T125
66

To the Graduate Council:

I am submitting herewith a thesis written by Roswitha T. Haas entitled "Typesetting Chemical Structure Formulas with the Text Formatter TeX/LaTeX." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Computer Science.

Kevin C. O'Kane, Major Professor

We have read this thesis
and recommend its acceptance:

Accepted for the Council:

Vice Provost
and Dean of The Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at The University of Tennessee, Knoxville, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of the source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature Roswitha T. Haas

Date July 2, 1986

TYPESETTING CHEMICAL STRUCTURE FORMULAS WITH THE TEXT FORMATTER TEX/LATEX

A Thesis

Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Roswitha T. Haas

August 1986

ACKNOWLEDGMENT

I appreciate very much the helpful advice and encouragement given to me by my thesis director, Dr. Kevin C. O'Kane.

I also recognize with gratitude all my computer science instructors, especially Dr. Charles P. Pfleeger, Dr. David W. Straight, and Mrs. Jane Wallace Mayo, for all they have taught me.

ABSTRACT

TeX is a widely used document formatting system. Its capabilities for mathematics typesetting make it attractive to use TeX in many scientific and technical fields. A TeX version that is intended to be fully useful for typesetting chemistry documents has to make the printing of chemical structure formulas about as convenient as the printing of complex mathematical expressions. The work described in this thesis constitutes a possible approach to enhancing TeX with structure-drawing capabilities.

The LaTeX picture environment, a simple graphics facility which is part of the LaTeX package of TeX macros, was used to draw the bond lines of chemical structures with branches and ring systems. TeX's integer variables and conditional facility provided the flexibility to draw many different structures from a limited number of basic fragments. Thirty-five macros were written to print these fragments – the most commonly occurring branching patterns, alicyclic systems, and heterocycles. Instructions are given for combining fragments from several macros to larger structures and to chemical equations. Diagrams of print quality are produced with a laser printer as output device.

This approach to generating printed chemical structure diagrams by computer has three main advantages: A device-independent ASCII file is sent to the output device, no graphics hardware is needed, and the highly acclaimed text formatting features of TeX are available to design the document containing the structure diagrams.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
	1. TeX/LaTeX's Capabilities for Typesetting Chemistry Texts	1
	2. The Output Technology	3
	3. Other Work with Computer-Generated Chemical Structure Diagrams	4
II.	ESSENTIAL FEATURES OF DOCUMENT PREPARATION WITH TEX AND LATEX	7
	1. Text Formatting with TeX	7
	2. The LaTeX Version of TeX	9
III.	TEX/LATEX CODE FOR COMPONENTS OF ORGANIC CHEMICAL STRUCTURE DIAGRAMS	11
	1. Conventions for Drawing the Diagrams	11
	2. Bond Line Drawing and Positioning	13
	A. Review of TeX/LaTeX Facilities for Line-Drawing	13
	B. Bonds in Structural Formulas Written on One Line	15
	C. Bonds in Acyclic Structures with Vertical Branches	16
	D. Bonds in Structures Containing Slanted Bond Lines	18
	3. Atomic Symbols and Condensed Structural Fragments	21
IV.	MACROS FOR TYPESETTING CHEMICAL STRUCTURE FRAGMENTS	24
	1. General Organization of a Structure Macro	24
	A. Box Constructions with Global Variables	25
	B. Use of TeX's Conditional Facility	27
	2. Use of Structure Macros	29
	A. Invoking the Structure Macros	29
	B. Displaying Macro-Generated Diagrams Within a Document	31

V. COMBINING STRUCTURES FROM SEVERAL MACROS ... 35
1. General Considerations for Combining Structures in this System ... 35
2. Combining Fragments to Form a New Structure ... 38
A. Attachment by the Subpicture Method ... 38
B. Attachment by Shifting the Coordinate System ... 42
3. Aligning Structures in an Equation ... 44

VI. THE COMPLETE SYSTEM OF MACROS – ITS DESIGN AND ITS USE ... 47
1. General Design Criteria ... 47
2. Macros of the System ... 48
A. Macros for Acyclic Fragments ... 48
i. Macro \cbranch[9] ... 48
ii. Macro \tbranch[7] ... 49
iii. Macro \ethene[4] ... 50
iv. Macro \upethene[4] ... 50
v. Macro \cright[7] ... 50
vi. Macro \cleft[7] ... 51
vii. Macro \chemup[7] ... 52
viii. Macro \cdown[7] ... 52
ix. Macro \csquare[5] ... 53
x. Macro \ccirc[4] ... 53
B. Macros for Alicyclic Ring Systems ... 54
i. Macro \threering[9] ... 54
ii. Macro \fourring[9] ... 54
iii. Macro \fivering[9] ... 55
iv. Macro \sixring[9] ... 56
v. Macro \sixringa[9] ... 57
vi. Macro \sixringb[9] ... 57
vii. Macro \chair[8] ... 58
viii. Macro \naphth[9] ... 59
ix. Macro \terpene[9] ... 59
x. Macro \steroid[9] ... 60
C. Macros for Heterocyclic Ring Systems ... 62
i. Macro \hetthree[8] ... 62
ii. Macro \hetifive[9] ... 62
iii. Macro \heticifive[9] ... 63
iv. Macro \pyrazole[8] ... 65
v. Macro \hetisix[9] ... 65

vi. Macro \pyrimidine[9] 66
vii. Macro \pyranose[9] 67
viii. Macro \furanose[8] 68
ix. Macro \purine[9] 68
D. General Utility Macros 70
i. Macro \fuseiv[9] 70
ii. Macro \fuseup[9] 71
iii. Macro \fuseiii[6] 72
iv. Macro \cto[3] 72
v. Macro \sbond[1] 73
vi. Macro \dbond[2] 73
vii. Macro \tbond[2] 73
3. Common Requirements for the Use of the System 74

VII. CONCLUSION 77

BIBLIOGRAPHY 80

APPENDICES 83

APPENDIX A. LATEX CODE FOR MACROS 84

APPENDIX B. COORDINATES OF POINTS OF ATTACHMENT 118

APPENDIX C. SLOPES OF LINES DRAWN WITH LATEX 120

VITA 121

LIST OF FIGURES

FIGURE PAGE

3.1: One-line structural formulas 16

3.2: One-line structure in picture environment 16

3.3: Vertical branches 17

3.4: Calculating position and length of double bond 19

3.5: Diagram using \multiput 20

3.6: Scaling(unitlength=0.07pt, 0.08pt, 0.15pt) 20

3.7: Right-justification of substituent formula 23

4.1: \cright{CH_3}{Q} {CH}{S}{$COOCH_3$}{S}{$COOCH_3$} 30

4.2: Structures drawn with the \cright macro 30

4.3: \sixring{OH}{Q}{Q}{Q} {NC}{Q}{CH_3}{NH}{D} .. 31

4.4: (S)-lactic acid 33

4.5: (R)-lactic acid 33

4.6: LaTeX code for two captions in one figure 33

5.1: Finding coordinates of points of attachment 37

5.2: Nicotine with LaTeX code 38

5.3: Combining ring structure and acyclic subpictures 39

5.4: Combining acyclic subpictures 40

5.5: Condensed formula fragment between rings 40

5.6: User-designed connecting bond line 41

5.7: Fusion of fully drawn rings 42

5.8: Vertical attachment by coordinate shifting 44

5.9: Alignment in a chemical equation 46

6.1: Statements at the beginning of a LaTeX file 75

CHAPTER I

INTRODUCTION

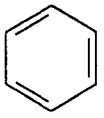
1. TEX/LATEX's CAPABILITIES FOR TYPESETTING CHEMISTRY TEXTS

TEX (TeX) is a computer language for document formatting developed by Donald Knuth (Knuth 84). TeX is in the public domain and widely used. The official 1985 TeX Users List includes 914 individuals and 76 institutions. The same list indicates that 26 different computers, mainframe and micro, and about 80 different output devices are used to run TeX and to produce the printed documents. One of the stated purposes of the TeX system is the typesetting of mathematical formulas. The American Mathematical Society has established a TeX Users Group and has published the semiannual newsletter called TUGboat since 1980.

In a collection of survey articles on document preparation systems (Nievergelt 82, p. 153), TeX is named as one of the three most interesting and influential pure formatting systems in current use; Scribe and the Unix-based formatters are judged to be the other two. This reference credits TeX in general with the capability of producing documents with the finest possible appearance, and specifically with superior models for the control of spacing and positioning, and for paragraph layout, as well as with a more unified specification language than that of other formatters (Nievergelt 82, pp. 153,171,190,191,194,206).

The fact that TeX can produce professional-looking documents, ready for the phototypesetter, with text containing mathematical formulas, makes it attractive to use TeX for typesetting in many scientific and technical fields. Chemists are using TeX as shown by a few entries in the users list. A survey of some of the publications of these authors established that they were dealing with the more mathematical branches of chemistry. To the knowledge of this

writer, there is no systematic published attempt to use TeX for typesetting chemical structure diagrams, especially those of organic chemistry, such as benzene:



The purpose of this thesis is to fill this gap and enhance TeX in such a way that its use for typesetting structure diagrams becomes practical and relatively convenient for an author or a secretary.

It is obvious that rudimentary graphics capabilities are necessary to produce most organic chemical structure diagrams. As a minimum, the system should be able to draw lines with slopes other than those of vertical and horizontal lines.

TeX as such does not have this feature. Graphics instructions can be included into a TeX file via a so-called special file. But with such a special file the TeX output file, the device-independent (DVI) file, could lose its device independence as an ASCII file. Knuth addresses the graphics issue in the TeXbook (Knuth 84, p. 229) where he states "The author anticipates that certain standards for common graphic operations will emerge in the TeX user community after careful experiments have been made ...." — Various approaches to graphics standards for TeX were recently discussed in the TeX users' newsletter (Spragens 85). Users from many scientific and technical fields are interested in merging text and graphics.

However, the LaTeX (LaTeX) version of TeX (Lamport 86) has simple graphics capacities. LaTeX adds to TeX a system of macros with the main purpose of simplifying the use of TeX for frequently occurring formatting tasks. The important LaTeX contribution for this thesis is the "picture environment,"

in which the user can specify a coordinate system and place objects into it. The objects include slanted lines, circles, and ovals. These are drawn with fonts that were designed for this purpose and are supplied with the LaTeX macros. Thus LaTeX graphics do not compromise the device independence of the TeX DVI file, they are generated by TeX itself, rather than at the level of the output device (Spragens 85, p. 68). This was the main reason for using LaTeX graphics for this thesis, despite some limitations that will be explained together with specific applications.

Since LaTeX is indispensable for this study, the convenient LaTeX macros are used wherever possible in producing the structure diagrams, even for tasks that TeX could have done without LaTeX. Where special features from plain TeX are used, it is stated explicitly.

2. THE OUTPUT TECHNOLOGY

The full benefits of sophisticated software such as TeX/LaTeX can only be realized with proper output technology. A laser printer, the IMAGEN Page Printer with a Canon LBP-CX laser engine, was used in this thesis. Its resolution of 300 dots/inch is high enough to give the printed text a typeset appearance and to produce smooth lines for the structure diagrams. By comparison, the resolution of direct impact dot matrix printers varies from 10 to about 160 dots/inch and jagged edges are distinctly noticeable on characters and lines. The resolution of an ink-jet printer is around 200 dots/inch.

Laser printers are also fast — the IMAGEN's specifications are eight pages per minute — , they are quiet, and they accept a variety of different types of paper, even transparency sheets. They can be used with mainframe or micro host computers.

Laser printers are now affordable for many offices and other users. They have already been called high quality typewriter replacements (Christian 85).

More important, the new output technology, combined with text formatting software, puts it within reach of the users to produce camera-ready copy of articles and even books. A substantial reduction in publication costs can result. As the high-resolution printers become more widely spread, more and more users will get to know formatting tools such as TeX . The enhancement to TeX developed in this thesis should be useful to some of them.

3. OTHER WORK WITH COMPUTER-GENERATED CHEMICAL STRUCTURE DIAGRAMS

Computer-generated structure diagrams are used in chemical information storage and retrieval, in visualizing and modeling chemical interactions, and in document preparation.

In many chemical information storage and retrieval systems, notably in the online version of the Chemical Abstracts Service (Dittmar 83), the structure diagram of a query result is displayed for the user along with the name and other properties of the substance. Often the user can also enter structures or substructures at the terminal as queries. In this application, each chemical substance is stored as a unique representation, in most cases as a connection table that lists for each atom other atoms to which it is connected. Ash's book is a recent reference that discusses these structure systems (Ash 85, chapter 5).

Where the study of chemical structures and reactions is the reason for generating and manipulating the diagrams, the structures can also be stored as unique connection tables, but more generic forms of internal representation are possible. A bibliography for molecular graphics for the years 1983/1984 (Morf-few 84) lists about 100 references, most of which deal with this area of computer modeling for chemistry. — Programs for microcomputers are available to display structures in the context of computer-assisted chemistry instruction. One approach that is taken for this purpose is to store the more common structures

and fragments as "shapes" — collections of lines, arrows, wedges, etc. (Zimmerman 84).

Generating structure diagrams for document preparation is the application that is of interest to this study. In principle, hard copy structure diagrams can be obtained from any of the applications mentioned before. However, there are no provisions to integrate the diagrams with text other than by cutting and pasting, and furthermore the aesthetics of the diagrams are not a primary concern as they should be in publishing documents.

There are also programs that are designed to generate structure diagrams specifically for reports and presentations, but do not attempt to integrate the diagrams with text (Watt 85, Kao 85). At least one commercial software package for microcomputers produces only the diagrams for publications, to be blended with separately generated text.

Commercial scientific word processors are the type of software most closely related to the software developed by this study. This writer knows of five packages that permit direct integration of structure diagrams into text. In a review of special-purpose word processors (Dickinson 86), four of these are evaluated with respect to quality of word processing, quality of graphics, printer requirements, documentation, and other criteria. All the scientific word processors are designed for personal computers, and with all of them the structures are "drawn" by the user on the screen, with a mouse or with crosshairs, by selecting graphics symbols, including larger structure fragments, from a menu. Thus, they require special hardware, graphics board, graphics monitor, and mouse, for the input of structure diagrams. The pricing of these software packages is also noteworthy, the most elaborate one sells for $ 1500, the least costly one for $ 495.

There are several reasons why a TeX/LaTeX chemical structure facility, as developed in this study, fills a need not satisfied by the commercial word

processors described above: TeX is a superior text- and mathematics formatting tool as discussed in section 1. of this chapter, TeX is running on a greater variety of computers, it produces device-independent ASCII files, a capability claimed by just one of the scientific word processors, and finally it is available at a much lower cost since it is in the public domain.

The TeX/LaTeX graphics are not entered by drawing on the screen, instead, LaTeX code is used to represent the graphics elements. This can be viewed as an advantage and as a disadvantage. It is an obvious advantage that no special graphics terminals are needed. In a publication on a graphics language for the Unix formatters (Kernighan 82, p. 19), more subtle advantages are mentioned: A linear, textual representation is well suited to machine manipulation and can convey information about the relationships of graphics objects that may get lost in an interactive system. — The disadvantage of having to input the chemical structure in the form of a line notation is the fact that the user can not see directly, during input, whether the diagram looks right.

CHAPTER II

ESSENTIAL FEATURES OF DOCUMENT PREPARATION WITH TEX AND LATEX

1. TEXT FORMATTING WITH TEX

An input file that is submitted to the TeX program for processing is considered to be a list of "tokens" by the program. A token can be either a single character or a so-called control sequence. As the single characters are read, one of fifteen category codes is associated with each character and it is processed according to its category. The control sequences are previously defined commands; they are identified as such by their first character, called an escape character. The backslash character is commonly used as the escape character.

Among the single characters, the letters of the alphabet, the digits, and the punctuation marks are the most frequently occurring ones; they are typeset as the respective characters in the style of the current font. — There are ten characters which are reserved for special purposes by TeX, the characters \ { } $ & # ^ _ ~ %. The backslash was mentioned above, { and } are used to begin and end a "group" of input text, $ symbols enclose input to be processed in math mode, & is used to set tabs, # designates a parameter in a macro, ^ and _ produce superscripts and subscripts respectively, a ~ instead of a space in the input text prevents a line break at this point, and the % causes the rest of the input on that line to be ignored. When it is desired to print these special characters as such, the input has to be \backslash \{ \} \$ \& \# \^ _ \~ \%.

Control sequences can have two formats. One form is called a control symbol and consists of the backslash followed by a single nonletter. Thus, nine of the special characters discussed in the previous paragraph can be printed

with control symbols. Many accents for foreign languages are also produced by control symbols. — The second type of control sequences is called a control word and consists of the backslash followed by one or more letters, followed by a space to mark the end of the control word. Control words are often used to produce characters that are not available from the keyboard, especially characters for mathematical expressions. Thus, \pi and \infty cause the symbols π and ∞ to be typeset. —Other control words are instructions for the formatting of the document. In this category, \indent starts a new paragraph and \vskip.5cm adds .5cm of extra vertical space after the current line. — As far as the TeX program is concerned, control sequences are either primitives of the TeX language or they are defined in terms of primitives. There are about 300 primitives and about 900 control sequences altogether in the standard version of TeX. Any number of control sequences can be defined by the individual users.

