	3.0	1.50	7	20	14.00670
11. Nar	7	BI3	sky_blue	3.0	1.48	11	7	14.00670
12. Nam	7	TG	sky_blue	3.0	1.45	28	20	14.00670
13. Npl3	7	TG	sky_blue	3.0	1.50	19	9	14.00670
14. N3+	7	TH	sky_blue	3.0	1.45	31	5	14.00670
15. Du	0	L1	magenta	0.0	0.00	26	20	0.00000
16. P3	15	TH	amber	2.1	1.75	12	14	30.97376
17. P3	D15	TH4	amber	2.1	1.75	12	20	30.97376
18. P4	15	TB	amber	2.1	1.75	26	20	30.97376
19. S3	16	BI4	yellow	2.5	1.70	10	12	32.06000
20. S2	16	L12	yellow	2.5	1.72	18	20	32.06000
21. SO	16	PY2	yellow	2.5	1.70	29	20	32.06000
22. SO2	16	TH4	yellow	2.5	1.70	30	13	32.06000
23. F	9	L1	green	4.0	1.30	16	16	18.99840
24. Cl	17	L1	green	3.0	1.65	15	17	35.45300
25. Br	35	L1	green	2.8	1.80	14	18	79.90400
26. I	53	L1	purple	2.5	2.05	17	19	126.90450
27. K	19	L1	magenta	0.8	2.80	22	20	39.09830
28. Na	11	L1	magenta	0.9	2.30	21	20	22.98977
29. FE	26	OH	magenta	0.9	2.30	26	20	55.84700
30. Li	3	L1	magenta	1.0	1.80	24	20	6.94100
31. Al	13	TH	dark_cyan	1.5	2.05	25	20	26.98154
32. Si	14	TH	yellow	2.8	2.10	27	20	28.08550
33. Ca	20	L2	magenta	1.0	2.75	23	20	40.09000

34. LP	0	L1	magenta	0.0	1.08	20	20	0.00000
35. ~	0	L1	pink	0.0	0.00	26	20	0.00000
36. *	0	L1	mint_green	0.0	0.00	26	20	0.00000
37. N2+	7	TG2	sky_blue	3.0	1.45	6	20	14.00670
38. B3	5	TH	yellow	3.0	1.45	26	20	10.81000
39. O3+	8	PYR	red	3.5	1.36	26	20	15.99400
40. C22	6	L22	light_grey	2.5	1.52	26	20	12.01100
41. N22+	7	L22	sky_blue	3.0	1.48	26	20	14.00670
42. O3-	8	L1	red	3.5	1.30	26	20	15.99400
43. P3D2	15	TH2	amber	2.1	1.75	12	20	30.97376
44. N1+	7	L23	sky_blue	3.0	1.50	7	20	14.00670

TYPE	ATN	LABEL	DESCRIPTION	MOLGEN EQUIVALENT
1.	6	C.3	CSP3	C3
2.	6	C.2	CSP2 ALKENE	C2
3.	6	C.nyl	CSP2 CARBONYL	C2
4.	6	C.1	CSP ALKYNE, C=C=O	C1
5.	1	H	EXCEPT ON N/O	H
6.	8	O.3	C-O-H, C-O-C	O3
7.	8	O.2	=O CARBONYL	O2
8.	7	N.3	NSP3	N3
9.	7	N.amd	NSP2 (AMIDE)	Nam
10.	7	N.1	NSP	N1
11.	9	F	FLUORIDE	F
12.	17	CL	CHLORIDE	Cl
13.	35	BR	BROMIDE	Br
14.	53	I	IODIDE	I
15.	16	S.3	-S- SULFIDE	S3
16.	16	S.3+	>S+ SULFONIUM	SO
17.	16	S.O	>S=O SULFOXIDE	SO
18.	16	S.O2	>SO2 SULFONE	SO2
19.	14	Si	SILANE	Si
20.	0	LP	LONE PAIR	LP
21.	1	H.ol	-OH ALCOHOL	H
22.	6	C.3cp	CYCLOPROPANE	C3
23.	1	H.n3	NH AMINE	H
24.	1	H.coo	COOH CARBOXYL	H
25.	15	P.3	>P- PHOSPHINE	P3
26.	5	B	TRIGONAL	Du
27.	5	B.3	TETRAHEDRAL	B3
28.	1	H.amd	VINYL ALCOHOL, AMIDE	H
29.	6	C.3.	RADICAL	C3
30.	6	C.3+	CARBONIUM ION	C3
31.	32	GE	GERMANIUM	Du
32.	50	SN	TIN	Du
33.	82	PB	LEAD (IV)	Du
34.	34	SE	SELENIUM	Du
35.	52	TE	TELLURIUM	Du
36.	1	D	DEUTERIUM	H
37.	7	N.2	-N= AZO, PYRIDINE	N2
38.	6	C.2cp	CSP2 CYCLOPROPENE	C2
39.	7	N.3+	NSP3 (AMMONIUM)	N3+
40.	7	N.pl3	NSP2 (PYRROLE)	Npl3
41.	8	O.fur	OSP2 (FURAN)	O3
42.	16	S.tio	SSP2 (THIOPHENE)	S3
43.	7	N.azy	-N=N-O AZOXY	N2

