NAME

q2m – convert cartesian coordinate table to Sybyl MOL file

SYNOPSIS

q2m [**-i** *inpfile*] [**-o** *outfile*] [**-s** *string*] [**-f** *optfile*] [**-n** *molname*] [**-d** *length*] [**-m** *unit*] [**-l** *link-list*] [**-u** *unlink-list*] [**-a** *maxatoms*] [**-b** *maxbonds*] [*inpfile* [*outfile*]]

DESCRIPTION

The program **q2m** converts a table of cartesian coordinates of atoms to a *SYBYL* **MOL** file format (*SYBYL* is a trademark of *Tripos Associates*). It creates appropriate atom types and bond types based on interatomic distances. The table of cartesian coordinates has to contain one column which identifies element either by its chemical symbol, or atomic mass, or atomic number. The MOL file so created can be read into *SYBYL* for display or further manipulation.

OPTIONs

Option arguments contain characters which are frequently conflicting with the shell interpretation. For safety, it is advised that inclose them in single quotes.

-i *inpfile*

File *inpfile* contains cartesian coordinates. If *inpfile* is not specified, standard input (stdin) is assumed. *inpfile* can be also specified without **-i** option letter as the first argument after all options.

-o *outfile*

File *outfile* will contain resulting **MOL** file. If *outfile* is not specified standard output (stdout) is assumed. *outfile* can be also specified without **-o** option together with *inpfile* after all options as the second argument.

-s *string*

The order and type of columns in the table is given by *string.* Column is a series of nonblank characters followed by one or more spaces or TABs. The following colum types are recognized: **M** - atomic mass, **A** - atomic number, **C** - chemical symbol (may be followed by digits or other characters), **S** - column to discard, **X** - X coordinate, **Y** - Y coordinate, **Z** - Z coordinate. Only one **M, A,** or **C** column may be specified. The **X, Y,** or **Z** columns cannot be repeated.

-f *optfile*

The options will be read from *optfile.* The *optfile* contains options in the same format as if they appeared on the command line.

-n *molname*

Using this option you can specify molecule name to be included in the **MOL** file or otherwise it will be set to "Unknown".

-d *length*

This is the *length* of longest possible bond [in Angstroms] in the molecule. All atoms closer to each other than *length* are considered bonded by a chemical bond. Exception is used for hydrogen atoms. Two hydrogen atoms are considered bonded only if they are closer than 0.8 Angstrom. If *length* is not specified 1.7 Angstrom is assumed.

-m *unit* The coordinates in the input file may be specified in *B* (bohrs, i.e. atomic units) *N* (nanometers) *A* (Angstroms). If this option is not specified, Angstroms are assumed.

-l *link-list*

Sometimes you want to include bonds which will not be found by a simple distance checking. You may specify a list of bonded atom pairs and a type of bond by which atoms are bonded. The *link-list* consist of entries of a type *ibj* separated by commas or spaces. *i* is the number of first atom, *j* is a number of second atom and *b* specifies a type of bond between them. The bond type *b* may one of the following: : (unspecified - when bond type is not known), - (single), = (double), # (triple), a (amide bond), r (aromatic), . (dummy - not a true bond, but marker rather), u (unusual - nobody yet knows what it is). This option may be used not only to add new bonds but also to specify a correct bond type for bonds which are found from interactomic distance comparisons. The *link-list* should be enclosed in single quotes for safety.

-u *unlink-list*
Sometimes you want to remove some bonds which were found by simple distance check. This option removes bonds between atom pairs specified by *unlink-list.* Atom numbers and their pairs can be separated by spaces or any other nondigit characters.

-a *maxatoms*
If your coordinate table may contain more than 100 atoms, you need to specify larger limit for number of atoms with this option.

-b *maxbonds*
If your molecule may contain more than 200 bonds, you need to specify larger limit for number of bonds with this option.

EXAMPLES

q2m -s SXYZM -m B -d 1.6 -l 1-4,1=2 -n propanal test.inp requests that file **test.inp** is converted to MOL file and sent to standard ouput. The coordinates are given originally in atomic units. The columns 2-4 contain cartesian coordinates, while 5-th column contains atomic masses. Atoms are considered bonded if they are closer than 1.6 Angstrom to each another. The single bond is specified between atoms 1 and 4, and double bond between atoms 1 and 2.

FILES

The program resides in /loclib/q2m directory. Examples of use are given in /loclib/q2m/test. The docs are in /loclib/q2m/doc.

SEE ALSO

To familiarize yourself with MOL file format read appropriate pages of SYBYL manual or use on-line help facility of SYBYL.

DIAGNOSTICS

Most error messages are self-explanatory.

BUGS

Program does not find some atoms types and you have to enter correct types by editing resulting **MOL** file. This program does not contain substructure searches necessary to match some of the SYBYL atom types, however, most basic atom types are assigned correctly. Report bugs to the author.

AUTHOR

Jan Labanowski, Ohio Supercomputer Center, 1224 Kinnear Rd, Columbus, OH 43212-1163, ph. 614-292-9279, FAX: 614-292-7168, e-mail: jkl@ccl.net, JKL@OHSTPY.BITNET.