To understand the processing of a list of tokens, it is important to be familiar with the TeX concept of a "box." Tex forms (imaginary) boxes on several levels. First, a box is formed for each printable character; this is the space where the output device will actually put the character. These boxes are then combined into a so-called horizontal box that will produce a line of text. The space between words in a horizontal box is stretchable and shrinkable which makes it possible for TeX to produce lines of equal width. The TeX term for this variable space is "glue." In turn, completed horizontal boxes are put together into vertical boxes until the vertical box of a paragraph or a page is full. The user would normally not interfere with TeX's box-making, unless it is desired to process a chunk of text in a special way. Then a construct such as \hbox{some text} can be applied to prevent a linebreak within this string of text, to raise or lower the text with respect to the line, or to treat the text in the box different from the surrounding text in other ways.

As TeX forms boxes from token lists it can be in different "modes." There

are basically three modes: horizontal, vertical, and math mode. TeX spends most of the processing time in horizontal mode, forming lines and paragraphs. When the user indicates a new paragraph by inserting a blankline into the input file, TeX switches to the vertical mode, forming a vertical box and attaching it to the previous vertical box on the page with the right amount of vertical glue. — Normally, the user does not have to be aware of the horizontal or vertical modes, unless an error message points out that a TeX feature was attempted in the wrong mode. — Math mode is requested explicitly by the user, through enclosing the respective input within $ symbols. Many keyboard characters have special interpretations in math mode, letters now denote italic letters, a hyphen (-) becomes a minus sign (−). Special characters such as Greek letters become available. All spacing in the input file is ignored and TeX follows its own rules for mathematics spacing. — The user can override these features of math mode when necessary.

TeX features that are specifically applicable to the project of this thesis are explained in detail in the respective chapters.

2. THE LATEX VERSION OF TEX

Each frequent user of TeX is expected to define individual control sequences (macros) and packages of control sequences for specific applications. This makes the use of TeX easier and more efficient, and it results in input files that reflect the logical structure of the document. The LaTeX version of TeX contains many of the definitions that various groups of users can apply to their formatting tasks. LaTeX macros are available for all levels of a document: words, sentences, paragraphs, sections, chapters, and the document as a whole.

As an example of a LaTeX macro dealing with short text strings, the LaTeX \mbox is very similar to TeX's \hbox, which is a primitive. An effect such as right justification of the string in the box can be generated by studying

and combining information from various parts of the TeXbook. LaTeX considers right justification one of the more commonly needed effects and makes it available as an option: \mbox[width][r]{text}.

For formatting larger amounts of input, the LaTeX definitions of "environments" are very convenient. The user encloses the input material within the keywords \begin{environment name} and \end{environment name} and LaTeX handles all the formatting specific for the environment automatically. There are environments for itemizing, enumerating, typesetting verse, and typesetting matrices and tables. The picture environment mentioned in chapter I will be discussed in the next chapter and used constantly throughout this thesis.

Progressing to still larger parts of a document, LaTeX offers macros for generating section- and subsection headings, including incremented numbering within a chapter. Numbering of footnotes, tables, and figures is also incremented by LaTeX.

The style of the document as a whole can be selected as article-, report-, and bookstyle. Chapter headings can be generated by the \chapter macro for reports and books, but not for articles. — Finally, there are options to make LaTeX generate cross references, and auxiliary files for producing a bibliography, a table of contents, and a listing of entries for the index.

As was stated for TeX, LaTeX features specifically applicable to this thesis are discussed in detail in the respective chapters.

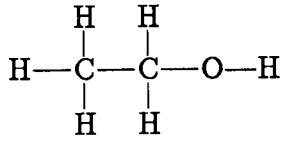
CHAPTER III

TEX/LATEX CODE FOR COMPONENTS OF ORGANIC CHEMICAL STRUCTURE DIAGRAMS

1. CONVENTIONS FOR DRAWING THE DIAGRAMS

The chemical structure of a molecule is defined by the spatial arrangement of the atoms and the bonding between them. Chemists use several standard methods for representing the structures two-dimensionally by diagrams called structural formulas; and this thesis will develop mechanisms for printing such diagrams using the TeX/LaTeX system.

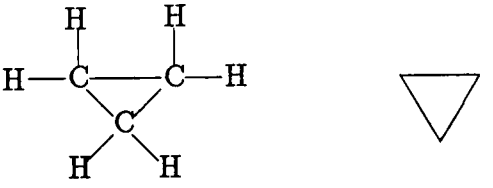
A very common structure representation, sometimes called a dash structural formula, uses the element symbols for the atoms and a dash for each covalent bond in the compound. Thus the dash represents the pair of shared electrons that constitutes the bond. Two dashes (=) represent a double bond and three dashes (≡) a triple bond. — It is usually neither necessary nor practical to represent each bond in a molecule explicitly by a dash. Some molecules and some bonds are so common that a complete dash formula would not be used except at a very introductory level of presenting chemical information. A condensed structural formula is one alternative. It does not contain dashes but uses the convention that atoms bonded to a carbon are written immediately after that carbon and otherwise atoms are written from left to right in the order in which they occur in the real structure. The following two structural formulas are a dash formula and a condensed formula, respectively, for the same compound, ethanol.



CH_3CH_2OH

Multiple bonds are usually not implied unless a very common group, such as the cyano group, is shown. It can be found as -C≡N or simply as -CN.

Another alternative to a complete dash formula is a diagram where the symbols for carbon and for hydrogen on carbon are not shown. Each corner and each open-ended bond in these diagrams implies a carbon atom with as many hydrogen atoms bonded to it as there are free valences. This representation is the customary one for ring structures (structures with a closed chain of atoms). Thus, the following two diagrams both represent the compound cyclopropane.



The three different kinds of structure representation can be combined in one diagram, such that in part of the diagram all bonds are represented by dashes and all atoms by an element symbol, in another part a condensed structural formula fragment is used, and in still another part a cyclic fragment with implied carbon and hydrogen atoms occurs.

The rest of this chapter describes how LaTeX can be used to position and typeset the bond lines and condensed formula strings that are the components of structure diagrams.

It should be mentioned that there are no binding rules for many aspects of the two-dimensional representations of a chemical structure. Structures and fragments of structures can be oriented in different ways depending on the availability of space, the emphasis given to a certain part of a structure, or the spatial relationship of the parts to each other. Thus, a cyclopropane ring can be represented in various orientations, $\triangle$, $\triangledown$, and others. Also, the angles between the bond lines can be different in different representations of one and the same

compound. Since most molecules do not have all their atoms lying in one plane it would not even be possible to reproduce all bond angles in a two-dimensional representation. The structures shown in this thesis adopt the orientations and bond angles found to prevail in Solomons' textbook (Solomons 84), the organic chemistry text used for several years at the University of Tennessee.

There are some methods to indicate the real, three-dimensional structure (the stereochemistry) of a molecule in the two-dimensional representation: A dashed line and a wedge instead of a full bond line mean that the real bond extends below or above the plane, respectively.

2. BOND LINE DRAWING AND POSITIONING

A. Review of TeX/LaTeX Facilities for Line-Drawing

The easiest way to produce horizontal and vertical lines representing chemical bonds is by the use of keyboard characters and simple control sequences provided by TeX. By typing one, two, or three hyphens, a normal hyphen, a medium dash designed for number ranges, and a punctuation dash are produced, - – —, respectively. When a hyphen is typed in TeX's math mode, it is interpreted as a minus sign and the spacing around it will be different from text mode. — The equal sign can represent a double bond for chemistry typesetting. It can be typed in text mode and in math mode, again resulting in different spacing around the symbol. — The control sequence \equiv can be used as a triple bond (≡). It has to be typed in math mode.

Vertical lines are available through the keyboard character or the control sequences \vert and \mid, all three to be entered in math mode. A double vertical bar is produced by \| or \Vert, again both in math mode.

The spacing around all these symbols can be controlled by adding extra (positive or negative) space with the horizontal spacing commands. The symbols, just as any other part of a line, can also be raised or lowered respective

to the normal baseline. The length and height of the symbols however depend on the font currently in use.

Where control of length and height of the bond lines is needed, TeX's or LaTeX's command sequences for printing horizontal and vertical "rules" can be used. The systems recognize several length units, including the inch, centimeter, millimeter, and printer point (Knuth 84, p. 57). One printer point (pt), an often used unit in typesetting, measures about 0.35 mm. — LaTeX's rule-printing command has the format

\rule[raise-length] {width} {height} .

Thus it can be used to produce horizontal and vertical rules. Using the \rule command one can also print multiple bond lines of user-controlled length, e. g. $=$, $\equiv$, with the short control sequences \dbond and \tbond defined in this thesis. The vertical spacing between the bonds depends on the current line spacing in the document and may have to be adjusted. The control sequences are set up for math mode.

When bond lines other than horizontal and vertical ones are to be printed, and when a coordinate system is needed to control placement of structure components relative to one another, LaTeX's picture environment (Lamport 86, pp. 101-111) is a necessity.

A picture environment uses length units which are dimensionless and have to be defined by the user before entering the environment. This is done by the \setlength command. In this study, \setlength {\unitlength} {0.1pt} is the definition used for most diagrams. Such a small unitlength was chosen to have fine control over the appearance of the diagram.

The picture environment starts with the statement

\begin{ picture} (width, height)

where picture width and height reserve space on the page and are specified in terms of unitlengths. Optionally, one can include the coordinates of the lower

left corner of the picture:

\begin{ picture} (width, height) (x_i,y_i).

The default value for these coordinates is (0,0).— Objects are placed into the picture with the \put command with their reference point at the coordinates (x,y): \put(x,y) { picture object} .

The picture objects of most interest to this study are straight lines. They are drawn by the \line command:

\line(x_s,y_s) {length}

where the coordinate pair specifies the slope of the line, and the nonnegative value of length specifies the length of the projection of the line on the x-axis for all nonvertical lines, and the length of the line for vertical lines. The reference point of a line is one of its ends. Thus the statement

\put(x,y) {\line(x_s,y_s) {len} }

draws a line that begins at (x,y), has a slope of y_s/x_s, and extends for length len as explained above.

Only a limited number of slopes is available through the line fonts in LaTeX. The possible values for x_s and y_s are integers between -6 and +6, inclusive. These values translate into 25 different absolute angle values, which are listed in appendix C.

B. Bonds in Structural Formulas Written on One Line

The application of some of the bond-drawing mechanisms for this simplest type of structural diagrams is illustrated in figure 3.1.

For figure 3.1a only keyboard characters and the TeX command \equiv were used to produce the bonds. Figure 3.1b shows a slight improvement through raising the triple bond. Figure 3.1c was printed using the \sbond, \dbond, and \tbond command sequences from this thesis, choosing a length of 14 pt for the bonds. It can be seen that each of the formulas in figure 3.1 is a creditable representation of the structure. Depending on the design of the page,

3.1a $CH \equiv C - CH = CH_2$

3.1b $CH{\equiv}C - CH = CH_2$

3.1c $CH{\equiv\equiv}C{-}CH{=}CH_2$

Figure 3.1: One-line structural formulas

the reason for displaying the structure at a particular place, and the emphasis put on features of the structure in the text, one would choose shorter or longer bonds and take more or less trouble to produce the structure.

The picture environment is not needed for one-line structural formulas, unless one of these formulas has to be attached to another structural fragment, as in figure 3.2. Then the coordinate system of the picture environment makes it possible to fit the two fragments together (see chapter V for details).

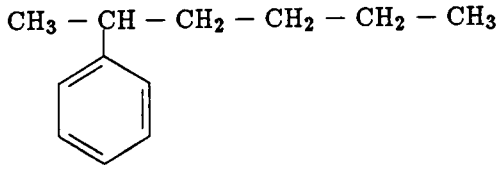


Figure 3.2: One-line structure in picture environment

C. Bonds in Acyclic Structures with Vertical Branches

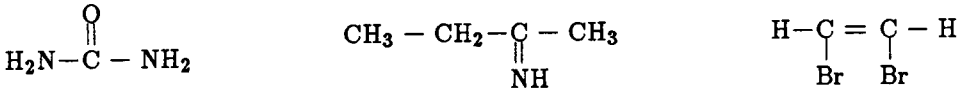
Structure diagrams with vertical, single- or double-bonded, branches, going up or down, are frequently seen. Several experiments with TeX and LaTeX were made to see how this type of structure can be handled. One method is to align the vertical bonds by using the mechanisms for tabbing or for printing tables and matrices. Here a structure such as the one shown in figure 3.3 is



Figure 3.3: Vertical branches

treated as a set of columns as indicated by the vertical dividing lines drawn into the second version of this structure in figure 3.3.

The structure diagram in figure 3.3 uses \vert for the vertical bonds and LaTeX's tabbing environment for the alignment. One can also use "rules" as the vertical bonds in order to give the horizontal and vertical bonds the same lengths. Furthermore, vertical bonds can also be double bonds. The following examples illustrate these features.

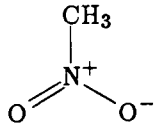


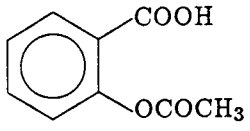
Similar structures were also generated with TeX's \halign mechanism which forms templates for the columns rather than setting tab stops. For the purpose of printing the structure diagrams, no clearcut advantage was seen in one or the other method of alignment. In each case the vertical spacing depends on the line spacing in the document.

The alternative method of producing these structures is the use of the picture environment. It provides better control over horizontal and vertical spacing and over bond lengths. Also, as illustrated in section 2B. of this chapter, using a picture environment makes it possible to attach one structural fragment to another at a specific place. Thus, although the picture environment is not necessary for drawing structures with vertical branches, it has several advantages, and writing LaTeX code for this implementation is not more difficult than writing the code for the tabbing method of alignment.

D. Bonds in Structures Containing Slanted Bond Lines

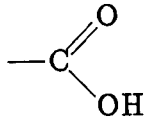
Structure diagrams with slanted bond lines are frequently used for acyclic compounds and have to be used to depict almost all cyclic structures. Two examples are shown here:





In developing diagrams for such structures in this thesis the conventions described in section 1. of this chapter are followed. Thus the symbol for carbon is not printed for the carbons that are ring members, but it is usually printed in acyclic structures, unless the acyclic structure fragment is a long chain, or space for the diagram is limited.

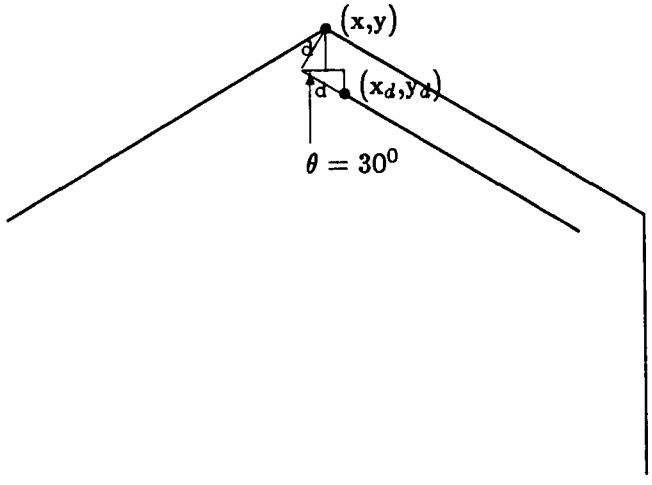
The picture environment is always needed for slanted lines. It was explained in section 2A. of this chapter that LaTeX can draw lines only with a finite number of slopes. This is not a severe limitation for creating the structure diagrams, since the conventions for structure representation allow variations in the angles. The representation does not have to reflect the true atomic coordinates. In fact many chemistry publications contain structure diagrams with angles significantly deviating from the real bond angles, even where those could have been used easily. Thus, Solomons' text (Solomons 84) shows the carboxylic acid group often in this form



with an angle of about 90^0 between the OH and doublebonded O, whereas the true angle is close to 120^0. — The angles used in this thesis for the regular hexagon of the sixring deviate by $\pm 1^0$ from 120^0 because of LaTeX's limited number of slopes. This difference is not big enough to be detected as a flaw.

To write the LaTeX statement for a slanted bond line, one chooses the origin and the slope and then uses trigonometric functions to calculate the LaTeX "length" of the line for the desired real length. Once the LaTeX length is determined, the coordinates of the end point of the line can be calculated in case the end point is needed as the origin of a connecting line. — The origin and length of slanted double bonds were also calculated with standard methods from trigonometry. As an example, figure 3.4 shows how coordinates of the origin were calculated for the inside part of a ring double bond that is at a distance d from the outside bond.

The LaTeX command \multiput is similar to \put and provides a shortcut



$$x_d = x - d\sin\theta + d\cos\theta$$

$$y_d = y - d\sin\theta - d\cos\theta$$

Figure 3.4: Calculating position and length of double bond.

for the coding of structures where several bond lines of the same slope and length occur at regular intervals. Multiput has the format

\multiput(x,y)(Δx,Δy) {n}{object} ,

where n is the number of objects, e. g. lines. A structure diagram for which several \multiput statements are appropriate is the structure of vitamin A shown in figure 3.5.

The size of objects in a picture environment can be scaled in a simple way by changing the unitlength. Figure 3.6 illustrates scaling and two problems associated with it. Changing the unitlength changes the length of the lines only, not the width of the lines or the size of text characters. Thus "it does not provide true magnification and reduction" (Lamport 86, p. 102). However, the size of the text characters can be varied separately, as will be discussed in the next section of this chapter.