44.	0	XXXX	Unknown	Du	
45.	7	N.1+	AZIDE (CENTER-N)	Du	
46.	7	N.o2	NO2 (NITRO)	Npl3	
47.	8	O.3-c	CARBOXYLATE	O3-	
48.	1	H.n3+	AMMONIUM	H	
49.	8	O.epx	EPOXY	O3	
50.	6	C.ben	BENZENE	CAR	
51.	2	He	HELIUM	Du	
52.	10	Ne	NEON	Du	
53.	18	Ar	ARGON	Du	
54.	36	Kr	KRYPTON	Du	
55.	54	Xe	XENON	Du	
56.	6	C.3cb	CYCLOBUTANE	C3	
57.	6	C.2cb	CYCLOBUTENE	C2	
58.	6	C.Ocb	CYCLOBUTANONE	C2	
59.	12	Mg	MAGNESIUM	Du	
60.	15	P.5	P (V)	P3D2	
61.	26	Fe.2	IRON (II)	Du	
62.	26	Fe.3	IRON (III)		Fe
63.	28	Ni.2	NICKEL (II)	Du	
64.	28	Ni.3	NICKEL (III)	Du	
65.	27	Co.2	COBALT (II)	Du	
66.	27	Co.3	COBALT (III)	Du	
67.	6	C.Ocp	CYCLOPROPANONE	C2	
68.	0	XXXXX	UNKNOWN	Du	
69.	8	O.3-n	AMINE OXIDE	O3-	
70.	8	O.2ke	KETONIUM OXYGEN		O2
71.	6	C.2ke	KETONIUM CARBON		Du
72.	7	N.imi	IMINE, OXIME	N2	
73.	0	XXXX	UNKNOWN	Du	

LABEL Atomic labels of used atomic types. This is the labeling which appears in the function ***[LABEL+] (16.13)***. An atom can have more types, e.g. the carbon atom occurs in the following ones:

C3	sp^3 hybridized carbon
C2	sp^2 hybridized carbon
C1	sp^1 hybridized carbon
CAR	aromatic carbon
C22	allelic carbon (the carbon with two double bonds)

ATN Atomic number

GEOM Spatial configuration of bonds for the given atomic type

TH, TH2, TH4	tetrahedral
BI4	planar
PYR	pyramidal
L1, L12, L22	linear
TG, TG2, TG3	trigonal
OH	octahedral
TB	trigonal bipyramid

COLOR The color in which the atom is displayed on the screen.

ELN Electronegativity

VDWRA van der Waals' radius

SYBYL_TYPE The number of the atomic type in Sybyl program corresponding to the atomic type in *Molgen*.

A_WEIGHT Atomic weight

General information

The *Database Of Molecules* is a constituent part of *Molgen*. It contains optimized three-dimensional molecular structures. Its importance consists in the two basic functions:

1) It is a source of a wide spectrum of substructures for the building of new molecules. For most of the structures, all geometric parameters have been optimized (bond lengths, angles, and torsion angles). For the optimization, the AM1 quantum chemistry method has been used, providing reliable geometric parameters and electronic distribution. The selection of substructures for the *Database Of Molecules* has been done with respect to as great coverage of likely combinations as possible. With the help of the *Database Of Molecules* and *Molgen* it is easy to create sets of fragments and molecules tailored to the special needs of the user.