The smallest diagram in figure 3.6 illustrates a limitation that is unfortunate for the printing of structure diagrams. The shortest slanted line that

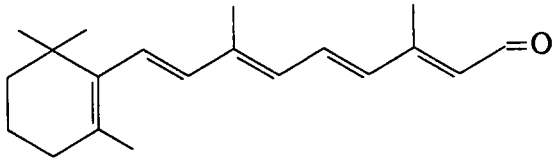


Figure 3.5: Diagram using \multiput

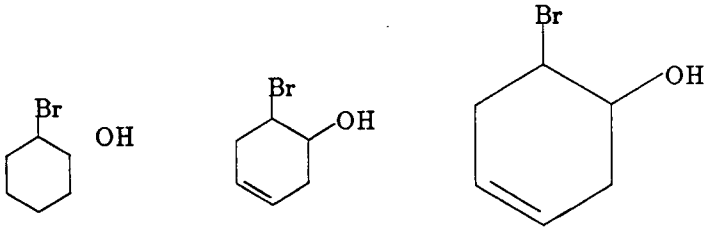


Figure 3.6: Scaling (unitlength=0.07pt, 0.08pt, 0.15pt)

can be printed by LaTeX's line fonts is one with an x-axis projection of about 3.6 mm. If a shorter slanted line is requested, LaTeX just prints nothing. A chemist would occasionally want to draw shorter lines, especially for the purpose of generating dashed lines indicating stereochemical features.

3. ATOMIC SYMBOLS AND CONDENSED STRUCTURAL FRAGMENTS

Special considerations for the printing of condensed structural fragments are required since many of them contain subscripts. TeX considers the printing of subscripts a part of mathematics typesetting which has to be done in the special math mode. It was pointed out in chapter I that typesetting of mathematics documents is one of the strong points of TeX; the fonts of type for the math mode are designed to agree with all conventions of high quality mathematics publishing. Each typestyle in math mode consists of a family of three fonts (Knuth 84, p. 153), a textfont for normal symbols, a scriptfont for first-level sub- and superscripts, and a scriptscriptfont for higher-level sub- and superscripts. When structural fragments such as C_2H_5 are typeset, the textfont is used for the C and the H.

As TeX enters math mode it selects textfont1 as the textfont unless otherwise instructed. Textfont1 is defined by the TeX macros as math italic, a typestyle that prints letters (not numbers) similar to the italic style, but with certain features adapted for mathematics typesetting. The italic style letters, lower and upper case, are the ones commonly seen in typeset mathematical formulas. Chemical formulas on the other hand are not usually printed with slanted letters. In this thesis, basically two methods were employed to produce chemistry-style letters in TeX's math mode which has to be used because of the presence of subscripts.

For a document that contains many chemical formulas it is convenient to redefine textfont1 at the beginning of the TeX input file. The statement

\textfont1=\tenrm was used at the beginning of this document and causes TeX to select the roman font as the textfont in math mode. The roman typestyle is the one normally used by TeX outside of math mode and it is the style in which this thesis is printed. The tenrm style, which is slightly smaller than the twelverm size of the text in this document, was chosen because it appears to look better for the chemical formulas which consist largely of capital letters. When different typesizes are used in this way, all the atomic symbols and formulas in any one structure, even those without subscripts, have to be printed in math mode so that they all have the same size. — It could be a problem with this method of selecting the roman font for math mode that the lowercase Greek letters (and some other symbols used in mathematics) are not available in this font. To print these one can temporarily redefine textfont1 to math italic with the statement \textfont1=\tenmi. One can also switch to a math font different from the default textfont1. Using one of LaTeX's font definitions, \small, a statement {\small$\theta $} will print the Greek letter.

Another method for avoiding the math italic style for letters in chemical formulas is to select the roman style in each individual instance where a formula has to be printed in math mode. A statement such as ${\rm C_2H_5}$ produces C_2H_5 at the size of type currently used in the document. When the typestyle is thus selected within math mode, enclosed by dollar signs, TeX changes the style of the letters of the alphabet only; the lowercase Greek letters and math symbols remain available.

The size of the letters in chemical formulas can be changed with the ten size declarations provided by LaTeX (Lamport 86, p. 200) or with TeX's declarations. (Some of TeX's declarations are not defined in LaTeX (Lamport 86, p.205)). The size declaration has to be written outside of math mode. One place in chemistry typesetting where a smaller typesize is desirable is the writing on reaction arrows. The size in the following example is scriptsize:

$$HC \equiv CH + H_2O \xrightarrow[18\%\ H_2SO_4,\ 90^0]{Hg^{++}} CH_3 - CHO$$

Finally, condensed structural formulas sometimes have to be right-justified to be attached to the main structural diagram. Figure 3.7 illustrates this for the positioning of the substituent in the 4-position of the pyrazole ring. LaTeX makes this positioning convenient with the \makebox command, especially in the picture environment where the command has the format

\makebox(width,height)[alignment]{content}

(Lamport 86, p. 104). The one-line piece of text that constitutes the content of the (imaginary) box can be aligned with the top, bottom, left side, or right side of the box.

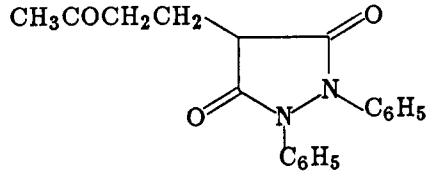


Figure 3.7: Right-justification of substituent formula

CHAPTER IV

MACROS FOR TYPESETTING CHEMICAL STRUCTURE FRAGMENTS

1. GENERAL ORGANIZATION OF A STRUCTURE MACRO

The macro facility in TeX/LaTeX was used to define mnemonics for typesetting frequently occurring structure fragments such as common ring structures and branching patterns. Chapter VI describes the complete system of macros designed for this thesis. All of the macros are defined with the LaTeX declaration \newcommand which has the format

\newcommand{\commandname} [n]{replacement text} .

(TeX calls these definitions macros, whereas LaTeX just uses the more general word "command.") In the definition, n is an integer from 1 to 9 and gives the number of arguments if any are used. The arguments are represented in the replacement text by parameters of the form #1, #2 etc.

A macroname, like the name of any control sequence in TeX, can contain letters only, not numerals. Where it was considered important in this thesis to indicate a numbering scheme in a macroname, either the full word for the number or Roman numerals in lower case letters were used. Thus \hetisix is a mnemonic for a hetero sixring with one hetero atom, and \fuseiv indicates a fusing fragment with four atoms. — In some macronames the mnemonic as such is preceded by "c" or "chem," for "chemistry." This was mainly done where the mnemonic was already used for a control sequence in TeX or LaTeX. Thus, chemical structure fragments that point "right" or "left" are drawn by the macros \cright and \cleft since TeX employs control sequences \right and \left with different meanings.

All the structure-drawing macro definitions in this thesis follow the same pattern in their organization. The actual structure-drawing LaTeX code is preceded by a "box" declaration. The diagrams are produced in a box because

it is sometimes desirable to move the diagram as a whole. In all but one of the macros, the box is the LaTeX "picture"; in the macro \tbranch, which uses the tabbing environment, the box is the LaTeX minipage. Box dimensions in the macros are made flexible through the use of global variables. — The structure-drawing code itself consists of unconditional and of conditional statements. The macro arguments are used to vary parts of the structure diagram, such as substituents and multiple bonds. — Those features of a structure macro that have not been discussed before are described in more detail in the following sections.

A. Box Constructions with Global Variables

Integer variables such as the ones used here for box dimensions have to be stored in one of TeX's 256 numerical registers and can be given symbolic names with TeX's \newcount declaration (Knuth 84, pp. 118-121). The variables used in the macros of this thesis are defined and initialized in the macro \initial which should be part of the preamble of an input file for chemistry typesetting using this system (see chapter VI for a summary of the preamble). A user can then change the variables by simple assignment, e. g. \xi=400. When several variables have been changed, it is convenient to reset all of them, including the unitlength, to their initial values with the macro \reinit.

The role of the LaTeX picture for line-drawing was discussed in chapter III, but the picture is also a box. As such it is processed in horizontal mode, as part of a line. Within a horizontal box, line breaks can never occur. — The macros use variables for all numerical parameters in the picture declaration. Thus, the picture declaration in the macros has the form

\begin{picture}(\pw, \pht)(\-xi, \-yi)

The picture width and height, \pw and \pht, specify the nominal size used by TeX to determine how much room to leave for the box. The diagram in the box can extend beyond these dimensions, but an adjoining box is typeset next to the

preceding one according to the specified width. The user needs control over the picture width in cases where several such boxes are put on one line, especially for the horizontal connection of structure fragments (see chapter V). — Control over the picture height is important because some chemical structures take up more vertical space than others. The picture width and height are initialized to 400 and 900 respectively, which is about 1.4×3.2 cm with the unitlength of 0.1 points.

Variables are used for the coordinates of the lower left corner, \xi and \yi, so that the user can change the placement of the diagram within the picture window. This is not often necessary for individual structures since they can be conveniently centered by the display mechanisms discussed later in this chapter; and also the whole picture can be shifted horizontally by adding horizontal space in front of it with the \hspace command. But shifting the diagram within the picture is employed for one of the methods of the horizontal and vertical connection of structure fragments discussed in chapter V. It was considered to be most convenient to put the minus signs in front of \xi and \yi in the declaration, since one thinks of the lower left corner of a coordinate system as having negative coordinates. With this declaration, an increase in the absolute \xi and \yi values shifts the diagram to the right and up. The coordinates \xi and \yi are initialized to 0 and 300 respectively, which places the coordinate origin about 1 cm above the bottom of the picture window with the unitlength of 0.1 points.

The minipage, the box used in the macro \tbranch, is a paragraph box, which allows line breaks. Only the width is specified for a paragraph box since the height is controlled by the number of lines that will be produced by a given amount of text. The variable used in this system for the width of paragraph boxes is \xbox. The number value assigned to \xbox is interpreted as printer points.

B. Use of TeX's Conditional Facility

TeX's conditional facility is very similar to those of other high-level languages; it has the form:

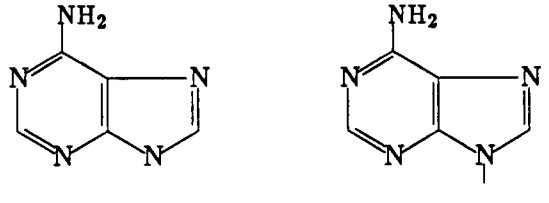
\if(condition)(true text) \else(false text) \fi

(Knuth 84, p. 207 ff.). Nesting is possible. The TeX \if primitive has over ten different forms for testing numbers, processing modes, or tokens. The form used in the chemical structure macros is \ifx⟨token$_1$⟩⟨token$_2$⟩, which tests for the equality of the (character code, category code) pair of two tokens.

In the structure macros, the \ifx tests the arguments. When the arguments are single characters, such as "S", "D", or "C" for single bond, double bond, and circle, respectively, the application is straightforward. The character "Q" is used as argument where "no action" — no substituent, no additional bond — is a desired option at a particular place in a structure diagram. Thus, the coding for ring positions where substituents are an option is:

```
\ifx#nQ
\else    draw a bond line
         print the substituent #n \fi
```

The parameter n represents the substituent formula. — The explicit no-action symbol makes it possible to distinguish three different cases at a particular ring position: no action, just a bond line extending from the ring, and a bond line with a substituent at the position. As an example, the purine macro was used with an argument of Q for the 9-position in the left-hand diagram and with an empty set argument in the right-hand diagram:



One would use the bond-line-only option in cases where another structure fragment in a picture is to be attached to the bond. — The character Q was chosen because it is not part of any element symbol and is not commonly used as a structural symbol otherwise.

When the parameter after the \ifx is substituted by a text string representing a multi-character substituent, TeX actually compares the first character of the string with its second character, since these are the first two tokens encountered. Thus, a substituent that begins with two identical characters always makes the condition true. In such cases, one has to ensure that the string as a whole is compared by enclosing it in a box, e. g. \mbox{$NNHC_6H_5$}.

TeX's conditional facility can also be used to impart some chemical intelligence to a macro by causing screen messages to be generated when the user supplies a combination of arguments that is chemically not possible. Such a combination would be a ring double bond (argument 7="D") and a circle denoting aromaticity (argument 9="C") for the carbon sixring. A section of code

```
\ifx#7D \ifx#9C
                    \message{Error: ...}
\fi \fi
```

will produce the message on the screen while the input file is processed by TeX to give the DVI file. The user can then correct the mistake and reprocess the input file before sending the DVI file to the output device. — Error messages were not placed into all macros, just into the \sixring macro to demonstrate this feature.

In addition to the simple \if statements TeX has an \ifcase construction of the form

```
\ifcase(number)(text for case 0) \or (text for case 1) \or ...
  \or (text for case n) \else (text for all other cases) \fi
```

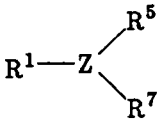
(Knuth 84, p 210). When the \ifcase statement is used in a macro and the case number is passed as an argument, many different actions can be requested through one argument. For a larger number of cases, the TeX code with \ifcase is somewhat more elegant than a series of individual \if statements. — A suitable application for the chemical structure macros is the placement of double bonds in various positions of a structure, where the position number, or a numeric code for a combination of positions, is passed as the case number. One version of the sixring macro, \sixringb, contains an \ifcase construct.

2. USE OF STRUCTURE MACROS

A. Invoking the Structure Macros

In order to typeset a chemical structure through invoking one of the macros in this system, the user must know how the structure will be oriented on the page, besides knowing, of course, the function of each argument. The structures in this system can not be rotated; they are oriented according to common practices in chemistry, but occasionally the user will have to adapt a model structure to the given orientation. Chapter VI shows for each macro a typical structure produced by it, to illustrate the orientation and the position numbers. In this section it will be demonstrated with two representative macros how this information is to be used.

The macro \cright typesets structures or structure fragments of the general form



The arguments 1, 3, 5, and 7 are the substituents or groups R^1, Z, R^5, and R^7. Arguments 2, 4, and 6 trigger the drawing of bonds between R^1 and Z, between Z and R^5, and between Z and R^7, respectively. The bonds can be single or double

bonds, from arguments "S" and "D" respectively, and the bond between R^1 and Z does not have to be present. A typical structure is shown in figure 4.1, where the caption is the invoking code. Since the second argument for figure 4.1 is "Q," no bond is drawn between R^1 and Z; and R^1, which is in a \makebox (see chapter III.3), is moved next to Z. Figure 4.2 shows two additional structures drawn with the \cright macro.

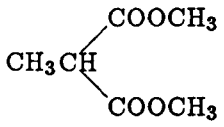
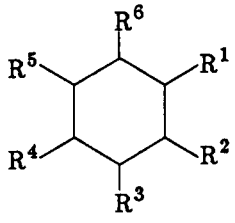


Figure 4.1: \cright{CH_3}{Q}{CH} {S}{$COOCH_3$}{S}{$COOCH_3$}



Figure 4.2: Structures drawn with the \cright macro

The macro \sixring typesets the very common carbon sixring. For this structure and all the ring structures, the user has to know how the system of macros assigns position numbers to the ring atoms. The assignment follows chemical nomenclature rules where applicable. However in the case of single-ring structures where all ring atoms are carbons, the assignment of position number 1 is arbitrary. In this system of macros, the sixring is numbered as follows:



The first six arguments are the formulas for the optional substituents in the respective positions. The user has to refer to the position assignment to supply the text strings in the correct form, e. g. a sulfonic acid group for positions 4 and 5 would be typed in as HO_3S, whereas it would be SO_3H for all other positions. — The remaining three arguments can produce alternating ring double bonds, but each of them has a second function. Argument 7 can produce a second substituent at position 1, argument 8 an outside double bond with substituent in position 3, and argument 9 a circle inside the ring denoting aromaticity. Where arguments have two functions in this way, the different structural features generated are of course mutually exclusive chemically. Figure 4.3 shows a sixring structure with the corresponding LaTeX code as caption. The substituent in position 3 is passed as argument 8 since it is not the regular single-bonded substituent represented by argument 3.

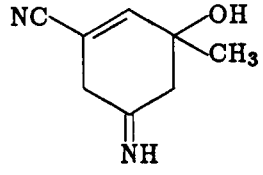


Figure 4.3: \sixring{OH}{Q}{Q}{Q} {NC}{Q}{CH_3}{NH}{D}

B. Displaying Macro-Generated Diagrams within a Document

Since the macro-generated diagrams constitute TeX boxes, the code for a diagram can be included anywhere in the input file and TeX will try to find a place for the diagram as a whole in the line and on the page. The diagrams, however, take up more space horizontally and vertically than a box of text; therefore TeX's line- and page breaking mechanisms would be strained, sometimes to such a degree that text squeezing or spreading would be apparent on the printed page.

A convenient way to display one or several structures on one line by themselves and centered, is the math display environment, enclosed by LaTeX with brackets in the form \[...\]. A modification of this environment puts consecutive equation numbers at the right edge of the line. When using math display one has to remember that the only spacing in math mode is around math operators, space inserted by the user is ignored. Also, a new paragraph can not be started in math display. — The diagrams in math display are printed at the place in the document where they are coded in the input file. Thus the display can still cause problems with pagebreaking, especially since the diagrams usually take up more vertical space than a math equation for which the environment is designed.

The LaTeX figure environment was specially designed for larger displays, such as the structure diagrams (Lamport 86, pp. 59, 60, 176, 177). A figure will usually not appear at the place in the document where the user has coded it; instead LaTeX finds space for it on the current page or the next one in such a way that an overfull page is never produced. The user has some control over the placement of figures with optional parameters such as top or bottom of a page, but LaTeX still makes the final decision, often producing surprising results. Since a figure is moved to a convenient place by LaTeX, the figure is called a "float." In addition to the automatic space-finding, the figure environment has the advantage that it makes captions possible. Furthermore, several displayed objects, each with a caption, can be included in one figure environment; figures 4.4 and 4.5 present an example, produced by the LaTeX code in figure 4.6. The code shows that each picture with its caption is enclosed in a \parbox.

It should also be mentioned that all the display mechanisms discussed above can be used in a two-column document style which is the format of most scientific chemistry journals.

If the same structure is to be printed many times in a document, processing

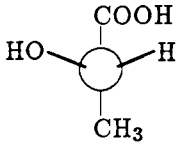


Figure 4.4: (S)−lactic acid

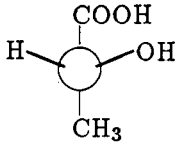


Figure 4.5: (R)−lactic acid

```
\begin{figure}
  \parbox{.4\textwidth}{\centering
    \ccirc{$COOH$}{$HO$}{$H$}{$CH_{3}$}
    %ccirc is a macro
  \caption{${\rm (S)-lactic acid}$}}
  \hfill
  \parbox{.4\textwidth}{\centering
    \ccirc{$COOH$}{$H$}{$OH$}{$CH_{3}$}
  \caption{${\rm (R)-lactic acid}$}}
\end{figure}
```

Figure 4.6: LaTeX code for two captions in one figure

time can be saved by storing it in a LaTeX \savebox. The structure-drawing code then has to be processed once only. Applied to the chemistry macros, a statement

$$\backslash\text{savebox1}\{\backslash\text{macroname } \{\text{arg}_1\}\{\text{arg}_2\}\ldots\ \}$$

will process the code and save the typeset structure. A \usebox1 statement, enclosed in a math display or figure environment, is then placed into the input file whereever the structure is to be printed. A \savebox can also be given a symbolic name (Lamport 86, p. 101).

CHAPTER V

COMBINING STRUCTURES FROM SEVERAL MACROS

1. GENERAL CONSIDERATIONS FOR COMBINING STRUCTURES IN THIS SYSTEM

Many individual structure diagrams can be typeset using just one of the macros together with condensed, one-line formulas; but often it will be necessary to combine the ring structure or the branched fragment from one macro with a structure part from another. To do this, the separate parts have to be precisely aligned horizontally and vertically. The typesetting of chemical equations containing structure diagrams also requires such alignments and is therefore included in this chapter. For the system of macros described here, alignment consists of moving each of the structure fragments as a whole, either within its picture box or together with the box.

In each structure diagram there are many different points to which other fragments can be attached. Similarly, one and the same structure can be aligned in different ways with others to produce a chemical equation. For these reasons, it was not considered feasible to develop a symbolic language for alignment, such as "attach(sixring) at(1) to(fivering) at(4)." Instead, this system lets the user manipulate the placement of the structures at a lower level by using some of the numerical coordinates from the macros. While it may be considered a disadvantage that the user has to extract information from the macros, this method also puts a lot more control into the hands of the user. The type of user anticipated for this system will probably prefer this mechanism to an overdose of user-friendliness. Outright manipulation of coordinates is also well suited for the textual method of structure input employed in this system, since it helps the user to visualize the result.

The information that is needed from a macro to form a new structure

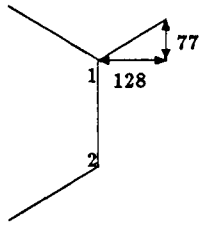
from fragments is the coordinate pair for the point of attachment in each fragment. A part of the code for the sixring and the relevant part of the structure is shown in figure 5.1 to illustrate briefly how these coordinate pairs are obtained: If a fragment is to be attached directly to a ring position, for example to position 1, the coordinate pair is found in the unconditional part of the code as the origin of the bondline beginning at position 1. (The code for the bond is located by finding the respective line comment.) The coordinate pair in this case would be (342,200). — A fragment can also be attached to the end of a bond extending from the ring. These bonds are optional and part of the conditional code. The optional bond extending from position 1 is located through the comment "substituent on 1." The x-coordinate at the end of this bond is $342 + 128 = 470$, since the length of the bond given in the code, 128 units, is the projection on the x-axis. The y increment from ring position 1 to the end of the bond is obtained from the slope of the line and the x increment of 128: $\Delta x(3/5) = 77$. Thus, the y-coordinate at the end of the bond is $200 + 77 = 277$ units. – Appendix B lists the coordinates of the more commonly used points of attachment for the system of macros described here.

Two conceptually different methods were used in this thesis to combine structure fragments from different macros. — One method follows a suggestion in the LaTeX manual (Lamport 86, p. 110) to put subpictures into an encompassing picture with the \put command:

\put(x,y){\begin{picture} ... \end{picture} }.

The reference point (x,y) is the lower left corner of the subpicture. When this technique is applied to the chemical structure macros, the macro invocation constitutes the subpicture. The user has to set up the encompassing picture and determine the coordinates of the reference points from the coordinates of the points of attachment between structure fragments.

The second method is somewhat less versatile; but there are applications




```
\begin{picture}(\pw,\pht)(-\xi,-\yi)
 .....
 \put(342,200)      {\line(0,-1) {200}} % bond from 1 to 2
 \ifx#1Q
 \else\put(342,200){\line(5,3)   {128}} % substituent on 1
      \put(475,250){#1}              \fi
 .....
\end{picture}
```

Figure 5.1: Finding coordinates of points of attachment

for which it is preferable. In this method, the individual picture boxes are put next to one another on one line or on successive lines. The fragments in the separate pictures are aligned by shifting the coordinate system, i. e. by changing the \xi and \yi values in the picture declaration, in one or more of the pictures.

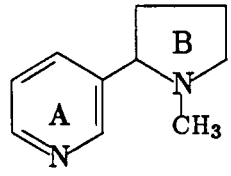
Finally, for the alignment of structures in a chemical equation, it is convenient to use a paragraph box construction (\parbox) in addition to coordinate shifting. LaTeX centers a paragraph box vertically on the current line which contains, in the case of the chemical equation, textual items such as plus symbols, condensed formulas, and reaction arrows.

Typical applications of all methods of combining structure fragments will be described in the rest of this chapter.

2. COMBINING FRAGMENTS TO FORM A NEW STRUCTURE

A. Attachment by the Subpicture Method

A simple example for this technique of structure-building is shown in figure 5.2, where two different heterocycles are fitted together to produce the structure of nicotine. The LaTeX code to be entered by the user for this structure is given underneath the diagram.




```
\begin{picture}(900,900)(0,0)
 \put(0,0)      {\hetisix{D}{Q}{}{Q}{Q}{Q}{D}{D}{N} }
 \put(470,277) {\hetifive{$CH_{3}$}{Q}{Q}{Q}{Q}{S}{S}{S}{N}}
\end{picture}
```

Figure 5.2: Nicotine structure with LaTeX code

The code in figure 5.2 illustrates how the user has to set up the encompassing picture with the \begin and \end statements and estimated values for the picture width and height, both 900 units (about 3cm) in this example. Then the picture box of the pyridine ring is placed at the origin of the outer picture. Next, the points of attachment are found in the respective macros as $x_{AB} = 470$, $y_{AB} = 277$ for pyridine and $x_{BA} = 0$, $y_{BA} = 0$ for pyrrolidine. The coordinates of the reference point in the outer picture where the inner picture with the pyrrolidine ring has to be placed then are

$$x = x_{AB} - x_{BA} = 470, \quad y = y_{AB} - y_{BA} = 277.$$

It is assumed that the lower left corner of both subpictures has the same coordinates, and this is the case when the macros are used. When more than two

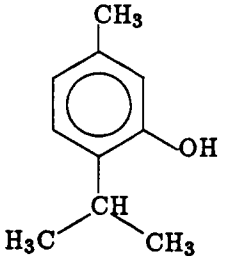
ring structures are combined one after the other, the calculation of the reference points is appropriately extended.

Since the macros for the various acyclic branched fragments also consist of picture boxes, these fragments can be used as subpictures together with ring structures and with other acyclic fragments. Thus the structure of thymol in figure 5.3 is produced from the \sixring and the \cdown macros. Again, the coordinates for the point of reference for the \cdown picture are calculated from the points of attachment:

$$x = x_{sixring} - x_{cdown} = 171 - 33 = 138$$

$$y = y_{sixring} - y_{cdown} = -103 - 220 = -323.$$

Figure 5.4, the structure of penicillic acid, combines three subpictures, one from the \cleft macro and two from the \cbranch macro which draws vertical branches.




```
\begin{picture}(500,1100)(0,-300)  % estimated dimensions
  \put(0,0)      {\sixring ...                }
  \put(138,-323){\cdown    ...                }
\end{picture}
```

Figure 5.3: Combining ring structure and acyclic subpictures

Many structures contain condensed formula fragments between ring diagrams. The structure of the anesthetic piridocaine is shown as an example in figure 5.5. In such a case one has to estimate the average horizontal space per character and move the second subpicture that much further to the right for

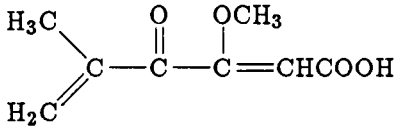


Figure 5.4: Combining acyclic subpictures

each character, including the subscripts, in the condensed formula fragment. In the ten point size, in which the characters in figure 5.5 are printed, the horizontal space per character is 6.8 points or 68 of the picture units.

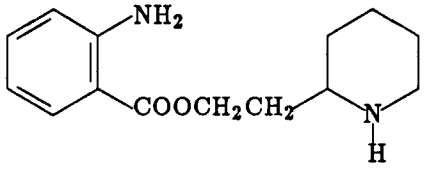
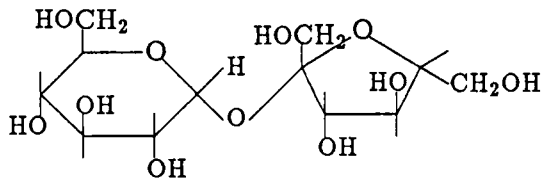


Figure 5.5: Condensed formula fragment between rings

Other special cases occur where a diagram would become too crowded when the two fragments are put next to one another. (This does not necessarily reflect steric hindrance in the real, three-dimensional chemical structure.) In such cases the user can design a longer bondline and put it into the outer picture between two points of attachment on macro-produced structure fragments. The structure of sucrose, shown in figure 5.6, illustrates this technique. For this structure, it was estimated that the x-offset between the bonding oxygen on glucose and the fructose ring should be at least 200 units to produce a diagram that does not appear crowded. Using this x-offset and a bonding angle of 45^0 (the angle used in the pyranose macro for glucose), the user can then easily calculate the point of reference for the fructose subpicture.

An important special case of combining structure fragments is the generation of fused ring systems. In a fused ring system more than one ring atom is shared between rings. — A simple method for producing such diagrams is

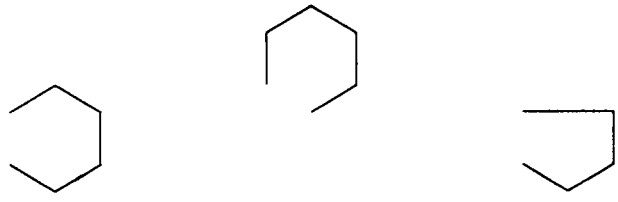



```
\put(0,0)      {\pyranose ....          }
\put(785,200)  {\line(1,1){200}}        }  % user-designed line
\put(985,100)  {\furanose ...           }
```

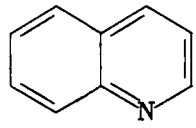
Figure 5.6: User-designed connecting bond line

to print the shared bondlines from individual ring structures precisely on top of each other. The bond lines have to have the same lengths, which is true in this system of macros for the five- and sixrings, the most frequently occurring ones. The shared lines don't appear to be heavier in the printed picture than other bond lines. Figure 5.7 shows the structure diagram of quinoline produced by this method together with the respective LaTeX code. These fused systems can of course include substituents and multiple bond variations at all positions where the original macros made them possible. — A relatively small number of single-ring fragments can produce a large number of fused systems in this way, among them the very common fused systems of anthracene, phenanthrene, chrysene, indene, indol, benzimidazole, quinoline, and acridine.

There are also some macros that draw fragments specifically designed for fusing. The following fragments



are produced by the \fuseiv, \fuseup, and \fuseiii macros. They can be attached




```
\begin{picture}(900,900)(0,0)
  \put(0,0)    {\sixring{Q}{Q}{Q}{Q}{Q}{Q}{S}{D}{D} }
  \put(342,0) {\hetisix{D}{Q}{Q}{Q}{Q}{Q}{D}{D}{N} }
\end{picture}
```

Figure 5.7: Fusion of fully drawn rings

to the five- and sixrings as subpictures. These fragments have the advantage that they can provide more options for double bond locations than a full ring structure within the constraint of nine arguments.

B. Attachment by Shifting the Coordinate System

This method is easy to use when a complex structure can be perceived as a series of fragments put next to one another horizontally, although not necessarily on exactly the same level. The structure of nicotine shown in figure 5.2 belongs to this category. As an alternative to the code listed in figure 5.2, the following LaTeX statements can be used to produce the nicotine diagram:

\pw = 470

\hetisix ...

\advance \yi by 277

\hetifive ... .

The first statement here sets the picture width \pw for the pyridine ring so that the rightside end of the picture box is at the x-coordinate of the point of attachment. Now LaTeX will put the next item on the line, in this case the picture box with the pyrrolidine ring, flush next to the pyridine box. When the

structure is printed in a math display environment (see chapter II) where no spacing between items on a line is applied, there will be no space between the picture boxes. When the structure is put into a figure environment only, without math display, normal spacing occurs as it would happen between words on a line. The user then has to request negative horizontal space between invoking the pyridine and the pyrrolidine macro to correct for the spacing. A statement \hspace{-11pt} produced the right correction for the typestyle of this document.

The statement \advance \yi by 277 causes the coordinate-shifting in the pyrrolidine picture. By increasing the y-coordinate, the pyrrolidine structure is shifted upwards so that the points of attachment of the two rings meet. In general, the coordinate shifts Δxi and Δyi applied to the second or any following picture are determined from the points of attachment (x_{AB},y_{AB}) and (x_{BA},y_{BA}) (the terminology used for figure 5.2) as follows:

$$\Delta xi = x_{BA},\ \Delta yi = y_{AB} - y_{BA}.$$

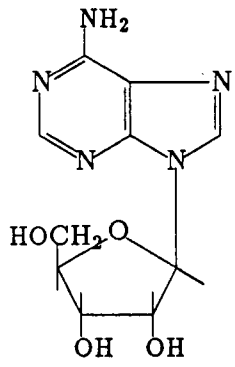
For vertical attachment, connecting one fragment to the lower end of another by coordinate shifting, the following steps are necessary: The points of attachment of the upper and the lower fragment are shifted to the bottom and to the top of their respective picture boxes and the x-coordinates of attachment are aligned. The new \xi and \yi values are then

$\backslash yi_{upper} = -y_{upper}$

$\backslash xi_{lower} = x_{upper} - x_{lower}$

$\backslash yi_{lower} = \backslash pht_{lower} - y_{lower} + 14$(correction for vertical spacing).

For the structure of adenosine shown in figure 5.8 the points of attachment on purine (at the bottom of N-9) and on deoxyribose (at the top of the long bond) have the coordinates (513,-130) and (448,380), respectively. Thus the structure was produced by the code given underneath the diagram. The horizontal space is used here instead of the centering option. The blank lines




```
\hspace{5cm}  \yi=130          \purine{ ... }
(blank line)
\hspace{5cm}  \xi=65   \yi=534  \furanose{ ... }
(blank line)
\caption{ ... }
```

Figure 5.8: Vertical attachment by coordinate shifting

after each ring structure code are necessary to inform LaTeX that the next item should not be printed on the same line.

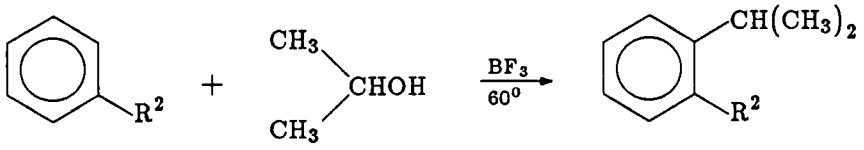
3. ALIGNING STRUCTURES IN AN EQUATION

In a chemical equation containing structure diagrams the various constituents of the equation have to be horizontally aligned. The equation is typeset in LaTeX's horizontal mode on one line, the current printline. Text items such as condensed formulas and plus symbols are put on the line as usual, their (imaginary) baseline determining the position of the line. The structure diagrams, as drawn by the macros, will not be vertically centered on the current line. They are drawn in picture boxes which are typeset on the line with the lower end of the (imaginary) box at the baseline of the current line. The picture boxes are positioned at this height without regard to the coordinates declared

for the lower left corner of the box. To line up the vertical middle of the diagram in the box with the text of the line, one would have to shift the diagram downwards beyond the bottom of the declared picture. While this can be done, it might result in a lack of space under the equation, since LaTeX reserves space only according to the declared dimensions of the picture. Paragraph boxes on the other hand are normally centered on the vertical center of the current line. Paragraph boxes containing the macros are positioned somewhat differently, with a point one third up from the bottom of the picture at the base of the current line. Thus there is one third of the declared picture below the base of the current line which yields enough vertical space to set off the equation from the succeeding text.