2) It is a source of information on the contained molecules. Two kinds of information are available:

   a) Those obtained from the quantum chemistry calculations (atomic charges, dipole moment, etc.).

   b) Those obtained from literature sources (melting points, density, occurrence, etc.).

The molecular structures are stored in files with the extension MOL and the corresponding information is in INF files.

Structure of the Database

The molecules contained in the *Database of Molecules* are divided into subdirectories according to structural types. The following groups are included (the actual names of directories are given in brackets):

1) Hydrocarbons (HYDRCARB)
    - Aromatic (AROMATIC)
    - Cyclic (CYCLIC)
    - Aliphatic (NOCYCLIC)
    - Spiranes (SPIRANES)

2) Derivatives (DERIVATE)
    - Acetals (ACETALS)
    - Aldehydes (ALDEHYDE)
    - Alcohols (ALCOHOLS)
    - Ethers (ETHERS)
    - Phenols (PHENOLS)
    - Halogen derivatives (HALOGEN)
    - Ketenes (KETENES)
    - Ketones (KETONES)
    - Acids (ACIDS)
    - Lactones (LACTONES)
    - Nitrogen containing derivatives (NITROGEN)
    - Sulphur (II) containing derivatives (SULPH_II)
    - Sulphur (IV) containing derivatives (SULPH_IV)
    - Sulphur (VI) containing derivatives (SULPH_VI)

3) Heterocycles (HETEROCY)
    - Aromatic (AROMATIC)
    - Non-aromatic (NOAROMAT)

4) Natural compounds (NATURAL)
    - Alkaloides (ALKALOID)

Alcohols (ALCOHOLS)

Aminoacids (AMINOACID)

Amines (AMINES)

Bioluminous materials (BIOLUMIN)

Ethereal oils (ETHEREAL)

Dyes (DYES)

Phytoalexines (FYTOALEX)

Hormones (HORMONES)

Nucleosides (NUCLEOSI)

Organic acids (ACIDS)

Carbohydrates (CARBOHYDR)

Vitamins (VITAMINES)

5) Silicon compounds (SILICON)

Silanes (SILANE)

Siloxanes (SILOXANE)

These basic groups are further divided to subgroups. The subdirectories are in the form of a directory tree. *Molgen* provides functions for convenient work with a directory tree (see ***PATH (2.7), IMPORT (2.8) and EXPORT (2.9)***).

Physical chemistry properties:

property	abbreviation	unit
heat of formation		kcal/mol
dipole moment		D (debye)
ionization potential		eV (electron volt)
melting point	mp	°C
boiling point	bp	°C
refraction index	nD[temperature in °C]	
density	h	g/cm^3

Other data can be given by means of:

1) keywords
2) text

Information from the *Database of Molecules* is available upon clicking the function ***INFO-M (16.8)***.

REFERENCES

1 *Ghose, A. K., Pritchett, A., and Crippen, G. M.*, Journal of Computational Chemistry 9 (1988) 80.

2 *Furet, P., Sele, A., and Cohen, N. C.*, Journal of Molecular Graphics 6 (1988) 182.

3 *Gasteiger J and Marsili M,* Tetrahedron, 36, 3219 (1980).

4 *Gavezzoti, A.*, Journal of the American Chemical Society 105 (1983) 5220.

5 *Silla E. et al.*, Journal of Molecular Graphics 8 (1990) 168.

6 *Bladon, P.*, Journal of Molecular Graphics 7, (1989) 130.

7 *French and Murphy*, Carbohydrate Research 27 (1973) 391.

8 *J. Gasteiger and M. Marsili,* Tetrahedron , 36 (1980) 3219.

9 *W.M.Meylan and P.H.Howard,* Environ. Toxicol. Chem. 10, (1991) 1283.

10 *T. Suzuki, K. Ohtaguchi and K. Koide,* Computers Chem. 16, 41 (1992).