The equation in figure 5.9 was typeset by putting each macro-drawn diagram into a \parbox. The y-coordinate of the lower end of the pictures is -300 as usual, which puts position 2 of the sixring and the CHOH part from the \cleft macro at the base of the current line. The LaTeX code for the equation is shown underneath it in the figure. The y-coordinates of the structures and of the reaction arrow, in this case drawn by a macro, could be shifted individually as well to change the alignment. The TeX control sequence \to ($\rightarrow$) can be used instead of the special reaction arrow for chemistry; ($\rightarrow$) is always centered on the line.

Getting good-looking horizontal spacings within the equation usually requires some experimenting. As previously mentioned, there is no inter-item spacing in math mode. Therefore more explicit horizontal space has to be added when chemical equations are typeset in the math display environment.




```
\hspace{1.5cm}
\parbox{40pt}{\sixring{Q}{$R^{2}$}{Q}{Q}{Q}{Q}{S}{S}{C} }
\hspace{1cm}  $+$  \hspace{1.5cm}
\parbox{40pt}{\cleft{$CH_{3}$}{S}{$CHOH$}{S}
                    {$CH_{3}$}{Q}{}        }
\parbox{40pt}{\cto{BF_{3}}{60_{0}}{3}      }
\hspace{3mm}
\parbox{40pt}{\sixring{$CH{(CH_3)}_2$}{$R_{2}$}{Q}{Q}{Q}{Q}
                    {S}{S}{C}  }
```

Figure 5.9: Alignment in a chemical equation

CHAPTER VI

THE COMPLETE SYSTEM OF MACROS — ITS DESIGN AND ITS USE

1. GENERAL DESIGN CRITERIA

LaTeX code can be written to typeset a structure diagram for any chemical compound in such a way that the diagram conforms to accepted practices in chemistry publications. The purpose of the macros in this thesis is to reduce the amount of low-level, bond-by-bond coding necessary to typeset a particular structure. The problem of designing a generally useful system of macros for this purpose has to be seen in the context of the large number of possible structures: More than 7 million chemical compounds are registered with the Chemical Abstracts Service, including 60,000 different ring systems, and innumerable additional structures are possible. The fragments to be typeset by the macros in this thesis were selected such that they could be helpful in producing the more common types of structures. The arguments of each macro in turn were selected with the goal of making the respective fragment as flexible as possible, such that the more common known structures of this type can be typeset by a particular macro. The selection criteria were informal, using the "expert knowledge" of the writer. — Where a related approach to displaying chemical structure diagrams was taken in previous work, a similar selection of fragments was made (Zimmerman 84)(Bendall 80,85). Both authors are chemists and present a collection of fragments without any attempt to justify their choices.

For more flexibility, the user of the system of macros presented in this thesis can effect many structural variations by defining an outer picture (see chapter V) and placing supplemental lines and atomic symbols into it in addition to the fragment(s) produced by the macros. It is estimated that the

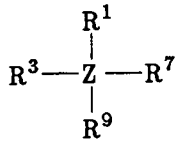
macros can provide shortcuts to the drawing of more than 50% of the structures shown in the widely used textbook by Solomons (Solomons 84).

The next section lists the individual macros in this system, each with a typical generic structure and directions for the use of all arguments. Specific selection criteria for the fragment as such and the arguments are mentioned in many cases. The fragments are listed in the traditional categories of organic chemistry: acyclic structures, alicyclic structures (rings where all ring members are carbon atoms), and heterocyclic structures. The number of arguments is given in brackets behind the macro name.

2. MACROS OF THE SYSTEM

A. Macros for Acyclic Fragments

i. Macro \cbranch[9]. This macro typesets structural fragments with vertical branches:



Arguments 1, 3, 7, 9: The substituent formulas for R^1, R^3, R^7, and R^9.

Argument 2: The bond between R^1 and Z, "S" for a single bond and "D"for a double bond. No action is taken for any other value of the argument.

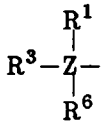
Argument 4: The bond between R^3 and Z, "S" for a single bond and "D" for a double bond. When the argument is "Q", no bond is drawn and R^3 is moved next to Z.

Argument 5: The center atom(s), Z. When the argument is a string of more than one character, argument 6 should not be "S" or "D," and argument 7 should be an empty set.

Argument 6: The bond between Z and R^7, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

Argument 8: The bond between Z and R^9, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

ii. Macro \tbranch[7]. This macro typesets structural fragments with vertical branches similar to \cbranch. The main reason for including \tbranch is to show the use of the LaTeX tabbing mechanism for printing structural fragments. Macro \cbranch is the preferred macro for structures of this type. In contrast to the other macros, \tbranch provides the math mode for the substituent formulas in the macro code. Therefore substituent formula arguments do not have to be enclosed by $ symbols.



Arguments 1 and 6: The substituent formulas for R^1 and R^6.

Argument 2: The bond between R^1 and Z, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

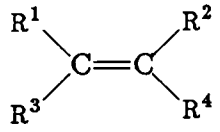
Argument 3: Atom symbols and bonds to the left of Z. Single bonds have to be typed in as hyphens, double bonds as equal signs.

Argument 4: The center atom Z and any bonds and atom symbols to its right. Single bonds have to be typed in as hyphens, double bonds as equal signs.

Argument 5: The bond between Z and R^6, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

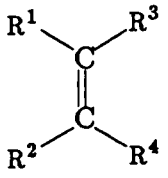
Argument 7: An integer number which is interpreted as printer points of negative space between lines. The correct number for a document with double spacing is 13.

iii. Macro \ethene[4]. This macro typesets an ethene fragment with four variable substituents:



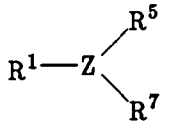
Arguments 1 – 4 are the substituent formulas represented by R^1, R^2, R^3, and R^4.

iv. Macro \upethene[4]. This macro is similar to \ethene, but it draws the ethene double bond vertically:



The arguments have the same meaning as they do for \ethene.

v. Macro \cright[7]. This macro typesets the following fragment which is often used for carboxylic acids and their derivatives:



Arguments 1,5,7: The substituent formulas R^1, R^5, and R^7.

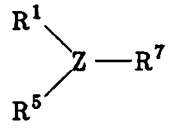
Argument 2: The bond between R^1 and Z, "S" for a single bond and "D" for a double bond. For an argument of "Q", no bond is drawn and R^1 is moved next to Z.

Argument 3: The center atom(s) Z.

Argument 4: The bond between Z and R^5, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

Argument 6: The bond between Z and R^7, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

vi. Macro \cleft[7]. This macro typesets a fragment similar to the one produced by \cright, but opening to the left:



Arguments 1, 5, 7: The substituent formulas R^1, R^5, and R^7.

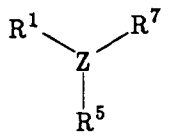
Argument 2: The bond between R^1 and Z, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

Argument 3: The center atom(s) Z. When the argument is a string of more than one character, argument 6 should not be "S" or "D", and argument 7 should be an empty set.

Argument 4: The bond between R^5 and Z, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

Argument 6: The bond between Z and R^7, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

vii. Macro \chemup[7]. This macro typesets the following fragment which can be used for small molecules with trigonal geometry:



Arguments 1, 5, 7: The substituent formulas R^1, R^5, and R^7.

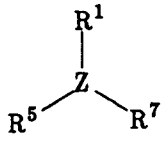
Argument 2: The bond between R^1 and Z, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

Argument 3: The center atom Z.

Argument 4: The bond between Z and R^5, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

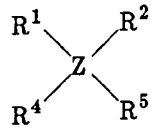
Argument 6: The bond between Z and R^7, "S" for a single bond and "D" for a double bond. No action is taken for any other value of the argument.

viii. Macro \cdown[7]. This macro typesets a fragment similar to the one produced by \chemup, but opening downwards:



The arguments have the same meaning as they do for \chemup.

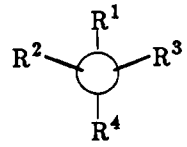
ix. Macro \csquare[5]. This macro typesets a fragment that is sometimes used when all four substituents on a center atom have to be shown explicitly:



Arguments 1, 2, 3, 4: The substituent formulas R^1, R^2, R^4, and R^5.

Argument 3: The center atom Z.

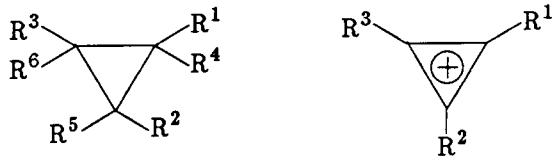
x. Macro \ccirc[4]. This macro typesets a fragment used to show the actual configuration at a tetrahedral atom. The tetrahedral atom itself is not shown and is assumed to be in the middle of the sphere represented by the circle. The bonds typeset as heavier lines and intersecting the circle are directed out of the plane of the paper, towards the viewer.



The arguments 1 – 4 are the substituent formulas R^1, R^2, R^3, and R^4.

B. Macros for Alicyclic Ring Systems

i. Macro \threering[9]. This macro typesets the cyclopropane ring. The aromatic cyclopropenyl cation is drawn with a circle enclosing a plus sign inside the ring. The ring positions to which R^1, R^2, and R^3 are attached are designated as position 1, 2, and 3, respectively:



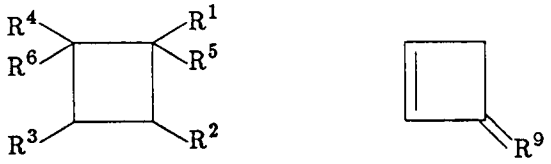
Arguments 1 – 6: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1, R^2, R^3, R^4, R^5, and R^6. The bond line to R^2 is straight if the circle is in the ring or if there is no second substituent at position 2, and slanted otherwise.

Argument 7: An argument of "D" typesets a second bond between ring positions 1 and 3. No action is taken for any other value of the argument.

Argument 8: An argument of "Q" causes no action. All other argument values cause an outside double bond to be drawn from ring position 2, and the argument itself to be put at the end of the double bond as the substituent formula.

Argument 9: An argument of "C" typesets a circle enclosing a plus sign inside the ring. All other argument values cause no action.

ii. Macro \fourring[9]. This macro typesets the cyclobutane ring. The ring positions to which R^1, R^2, R^3, and R^4 are attached are designated position 1, 2, 3, and 4, respectively.



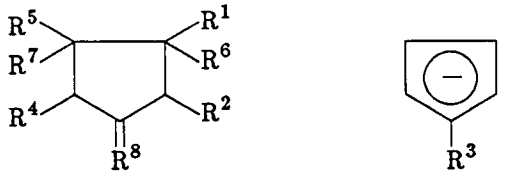
Arguments 1 – 6: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1, R^2, R^3, R^4, R^5, and R^6.

Argument 7: An argument of "D" typesets a second bond between ring positions 1 and 2. No action is taken for any other value of the argument.

Argument 8: An argument of "D" typesets a second bond between ring positions 3 and 4. No action is taken for any other value of the argument.

Argument 9: An argument of "Q" causes no action. All other argument values cause an outside double bond to be drawn from ring position 2, and the argument itself to be put at the end of the double bond as substituent formula R^9.

iii. Macro\fivering[9]. This macro typesets the cyclopentane ring. The aromatic cyclopentadienyl anion is drawn with a circle enclosing a minus sign inside the ring. The ring positions to which R^1, R^2, R^3, R^4, and R^5 are attached are designated as position 1, 2, 3, 4, and 5, respectively:



Arguments 1 – 5: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1, R^2, R^3, R^4, and R^5.

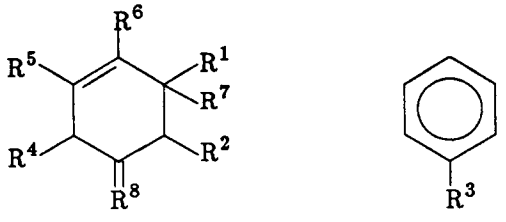
Argument 6: An argument of "D" typesets a second bond between ring positions 1 and 2. An argument of "S" causes no action. All other argument values are used as the substituent formula R^6.

Argument 7: An argument of "D" typesets a second bond between ring positions 4 and 5. An argument of "S" causes no action. All other argument values are used as the substituent formula R^7.

Argument 8: An argument of "Q" causes no action. All other argument values cause an outside double bond to be drawn from ring position 3, and the argument itself to be put at the end of the double bond as substituent formula R^8.

Argument 9: An argument of "C" typesets a circle enclosing a minus sign inside the ring. No action is taken for any other value of the argument.

iv. Macro\sixring[9]. This macro typesets a carbon sixring as a regular hexagon. A benzene ring can be drawn with alternating double bonds or with a circle inside the ring. The ring positions to which R^1, R^2, R^3, R^4, R^5, and R^6 are attached are designated as position 1, 2, 3, 4, 5, and 6, respectively:



Arguments 1 – 6: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1, R^2, R^3, R^4, R^5, and R^6.

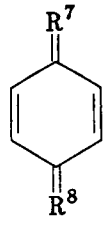
Argument 7: An argument of "D" typesets a second bond between ring positions 1 and 2. An argument of "S" causes no action. All other argument values are used as the substituent formula R^7.

Argument 8: An argument of "D" typesets a second bond between ring positions 3 and 4. An argument of "S" causes no action. All other argument values cause an outside double bond to be drawn from ring position 3 and the argument itself to be put at the end of the double bond as substituent formula R^8.

Argument 9: An argument of "D" typesets a second bond between ring positions 5 and 6. An argument of "C" typesets a circle inside the ring. No action is taken for any other value of the argument.

v. Macro \sixringa[9]. This macro differs from \sixring only in the positions of the double bonds. A value of "D" for arguments 7, 8, and 9 puts a double bond between ring positions 1 and 6, ring positions 2 and 3, and ring positions 4 and 5, respectively. Since the carbon sixring is so common, more options are needed for it than for the other rings.

vi. Macro \sixringb[9]. This macro is also very similar to \sixring, but it allows all 17 chemically possible combinations of double bonds, including the three quinoid structures that can not be typeset with \sixring or \sixringa.



Arguments 1 – 6: These arguments have the same meaning as in \sixring.

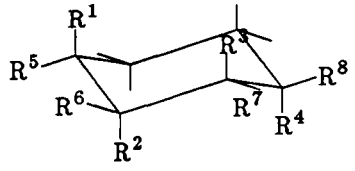
Argument 7: An argument of "Q" causes no action. All other argument values cause an outside double bond to be drawn from ring position 6

and the argument itself to be put at the end of the double bond as substituent formula R^7.

Argument 8: An argument of "Q" causes no action. All other argument values cause an outside double bond to be drawn from ring position 3 and the argument itself to be put at the end of the double bond as substituent formula R^8.

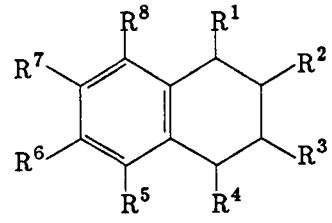
Argument 9: An integer number. The number zero causes the circle to be drawn inside the ring. All other integers are interpreted as a combination of ring double bonds according to the bit pattern corresponding to the decimal integer: A bit pattern of 000001 is interpreted as a double bond beginning at ring position 1, a bit pattern of 100000 (integer 32) as a double bond beginning at ring position 6. Thus, argument 9 for the diagram shown above is 9 (001001). No action occurs for argument values that correspond to a combination of double bonds which is chemically not possible, namely any combination with two adjoining double bonds.

vii. Macro\chair[8]. This macro typesets the saturated carbon sixring in its most favorable conformation. The axial and equatorial bond lines to the substituents are always drawn by this macro, even when there is no substituent in a particular position. This is the usual practice in drawing the chair form.



The eight arguments represent the respective substituent formulas R^1 – R^8.

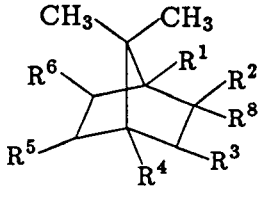
viii. Macro \naphth[9]. This macro typesets the aromatic naphthalene ring system, the fully saturated decalin ring system, and the 1,2,3,4-tetrahydronaphthalene shown in the diagram. The position numbers 1 – 8 are specified by the nomenclature rules of chemistry.

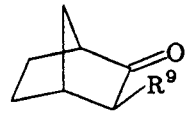


Arguments 1 – 8: An argument of Q causes no action. All other argument values are used as the respective substituent formulas R^1 – R^8.

Argument 9: A value of "S" typesets the ring system with no double bonds (decalin). A value of "D" typesets the aromatic system naphthalene with alternating double bonds. All other argument values draw the partially saturated system shown above.

ix. Macro \terpene[9]. This macro typesets the bicyclo(2.2.1)heptane ring system found in such terpenes as borneol, camphor, and fenchol. The position numbers 1 – 7 are specified by the nomenclature rules of chemistry.





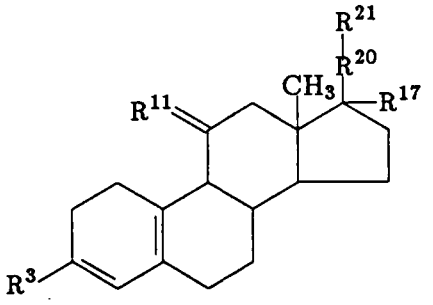
Arguments 1 – 6: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1 – R^6.

Argument 7: An argument of "M" prints two methyl groups on bonds extending from carbon #7. All other arguments cause no action.

Argument 8: An argument of "Q" causes no action. An argument of "O" prints an oxo group at carbon #2. All other argument values are used as a second substituent on carbon #2, shown as R^8.

Argument 9: An argument of "Q" causes no action. An argument of "D" prints a second bond between positions 2 and 3. All other argument values are used as a second substituent on carbon #3, shown as R^9.

x. Macro \steroid[9]. This macro typesets the steroid skeleton. The position numbers are specified by the nomenclature rules. The arguments are selected such that common types of steroids can be printed. Cholesterol, estradiol, progesterone, and cortisone are some of the steroids that can be produced.



Argument 1: An argument of "D" prints a second bond between positions 1 and 2. An argument of "Q" causes no action. All other argument values cause an outside double bond to be drawn from position 11 and the argument itself to be put at the end of the double bond as substituent formula R^{11}.

Argument 2: An argument of "D" prints a second bond between positions 3 and 4 (this double bond is shown in the diagram). An argument of "Q" causes no action. All other argument values cause an outside double bond to be drawn from position 3 and the argument itself to be put at the end of the double bond.

Argument 3: An argument of "Q" causes no action. All other argument values cause a single bond to be drawn from position 3 and the argument itself to be put at the end of the bond as substituent formula R^3.

Argument 4: An argument of "D" prints a second bond between positions 4 and 5. All other argument values cause no action.

Argument 5: An argument of "D" prints a second bond between positions 5 and 6. An argument of "Q" causes no action. All other argument values cause an outside double bond to be drawn from position 17 and the argument itself to be put at the end of the double bond.

Argument 6: An argument of "D" prints a second bond between positions 5 and 10 (shown in the diagram). An argument of "M" prints the methyl group containing carbon #19 and the bond to it. No action is taken for any other value of the argument.

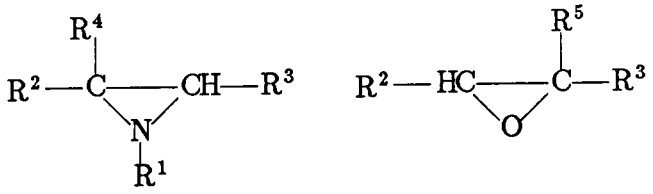
Argument 7: An argument of "Q" causes no action. All other argument values print the substituent formula beginning with carbon #20, represented by R^{20} in the diagram, and the bond to it.

Argument 8: An argument of "Q" causes no action. All other argument values print the substituent formula beginning with carbon #21, represented by R^{21} in the diagram, and the bond to it.

Argument 9: An argument of "Q" causes no action. All other argument values print the second substituent on carbon #17 and the bond to it. This substituent is shown in the diagram as R^{17}.

C. Macros for Heterocyclic Ring Systems

i. Macro \hetthree[8]. This macro typesets a 3-membered ring with one hetero atom. The common ring structures of this type are epoxides (oxirane) and ethylene imine (aziridine). Ring positions 1, 2, and 3 are the positions to which R^1, R^2, and R^3 are attached.



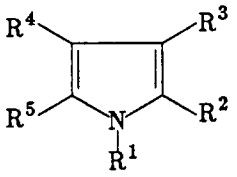
Arguments 1 – 5: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1 – R^5.

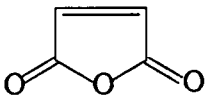
Argument 6: An argument of "S" typesets a bond to the left of ring atom #2. An argument of "H" puts —H to the left of ring atom #2. For all other argument values, no bond is drawn and R^2 is moved next to ring atom #2.

Argument 7: An argument of "S" typesets a bond to the right of ring atom #3. An argument of "H" puts H— to the right of ring atom #3. For all other argument values, no bond is drawn and R^3 is moved next to ring atom #3.

Argument 8: The atom symbol for the hetero atom.

ii. Macro \hetifive[9]. This macro typesets 5-membered rings with one hetero atom. Thus it can be used to print the pyrrole, furan, and thiophene ring systems, and their hydrogenated versions. The arguments are selected such that common compounds like proline, pyrrolidone, maleic anhydride, and vitamin C can be printed. Ring positions 1, 2, 3, 4, and 5 are the positions to which R^1 – R^5 are attached.





Arguments 1,3,4: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1, R^3, and R^4.

Argument 2: An argument value of "Q" causes no action. An argument value of "O" puts an outside double bond with an O in ring position 2. All other argument values are used as the substituent formula R^2 with a single bond.

Argument 5: An argument value of "Q" causes no action. An argument value of "O" puts an outside double bond with an O in ring position 5. All other argument values are used as the substituent formula R^5 with a single bond.

Argument 6: An argument of "D" draws a second bond between ring positions 2 and 3. All other argument values cause no action.

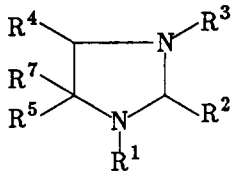
Argument 7: An argument of "D" draws a second bond between ring positions 3 and 4. All other argument values cause no action.

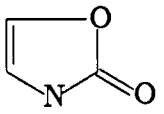
Argument 8: An argument of "D" draws a second bond between ring positions 4 and 5. All other argument values cause no action.

Argument 9: The atomic symbol of the hetero atom.

iii. Macro \heticifive[9]. This macro typesets a 5-membered ring with 2 hetero atoms separated by a carbon atom. Thus it can be used to print ring systems such as imidazole, thiazole, and oxazole. The arguments were selected

by considering actually occurring compounds containing these ring systems. Ring positions 1, 2, 3, 4, and 5 are the positions to which R^1 – R^5 are attached.





Arguments 1, 3, 5: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1, R^3, and R^5.

Argument 2: An argument of "Q"causes no action. An argument of "O" puts an outside double bond with an O in ring position 2. All other argument values are used as the substituent formula R^2 with a single bond.

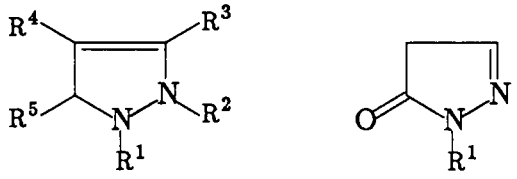
Argument 4: An argument of "Q"causes no action. An argument of "O" puts an outside double bond with an O in ring position 4. All other argument values are used as the substituent formula R^4 with a single bond.

Argument 6: An argument of "D" draws a second bond between ring positions 2 and 3. All other argument values cause no action.

Argument 7: An argument of "Q" causes no action. An argument of "D" draws a second bond between ring positions 4 and 5. All other argument values are used as the substituent formula R^7, the second substituent at ring position 5.

Arguments 8 and 9: The atomic symbols of the hetero atoms in position 1 and 3, respectively.

iv. Macro \pyrazole[8]. The pyrazole ring is found in a number of drugs, such as aminopyrine. Ring positions 1, 2, 3, 4, and 5 are the positions to which $R^1 - R^5$ are attached.



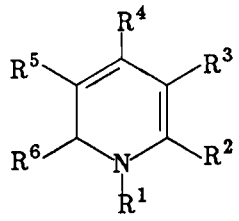
Arguments 1, 2, 4: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1, R^2, and R^4.

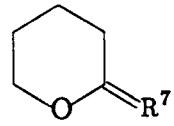
Argument 3: An argument of "Q" causes no action. An argument of "O" puts an outside double bond with an O in ring position 3. All other argument values are used as the substituent formula R^3 with a single bond.

Argument 5: An argument of "Q" causes no action. An argument of "O" puts an outside double bond with an O in ring position 5. All other argument values are used as the substituent formula R^5 with a single bond.

Arguments 6, 7, 8: An argument of "D" draws a second bond between ring positions 2 and 3, ring positions 3 and 4, and ring positions 4 and 5, respectively. All other argument values cause no action.

v. Macro \hetisix[9]. This macro typesets a six-membered ring with one hetero atom. Thus it can be used to print ring systems such as pyridine and pyran. The arguments were selected by considering actually occurring compounds such as the B vitamins niacin and pyridoxine and the coumarin ring system. Ring positions 1 – 6 are the positions to which $R^1 - R^6$ are attached.





Argument 1: An argument of "Q" causes no action. An argument of "D" prints a second bond between positions 1 and 6. All other arguments values are used as the substituent formula R^1.

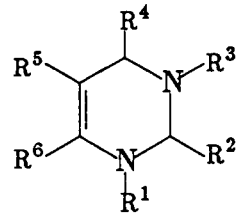
Arguments 2 – 6: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^2 – R^6.

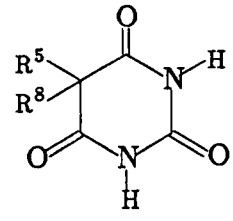
Argument 7: An argument of "Q" causes no action. An argument of "D" prints a second bond between positions 2 and 3. All other argument values cause an outside double bond to be drawn from position 2 and the argument itself to be put at the end of the double bond as R^7.

Argument 8: An argument of "D" prints a second bond between positions 4 and 5. All other argument values cause no action.

Argument 9: The atomic symbol of the hetero atom.

vi. Macro \pyrimidine[9]. The pyrimidine ring occurs in such important compounds as cytosine, thymine, uracil, and the barbiturates. The arguments of the macro were selected such that these compounds can be typeset. Ring positions 1 – 6 are the positions to which R^1 – R^6 are attached.





Arguments 1, 3, 5: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1, R^3, and R^5.

Argument 2: An argument of "Q" causes no action. An argument of "O" causes an outside double bond with an O to be drawn at position 2. All other argument values are used as the substituent formula R^2 with a single bond.

Argument 4: An argument of "Q" causes no action. An argument of "O" causes an outside double bond with an O to be drawn at position 4. All other argument values are used as the substituent formula R^4 with a single bond.

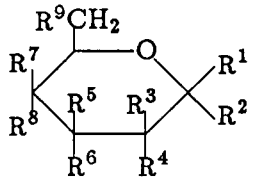
Argument 6: An argument of "Q" causes no action. An argument of "O" causes an outside double bond with an O to be drawn at position 6. All other argument values are used as the substituent formula R^6 with a single bond.

Argument 7: An argument of "D" prints a second bond between positions 1 and 2. All other argument values cause no action.

Argument 8: An argument of "Q" causes no action. An argument of "D" prints a second bond between positions 3 and 4. All other argument values are used as the second substituent in position 5, R^8.

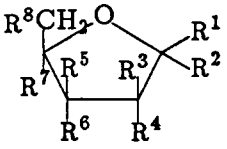
Argument 9: An argument of "D" prints a second bond between positions 5 and 6. All other argument values cause no action.

vii. Macro \pyranose[9]. This macro was developed specifically for monosaccharide structures. Carbon #1 is at the position to which R^1 is attached. Structures from this macro look best with substituent formulas in 10 point size (shown) or even smaller.



Arguments 1 – 9 are used as the respective substituent formulas R^1 – R^9. An argument of "Q" causes no action.

viii. Macro \furanose[8]. This macro was also developed specifically for monosaccharide structures. Carbon #1 is at the position to which R^1 is attached. Structures look best with substituent formulas in 10 point size (shown) or even smaller.

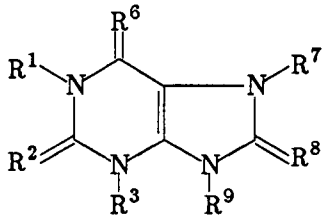


Argument 1: An argument of "Q" causes no action. An argument of "N" prints a long vertical bond at position 1, used for attachment to purine and pyrimidine bases to form nucleosides. All other argument values are used as the substituent formula R^1.

Arguments 2 – 8 An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^2 – R^8.

ix. Macro \purine[9]. The purine ring system occurs in such important compounds as adenine, guanine, caffeine, and uric acid. The arguments of the macro were selected such that these compounds can be typeset. Positions 1,

2, 3, 6, 7, 8, and 9 are indicated in the following diagram by the respective substituent formulas.



Arguments 1, 3, 6, 7, 9: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1 ... R^9.

Argument 2: An argument of "D" prints a second bond between positions 2 and 3. All other argument values cause an outside double bond to be printed at position 2 and the argument itself to be put at the end of the double bond as the substituent formula R^2.

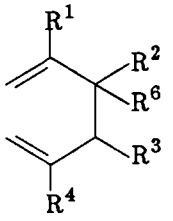
Argument 4: An argument of "D" prints a second bond between positions 1 and 6. All other argument values cause no action.

Argument 5: An argument of "Q" causes no action. . All other argument values cause an outside double bond to be printed at position 6 and the argument itself to be put at the end of the double bond as the substituent formula R^6.

Argument 8: An argument of "D" prints a second bond between positions 7 and 8. All other argument values cause an outside double bond to be printed at position 8 and the argument itself to be put at the end of the double bond as the substituent formula R^8.

D. General Utility Macros

i. Macro \fuseiv[9]. This macro typesets a fragment that is designed to be connected at two places to another ring system with the effect of fusing an additional sixring to that system. The fragment can be fused to positions 1 and 2 of the carbon fivering and the carbon sixring, and to positions 2 and 3 of the \hetifive and \hetisix rings without changing the unitlength and the \yi coordinate.



Arguments 1 – 4: An argument of "Q" causes no action. All other argument values are used as the substituent formulas R^1 – R^4.

Argument 5: An argument of "D" prints a second bond between the upper point of attachment and position 1 (this double bond is shown in the diagram). All other argument values cause no action.

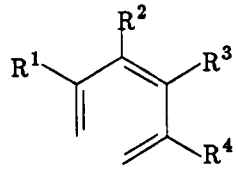
Argument 6: An argument of "Q" causes no action. An argument of "D" prints a second bond between positions 1 and 2. All other argument values are used as the substituent formula R^6.

Argument 7: An argument of "D" prints a second bond between positions 2 and 3. All other argument values cause no action.

Argument 8: An argument of "Q" causes no action. An argument of "D" prints a second bond between positions 3 and 4. All other argument values are used as a second substituent in position 3 (not shown in the diagram).

Argument 9: An argument of "D" prints a second bond from the lower point of attachment to position 4 (this double bond is shown in the diagram). All other argument values cause no action.

ii. Macro \fuseup[9]. This macro typesets a fragment that is designed to be connected at two places to another ring system with the effect of fusing an additional sixring to that system at an angle. The fragment can be fused to positions 1 and 6 of the carbon sixring and positions 3 and 4 of the \hetisix rings without changing the unitlength and the \yi coordinate.



Arguments 1 – 4: An argument of "Q" causes no action. All other argument values are used as the respective substituent formulas R^1 – R^4.

Argument 5: An argument of "D" prints a second bond from the upper point of attachment to position 1 (the resulting double bond is shown in the diagram). All other argument values cause no action.

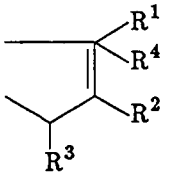
Argument 6: An argument of "D" prints a second bond between positions 1 and 2. All other argument values cause no action.

Argument 7: An argument of "D" prints a second bond between positions 2 and 3 (the resulting double bond is shown in the diagram). All other argument values cause no action.

Argument 8: An argument of "D" prints a second bond between positions 3 and 4. All other argument values cause no action.

Argument 9: An argument of "D" prints a second bond between position 4 and the lower point of attachment (the resulting double bond is shown in the diagram). All other argument values cause no action.

iii. Macro \fuseiii[6]. This macro typesets a fragment that is designed to be connected at two places to another ring system with the effect of fusing an additional fivering to that system. The fragment can be fused to positions 1 and 2 of the carbon fivering and sixring, and to positions 2 and 3 of the \hetifive and \hetisix rings without changing the unitlength and the \yi coordinate.

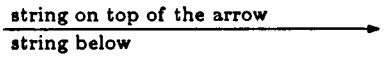


Arguments 1 – 4: An argument of "Q" causes no action. All other arguments are used as the respective substituent formulas R^1 – R^4.

Argument 5: An argument of "Q" causes no action. All other argument values are used as a second substituent in position 2 (not shown in the diagram).

Argument 6: An argument of "D" prints a second bond between positions 1 and 2. All other argument values cause no action.

iv. Macro \cto[3]. This macro draws a reaction arrow and puts the requested character strings representing reagents and reaction conditions on top and below the arrow, respectively. The arrow is made long enough to accommodate the longer of the strings. The vertical position of the arrow can be changed by changing the \yi value.



Arguments 1 and 2: The character strings above and below the arrow, respectively.

Argument 3: An integer, the number of characters – including subscripts – in the longer string.

v. Macro \sbond[1]. This macro draws a horizontal single bond of a specified length, vertically centered on a line. It should be used for structural formulas that do not use the picture environment and are written on one line.

—

The argument is an integer, expressing the length of the bond in printer points (1 pt = .35 mm).

vi. Macro \dbond[2]. This macro draws a horizontal double bond of a specified length. It should be used for structural formulas that do not use the picture environment and are written on one line.

═

Argument 1: An integer, expressing the length of the bond in printer points.

Argument 2: An integer, expressing the amount of vertical space by which the bonds have to be pushed together to give the desired vertical distance. In a document with double spacing, the number 19 produced the spacing in the double bond shown above.

vii. Macro \tbond[2]. This macro is similar to \dbond, except that it draws a triple bond:

≡

The meaning of the arguments is the same as in \dbond. The number 20 was used as argument 2.

3. COMMON REQUIREMENTS FOR THE USE OF THE SYSTEM

So far in this thesis it has been explained how to write LaTeX code to produce a chemical structure diagram at a particular place in a document. This section will discuss the mandatory and the optional statements at the beginning of an input file that make the system of macros accessible and its use more practical and convenient. Figure 6.1 contains these statements together with the two required declarations at the beginning of a LaTeX file, lines (1) and (9). (The line numbers are for reference only, they are not used in the input file.)

The part of the input file preceding the \begin{document} statement is called the "preamble" in the LaTeX Manual. In addition to the statements shown here, the preamble usually contains declarations pertaining to text formatting details such as margin width, text height on a page, and space between lines.

The statement in line (2) of figure 6.1 is necessary if the structure-drawing macros of this thesis are to be used for the preparation of a document. This statement reads the file init.tex into TeX's memory, a file that contains two short macros, \initial and \reinit. Macro \initial defines the command sequences \xi, \yi, \pw, \pht, \xbox, and \len as integer variables and assigns a count register to each of them. The use of the first four variables in the picture declaration and the use of \xbox in a minipage or parbox environment was explained in chapter III. The counter \len is a general purpose integer variable for the user. All the variables except \len are also given initial values. — Furthermore, the unitlength for the picture environments is set to 0.1 printer points in \initial. This is the recommended unitlength for the chemical structure diagrams, but it can be changed anywhere in the document. — Line (11) from figure 6.1 expands \initial.

The macro \reinit simply resets all the parameters to their initial values

```
(1)  \documentstyle[12pt]{report}
(2)  \input{init.tex}
(3)  \input{rings1.tex}
(4)  \setcounter{totalnumber}{4}
(5)  \setcounter{topnumber{2}
(6)  \setcounter{bottomnumber}{2}
(7)  \renewcommand{\topfraction}{.5}
(8)  \renewcommand{\bottomfraction}{.5}
(9)  \begin{document}
(10) \textfont1=\tenrm
(11) \initial
```

Figure 6.1: Statements at the beginning of a LaTeX file

from \initial. It is a convenience, especially for cases where more than one variable needs to be reset.

The statement in line (3) of figure 6.1 reads a file with structure-drawing macros. There will usually be several such statements, reading in different macros or sets of macros. At the installation where this work was done the TeX memory is not large enough to read in all the chemistry macros. TeX's own macros already take up 13% of the reserved memory and when the LaTeX macros are added, two thirds of the memory are used before an input file is processed. — Reading in just part of the structure-drawing macros for any given document has the advantage of saving processing time.

Lines (4) – (8) in figure 6.1 affect the placement of "floats" on the page. The only floats discussed in this thesis are the diagrams produced in the figure environment (see chapter III). In defining the style of a document – the report style is designated by line (1) – the LaTeX program sets default values for the maximum total number of floats on a page (three), the maximum number

of floats at the top of the page (two), and at the bottom of the page (one). These values can be changed for documents with an unusually large number of figures. Thus, lines (4) – (6) increase the maximum number of floats to 4, evenly distributed on the page. It is then necessary to change the counters \topfraction and \bottomfraction to reflect the distribution of figures on the page.

Finally, line(10) is the optional redefinition of the math textfont, discussed in chapter III. This definition can be changed anywhere in the document.

CHAPTER VII

CONCLUSION

The object of this thesis was two-fold: to demonstrate that TeX/LaTeX can be used for typesetting chemical structure diagrams within a document, and to make it practical and relatively convenient to produce a large variety of structures.

The diagrams shown throughout this thesis illustrate that the first goal has been accomplished to a large degree. The picture environment, which is used for the majority of the structures, gives the programmer enough control to place lines of a desired length in a specified position, and to join lines at a precise point. The fact that only a limited number of different slopes is possible for the lines did not create problems for this application of the LaTeX graphics. — It is also useful that the geometric figures which represent the chemical structures can be scaled by changing the unitlength. — Some deficiencies of the LaTeX graphics for the work described here were noted, but they are not critical. It is unfortunate that the shortest slanted line that can be drawn is not very short (about 3.6 mm), and that there is no mechanism for producing the dashed lines (other than horizontal and vertical ones) and wedges that indicate the three-dimensional configuration of a molecule. But some of the software developed for drawing chemical structures by other authors does not have the stereochemical capability either.

The fact that the LaTeX graphics proved to be suitable for this application provided a very important benefit: The structure-drawing code is completely based on TeX and thus as device independent as any TeX DVI file.

Although most of the structures were produced in the picture environment, it should be mentioned that the tabbing environment is suitable for printing some types of structures.

Another part of the first objective was the positioning and printing of condensed substituent formulas. Here LaTeX assisted with the \makebox macro for the placement of textstrings and TeX made it possible to produce the subscripts of the math mode together with a typestyle that is customary in chemistry.

The integration of the diagrams into the text of a document was still another part of the first objective. It was to be expected that LaTeX's figure environment would be suitable for displaying the structures and this was true. But in addition to it the math display environment proved to be a good mechanism for placing the structures on the page, although the environment was designed for mathematical equations.

The second goal was approached by making flexible structure fragments available through TeX's macro facility. Flexibility was built into the structure macros in several ways, most importantly through the combination of macro parameters and TeX's conditional statements. Although the LaTeX Manual does not guarantee that "plain TeX" features will work within LaTeX environments (Lamport 86, p. 204), the conditional code could be used in LaTeX's picture environment and produced the desired result of varying bonds, multiple bonds, substituents, and hetero atoms in a given structure fragment. The capability of varying substituents and bonds extends further to combining fragments from different macros into one structure. — Added flexibility in the macros is provided through a variable unitlength and variable coordinates and picture size.

It would not be realistic to expect that a system of macros can contribute building blocks for each of the infinitely many possible chemical structures; but the set of macros presented here facilitates the typesetting of a large number of common structures. — It should be emphasized again that LaTeX code as such can typeset almost any structure diagram. Even where none of the macros can be used as a fragment for a structure, the code in the macros can still assist in

building up the structure. As an example, the user may not find the complete ring system needed, but parts of the ring system could occur in different macros.

When the method of computer-based structure drawing presented here is compared with the interactive methods reviewed in chapter I, it can certainly be said that the quality of the printed structures measures up well enough. Where convenience is concerned, one would have to say that neither the use of the LaTeX macros nor interactive structure drawing is trivially easy, as this writer knows from experience with both methods. — Typesetting chemical structures with TeX/LaTeX can be the method of choice where TeX's superior text formatting capabilities and its device indepence are important considerations.

Finally, this system of structure drawing could be enhanced in several ways. An obvious extension would be a larger number of macros. The chemical intelligence in the system could be increased with additional screen messages. One might also consider writing computer programs to produce the structure-drawing LaTeX code.

BIBLIOGRAPHY

BIBLIOGRAPHY

Ash, J. E., et al. *Communication, Storage and Retrieval of Chemical Information.* Ellis Horwood Limited, Chichester, England, 1985.

Bendall, V. I. Display of organic molecules using BASIC. *J. Chem. Educ.* **1980**, *57*, 250.

Bendall, V. I. *CHEMUTIL-2. A chemistry programming utility.* Project Seraphim, Ypsilanti, MI, 1985.

Blake, J. E., Farmer, N. A., and Haines, R. C. An interactive computer graphics system for processing chemical structure diagrams. *J. Chem. Inf. Comput. Sci.* **1977**, *17*, 223–228.

Christian, K. Lasers get graphic. *PC Magazine* **1985**, *4*(19), 211–214.

Cooke-Fox, D. I., Kirby, G. H., and Rayner, J. D. From names to diagrams — by computer. *Chem. Br.* **1985**, *21*, 467–471.

Davis, M. Chemical writing on inexpensive home computers. *Chem. Br.* **1986**, *22*, 447–449.

Dickinson, J. Laser printers. *PC Magazine* **1985**, *4*(19), 208–209.

Dickinson, J. The business of Words: Special purpose. (A review of scientific word processors.) *PC Magazine* **1986**, *5*(4), 177–198.

Dittmar, P. G., Mockus, J., and Couvreur, K. M. An algorithmic computer graphics program for generating chemical structure diagrams. *J. Chem. Inf. Comput. Sci.* **1977**, *17*, 186–192.

Dittmar, P. G., et al. The CAS ONLINE search system. 1. General system design and selection, generation, and use of search screens. *J. Chem. Inf. Comput. Sci.* **1983**, *23*, 93–102.

Goodson, A. L. Graphical representation of chemical structures in Chemical Abstracts Service publications. *J. Chem. Inf. Comput. Sci.* **1980**, *20*, 212–217.

Harts, B. The little engine that could. (Laser printers.) *PC Magazine* **1985**, *4*(19), 210–211.

Johnston, M. D. MOLPIX — A program for generating and displaying molecular structures. *J. Chem. Educ.* **1985**, *62*, 413.

Kernighan, B. W. PIC — A language for typesetting graphics. *Software — Practice and Experience* **1982**, *12*, 1–21.

Knuth, D. *TeX and Metafont — new directions in typesetting.* Digital Press, Bedford, MA, 1979.

Knuth, D. *The TeXbook.* Addison-Wesley Publishing Company, Reading, MA, 1984.

Knuth, D., interviewed by Vose, G. M. and Williams G. Computer science considerations. Donald Knuth speaks on his involvement with digital typography. *Byte* **1986**, *11*(2), 169–172.

Lamport, L. *LaTeX. A Document Preparation System.* Addison-Wesley Publishing Company, Reading, MA, 1986.

Lynch, M. F., Barnard, J. M., and Welford, S. M. Generic structure storage and retrieval. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 264–270.

MacKay, P. A. Typesetting problem scripts. *Byte* **1986**, *11*(2), 201–218.

Morffew, A. J. Bibliography for molecular graphics 1983/84. *Journal of Molecular Graphics* **1984**, *2*(4), 124–128.

Nievergelt, J., et al., eds. *Document Preparation Systems.* North-Holland Publishing Company, Amsterdam-New York-Oxford, 1982.

Ryan, A. W., and Stobaugh, R. E. The Chemical Abstracts Service Chemical Registry System. 9. Input structure conventions. *J. Chem. Inf. Comput. Sci.* **1982**, *22*, 22–28.

Solomons, T. W. G. *Organic Chemistry.* Third Edition. John Wiley & Sons, New York, 1984.

Spragens, A. Graphics commands for TeX. Discussion at the TeXhax Conference. *TUGboat* **1985**, *6*(2), 66–69.

Varian, H. R. PCTeX and MicroTeX. *Byte* **1986**, *11*(4), 267–272.

Watt, L., and Kao, J. Chemical Draftsman(CDRAFT). *Comput. Chem.* **1985**, *9*(4), 269–277.

White, W. G., Swanson, S. M., and Meyer, E. F. Molecular input to a computer. *J. Chem. Educ.* **1982**, *59*, 515–516.

Wiswesser, W. J. Historic development of chemical notations. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 258–263.

Zimmerman, S. S. and Larson, M. A. Chemical symbols and special characters on the Apple microcomputer. *Comput. Chem.* **1984**, *8*(1), 13–27.

APPENDICES

APPENDIX A

LATEX CODE FOR MACROS

```
\newcommand{\initial}  {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Macro initial declares variables and initializes the      %
% variables and the unitlength.  Macro reinit resets the    %
% values of the variables and the unitlength to the         %
% original values.                                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
  \setlength{\unitlength}{.1pt}
  \newcount\xi          \newcount\yi
  \xi=0                 \yi=300
  % coordinates of lower left corner
  \newcount\pht         \pht=900           % picture height
  \newcount\pw          \pw=400            % picture width
  \newcount\xbox        \xbox=50           % width of minipage
  \newcount\len  }     % general purpose variable
  % end macro initial
\newcommand{\reinit}   {\xi=0   \yi=300    \xbox=50
            \setlength{\unitlength}{.1pt}
            \pht=900 \pw=400  }   % end macro reinit


\newcommand{\cbranch}[9]    {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The cbranch macro draws vertical branches as single and   %
% double bonds, up and down.                                %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
\begin{picture}(\pw,\pht)(-\xi,-\yi)
             \put(0,200)     {#1}                  % upper subst.
 \ifx#2S     \put(40,85)     {\line(0,1)  {100}}  % single up
 \else\ifx#2D\multiput(27,85)(26,0){2}             % double up
                             {\line(0,1)  {100}} \fi \fi
 \ifx#4Q     \put(-305,0)    {\makebox(300,87)[r]{#3}}
                             % left substituent without bond
 \else       \put(-455,0)    {\makebox(300,87)[r]{#3}} \fi
                             % left substituent with bond
 \ifx#4S     \put(-150,33)   {\line(1,0)  {140}}  % single left
 \else\ifx#4D\multiput(-150,20)(0,26){2}           % double left
```

```
                        {\line(1,0)  {140}} \fi \fi
          \put(0,0)     {#5}                % center
                                            %  atom(s)
 \ifx#6S  \put(90,33)   {\line(1,0)  {140}} % single right
 \else\ifx#6D\multiput(90,20)(0,26){2}      % double right
                        {\line(1,0)  {140}} \fi \fi
          \put(240,0)   {#7}                % right subst.
 \ifx#8S  \put(40,-15)  {\line(0,-1) {100}} % single down
 \else\ifx#8D\multiput(27,-15)(26,0){2}     % double down
                        {\line(0,-1) {100}} \fi \fi
          \put(0,-210)  {#9}                % lower subst.
\end{picture}         }  % end cbranch macro
```

```
\newcommand{\tbranch}[7]    {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Macro tbranch draws structures with vertical branches,   %
% single or double bonds, going up or down.                %
% This macro uses the LaTeX tabbing mechanism.             %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
\begin{minipage}{\xbox pt}
 \begin{tabbing}
  $#3$\= $#4$\+ \kill
        $#1$   \\ [-#7pt]      % print top subst.
  \ifx#2S                      % vertical bond going up
        \hspace{4pt}\rule{0.4pt}{8pt}  \\ [-#7pt]\fi
  \ifx#2D
        \hspace{2pt}\rule{0.4pt}{8pt}
        \hspace{-2pt}\rule{0.4pt}{8pt} \\ [-#7pt]\fi
        \- \kill
  $#3$\> $#4$\+ \\ [-#7pt]     % substituents on print line
  \ifx#5S                      % vertical bond going down
        \hspace{4pt}\rule{0.4pt}{8pt}  \\ [-#7pt]\fi
  \ifx#5D
        \hspace{2pt}\rule{0.4pt}{8pt}
        \hspace{-2pt}\rule{0.4pt}{8pt} \\ [-#7pt]\fi
        $#6$
 \end{tabbing}
\end{minipage}    }            % end tbranch macro
```

```
\newcommand{\ethene}[4]     {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a horizontal ethene fragment with    %
% four variable substituents.                               %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(-405,170)  {\makebox(300,87)[r]{#1}}          % upper left
                                                     %  subst.
  \put(-405,-185) {\makebox(300,87)[r]{#3}}          % lower left
                                                     %  subst.
  \put(0,70)      {\line(-1,1)          {100}}       % NW bond
  \put(0,0)       {\line(-1,-1)         {100}}       % SW bond
  \put(0,0)       {C}                                % left C
  \multiput(90,20)(0,25){2} {\line(1,0){140}}        % double bond
  \put(240,0)     {C}                                % right C
  \put(315,70)    {\line(1,1)           {100}}       % NE bond
  \put(315,0)     {\line(1,-1)          {100}}       % SE bond
  \put(425,170)   {#2}                               % upper right
                                                     %  subst.
  \put(425,-170)  {#4}                               % lower right
                                                     %  subst.
 \end{picture}              }                  % end ethene macro


\newcommand{\upethene}[4]    {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a vertical ethene fragment with four   %
% variable substituents.                                      %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(-430,360)          {\makebox(300,87)[r]{#1}} % NW subst.
  \put(-430,-150)         {\makebox(300,87)[r]{#2}} % SW subst.
  \put(210,370)           {#3}                      % NE subst.
  \put(210,-140)          {#4}                      % SE subst.
  \put(0,300)             {\line(-5,3)  {121}}      % NW bond
  \put(0,230)             {C}                       % upper C
  \put(20,80)             {\line(0,1)   {140}}      % vertical
  \put(46,80)             {\line(0,1)   {140}}      %  d. bond
  \put(0,0)               {C}                       % lower C
  \put(0,0)               {\line(-5,-3) {121}}      % SW bond
```

```
   \put(80,300)              {\line(5,3)   {121}}      % NE bond
   \put(80,0)                {\line(5,-3)  {121}}      % SE bond
  \end{picture}              }              % end upethene macro


\newcommand{\cright}[7]        {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a trigonal fragment, opening to the     %
% right. The fragment has a variable center atom and three    %
% variable substituents. Bonds can be single or double.      %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
  \begin{picture}(\pw,\pht)(-\xi,-\yi)
   \ifx#2Q \put(-305,-15){\makebox(300,87)[r]{#1}}   % left sub.
    \else  \put(-455,-15){\makebox(300,87)[r]{#1}} \fi
   \ifx#2S \put(-150,33)     {\line(1,0)   {140}} \fi % single
                                                      % hor. bond
   \ifx#2D \put(-150,20)     {\line(1,0)   {140}}     % hor.
           \put(-150,46)     {\line(1,0)   {140}} \fi %  d. bond
   \put(0,0)                 {#3}                     % center
                                                      %  atoms
   \ifx#4S \put(80,70)       {\line(1,1)   {100}} \fi % NE single
                                                      %  bond
   \ifx#4D \put(71,79)       {\line(1,1)   {100}}
           \put(89,61)       {\line(1,1)   {100}} \fi % NE double
   \put(185,170)             {#5}                     % NE subst.
   \ifx#6S \put(80,0)        {\line(1,-1)  {100}} \fi % SE single
   \ifx#6D \put(71,-9)       {\line(1,-1)  {100}}     % SE double
           \put(89,9)        {\line(1,-1)  {100}} \fi %  bond
   \put(185,-170)            {#7}                     % SE subst.
  \end{picture}              }              % end cright macro


\newcommand{\cleft}[7]         {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a trigonal fragment, opening to the     %
% left. The fragment has a variable center atom and three     %
% variable substituents. Bonds can be single or double.      %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
  \begin{picture}(\pw,\pht)(-\xi,-\yi)
   \put(-405,160)            {\makebox(300,87)[r]{#1}} % NW subst.
   \ifx#2S \put(0,70)        {\line(-1,1)  {100}} \fi  % NW single
```

```
                                                 %  bond
  \ifx#2D \put(9,79)        {\line(-1,1)  {100}}       % NW double
          \put(-9,61)       {\line(-1,1)  {100}} \fi   %  bond
  \put(0,0)                 {#3}                       % center
                                                       %  atoms(s)
  \ifx#4S \put(0,0)         {\line(-1,-1) {100}} \fi   % SW single
  \ifx#4D \put(-9,9)        {\line(-1,-1) {100}}       % SW double
          \put(9,-9)        {\line(-1,-1) {100}} \fi   %  bond
  \put(-405,-185)           {\makebox(300,87)[r]{#5}}  % SW subst.
  \ifx#6S \put(90,33)       {\line(1,0)   {140}} \fi   % hor.
                                                       %  single
  \ifx#6D \put(90,20)       {\line(1,0)   {140}}       %  double
          \put(90,46)       {\line(1,0)   {140}} \fi   %  bond
  \put(240,0)               {#7}                       % right sub.
 \end{picture}              }                  % end cleft macro


\newcommand{\chemup}[7]       {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a trigonal fragment, opening         %
% upwards.  The fragment has a variable center atom and    %
% three variable substituents.  Bonds are single or        %
% double.                                                  %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(-430,130)            {\makebox(300,87)[r]{#1}}  % NW subst.
  \ifx#2S \put(0,70)        {\line(-5,3)  {121}} \fi   % NW single
  \ifx#2D \put(7,81)        {\line(-5,3)  {121}}       % NW double
          \put(-7,59)       {\line(-5,3)  {121}} \fi   %    bond
  \put(0,0)                 {#3}                       % center
                                                       %  atom(s)
  \ifx#4S \put(33,-10)      {\line(0,-1)  {140}} \fi   % vertical
                                                       %  single
  \ifx#4D \put(20,-10)      {\line(0,-1)  {140}}       % vertical
          \put(46,-10)      {\line(0,-1)  {140}} \fi   %  double
  \put(0,-230)              {#5}                       % bottom
                                                       %  subst.
  \ifx#6S \put(80,70)       {\line(5,3)   {121}} \fi   % NE single
  \ifx#6D \put(73,81)       {\line(5,3)   {121}}       % NE double
          \put(87,59)       {\line(5,3)   {121}} \fi   %    bond
  \put(210,140)             {#7}                       % NE subst.
 \end{picture}              }                  % end chemup macro
```

```
\newcommand{\cdown}[7]      {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a trigonal fragment, opening                          %
% downwards. The fragment has a variable center atom and                     %
% three variable substituents. Bonds are single or double.                  %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(0,230)               {#1}                          % upper sub.
  \ifx#2S \put(33,80)       {\line(0,1)   {140}} \fi      % vert.
                                                          %  single
  \ifx#2D \put(20,80)       {\line(0,1)   {140}}          %  double
          \put(46,80)       {\line(0,1)   {140}} \fi
  \put(0,0)                 {#3}                          % center
                                                          %  atom(s)
  \ifx#4S \put(0,0)         {\line(-5,-3) {121}} \fi      % SW single
  \ifx#4D \put(-7,11)       {\line(-5,-3) {121}}          % SW double
          \put(7,-11)       {\line(-5,-3) {121}} \fi      %    bond
  \put(-430,-150)           {\makebox(300,87)[r]{#5}}     % SW subst.
  \ifx#6S \put(80,0)        {\line(5,-3)  {121}} \fi      % SE bond
  \ifx#6D \put(87,11)       {\line(5,-3)  {121}}          % SE double
          \put(73,-11)      {\line(5,-3)  {121}} \fi      %    bond
  \put(210,-140)            {#7}                          % SE subst.
 \end{picture}              }              % end cdown macro


\newcommand{\csquare}[5]    {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a fragment that consists of a variable                 %
% center atom with four variable substituents pointing to                    %
% the four corners of a square.                                              %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(-405,160)            {\makebox(300,87)[r]{#1}}     % NW subst.
  \put(0,70)                {\line(-1,1)  {100}}          % NW bond
  \put(0,0)                 {#3}                          % center
                                                          %  atom
  \put(0,0)                 {\line(-1,-1) {100}}          % SW bond
  \put(-405,-185)           {\makebox(300,87)[r]{#4}}     % SW subst.
  \put(80,70)               {\line(1,1)   {100}}          % NE bond
```

```
  \put(185,170)          {#2}                          % NE subst.
  \put(80,0)             {\line(1,-1)  {100}}          % SE bond
  \put(185,-170)         {#5}                          % SE subst.
\end{picture}            }                        % end csquare macro


\newcommand{\ccirc}[4]     {
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 % The ccirc macro draws a circle with 2 substituents          %
 % infront of the circle and 2 behind it to give a             %
 % threedimensional impression.                                %
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 %
\begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(90,0)     {\circle{180}}
  \put(90,90)    {\line(0,1)   {70}}         % behind and up
  \put(60,170)   {#1}
  \thicklines
  \put(30,10)    {\line(-5,2)  {140}}        % in front
  \put(-415,30)  {\makebox(300,87)[r]{#2}}   %  and left
  \put(150,10)   {\line(5,2)   {140}}        % in front
  \put(300,30)   {#3}                        %  and right
  \thinlines
  \put(90,-90)   {\line(0,-1) {90}}          % behind and
  \put(60,-260)  {#4}                        %  down
\end{picture}     }                 % end ccirc macro
```

```
\newcommand{\threering}[9]      {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The threering macro typesets the cyclopropane ring      %
% with optional substituents, an optional double bond,    %
% and a plus inside the ring for aromaticity.              %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
   \begin{picture}(\pw,\pht)(-\xi,-\yi)
     \put(300,0)      {\line(-3,-5) {150}}     % bond 1 to 2
     \put(150,-244)   {\line(-3,5)  {150}}     % bond 2 to 3
     \put(0,0)        {\line(1,0)   {300}}     % bond 3 to 1
     \ifx#7D \put(40,-40){\line(1,0){220}}\fi % double 3 to 1
     \ifx#1Q                                   % subst. on 1
       \else\put(300,0)     {\line(5,3)   {128}}
            \put(433,50)    {#1}          \fi
     \ifx#2Q                                   % subst. on 2
       \else\ifx#9C \put(150,-244) {\line(0,-1){100}}
                    \put(114,-424) {#2}  % on straight bond
                \else\put(150,-244) {\line(5,-3)  {128}}
                    \put(283,-344) {#2} \fi \fi % slanted
     \ifx#2Q                                   % subst. on 2
       \else\ifx#5Q \put(150,-244) {\line(0,-1){100}}
                    \put(114,-424) {#2}  % on straight bond
                \else\put(150,-244) {\line(5,-3)  {128}}
                    \put(283,-344) {#2} \fi \fi % slanted
     \ifx#3Q                                   % subst. on 3
       \else\put(0,0)       {\line(-5,3)  {128}}
            \put(-430,34) {\makebox(300,87)[r]{#3}} \fi
     \ifx#4Q                                   % second subst.
       \else\put(300,0)     {\line(5,-3)  {128}} %  on 1
            \put(433,-100) {#4}          \fi
     \ifx#5Q                                   % second subst.
       \else\put(150,-244) {\line(-5,-3) {128}} %  on 2
            \put(-280,-360){\makebox(300,87)[r]{#5}}  \fi
     \ifx#6Q                                   % second subst.
       \else\put(0,0)       {\line(-5,-3) {128}} %  on 3
            \put(-430,-116){\makebox(300,87)[r]{#6}} \fi
     \ifx#8Q
       \else\multiput(135,-244)(30,0){2}       % outside
            {\line(0,-1)    {100}}             %  double
            \put(114,-424) {#8}           \fi %  on 2
     \ifx#9C \put(150,-90){\circle{120}}    % circle with +
```

```
          \put(110,-120){+}              \fi
    \end{picture}              }   % end cycloprop. macro


\newcommand{\fourring}[9]      {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The fourring macro typesets the cyclobutane ring with      %
% optional substituents and double bonds.                    %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
   \begin{picture}(\pw,\pht)(-\xi,-\yi)
    \put(300,300)   {\line(0,-1)  {300}}      % bond 1 to 2
    \put(300,0)     {\line(-1,0)  {300}}      %      2 to 3
    \put(0,0)       {\line(0,1)   {300}}      %      3 to 4
    \put(0,300)     {\line(1,0)   {300}}      %      4 to 1
    \ifx#7D\put(260,260){\line(0,-1){220}}\fi % double 1 to 2
    \ifx#8D\put(40,40)  {\line(0,1) {220}}\fi % double 3 to 4
    \ifx#1Q                                   % subst. on 1
      \else\put(300,300){\line(5,3) {128}}
           \put(433,350){#1}              \fi
    \ifx#2Q                                   % subst. on 2
      \else\put(300,0)  {\line(5,-3){128}}
           \put(433,-100){#2}             \fi
    \ifx#3Q                                   % subst. on 3
      \else\put(0,0)    {\line(-5,-3){128}}
           \put(-430,-116){\makebox(300,87)[r]{#3}} \fi
    \ifx#4Q                                   % subst. on 4
      \else\put(0,300)  {\line(-5,3) {128}}
           \put(-430,334){\makebox(300,87)[r]{#4}}  \fi
    \ifx#5Q                                   % second subst.
      \else\put(300,300){\line(5,-3) {128}}   %  on 1
           \put(433,200){#5}              \fi
    \ifx#6Q                                   % second subst.
      \else\put(0,300)  {\line(-5,-3){128}}   %  on 4
           \put(-430,184){\makebox(300,87)[r]{#6}}  \fi
    \ifx#9Q                                   % outs. double
      \else\multiput(280,-5)(20,30){2}        %  and subst.
           {\line(1,-1){100}} \put(405,-140){#9}   \fi % on 2
  \end{picture}          }              % end cyclobutane macro
```

```
\newcommand{\fivering}[9]    {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets the cyclopentane ring with optional     %
% substituents and double bonds. A minus sign in a circle     %
% can be put inside the ring to denote aromaticity.           %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
   \begin{picture}(\pw,\pht)(-\xi,-\yi)
    \put(342,200)    {\line(0,-1)   {200}}       % bond 1 to 2
    \put(342,0)      {\line(-5,-3)  {171}}       % bond 2 to 3
    \put(171,-103)   {\line(-5,3)   {171}}       % bond 3 to 4
    \put(0,0)        {\line(0,1)    {200}}       % bond 4 to 5
    \put(0,200)      {\line(1,0)    {342}}       % bond 5 to 1
    \ifx#1Q                                      % subst. on 1
      \else\put(342,200)   {\line(5,3)  {128}}
           \put(475,250)   {#1}              \fi
    \ifx#2Q                                      % subst. on 2
      \else\put(342,0)     {\line(5,-3) {128}}
           \put(475,-100)  {#2}              \fi
    \ifx#3Q                                      % subst. on 3
      \else\put(171,-103)  {\line(0,-1) {100}}
           \put(150,-283)  {#3}              \fi
    \ifx#4Q                                      % subst. on 4
      \else\put(0,0)       {\line(-5,-3){128}}
           \put(-430,-116) {\makebox(300,87)[r]{#4}}  \fi
    \ifx#5Q                                      % subst. on 5
      \else\put(0,200)     {\line(-5,3) {128}}
           \put(-430,234)  {\makebox(300,87)[r]{#5}}  \fi
    \ifx#6D\put(316,174)   {\line(0,-1) {148}}   % double 1,2
      \else\ifx#6S
             \else\put(342,200) {\line(5,-3) {128}}
                  \put(475,100){#6}   \fi        % second sub.
    \fi                                          %  on 1
    \ifx#7D\put(26,26)     {\line(0,1)  {148}}   % double 4,5
      \else\ifx#7S
             \else\put(0,200){\line(-5,-3){128}} % second sub.
                  \put(-430,84){\makebox(300,87)[r]{#7}}  \fi
    \fi                                          %  on 5
    \ifx#8Q                                      % outs. double
      \else\multiput(156,-103)(30,0){2}          %  and subst.
           {\line(0,-1) {100}} \put(135,-283){#8}  \fi  % on 3
    \ifx#9C\put(171,60)    {\circle{210}}        % circle and
```

```
          \put(130,35)    {$-$}       \fi       %  minus
   \end{picture}          }                     % end 5-ring macro


\newcommand{\sixring}[9]   {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The sixring macro draws standard carbon sixrings in the    %
% shape of a regular hexagon. There are optional ring        %
% double bonds and substituents.                             %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
\begin{picture}(\pw,\pht)(-\xi,-\yi)
 \put(342,200)  {\line(0,-1)  {200}}          % bond 1 to 2
 \put(342,0)    {\line(-5,-3) {171}}          % bond 2 to 3
 \put(171,-103) {\line(-5,3)  {171}}          % bond 3 to 4
 \put(0,0)      {\line(0,1)   {200}}          % bond 4 to 5
 \put(0,200)    {\line(5,3)   {171}}          % bond 5 to 6
 \put(171,303)  {\line(5,-3)  {171}}          % bond 6 to 1
 \ifx#7D                                      % d. bond 1,2
   \put(316,174)   {\line(0,-1)  {148}}
   \else \ifx#7S                              % 2. sub. on 1
         \else\put(342,200) {\line(5,-3) {128}}
              \put(475,100) {#7}    \fi  \fi
 \ifx#8D
   \put(162,-67)   {\line(-5,3)  {126}}       % double 3 to 4
   \else\ifx#8S
        \else\put(156,-203) {\line(0,1){100}} % outside
                                              %  double and
             \put(186,-203) {\line(0,1){100}} %  subst. on 3
             \put(135,-283) {#8}    \fi  \fi
 \ifx#9D
   \put(36,191)    {\line(5,3)   {126}} \fi   % double 5 to 6
 \ifx#9C \put(171,100) {\circle{250}}   \fi   % circle for
                                              %  aromaticity
 \ifx#1Q                                      % subst. on 1
   \else\put(342,200)  {\line(5,3){128}}
        \put(475,250)  {#1}             \fi
 \ifx#2Q
   \else\put(342,0)    {\line(5,-3){128}}     % subst. on 2
        \put(475,-100)    {#2}          \fi
 \ifx#3Q
   \else\put(171,-203) {\line(0,1){100}}
        \put(150,-283) {#3}             \fi   % subst. on 3
```

```
  \ifx#4Q
     \else\put(0,0)      {\line(-5,-3){128}}     % subst. on 4
          \put(-430,-116){\makebox(300,87)[r]{#4}}  \fi
  \ifx#5Q
    \else\put(0,200)     {\line(-5,3){128}}      % subst. on 5
         \put(-430,234) {\makebox(300,87)[r]{#5}}   \fi
  \ifx#6Q
    \else\put(171,303)  {\line(0,1){100}}       % subst. on 6
         \put(150,410)  {#6}              \fi
  \ifx#7D  \ifx#9C \message{Error: ring double bond
                            simultaneous with circle}
  \fi   \fi
 \end{picture}           }                % end sixring macro


\newcommand{\sixringa}[9]   {
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 % This macro differs from the original sixring macro only    %
 % in the position of the double bonds in the ring.           %
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 %
   \begin{picture}(\pw,\pht)(-\xi,-\yi)
     \put(342,200)  {\line(0,-1)  {200}}          % bond 1 to 2
     \put(342,0)    {\line(-5,-3) {171}}          % bond 2 to 3
     \put(171,-103) {\line(-5,3)  {171}}          % bond 3 to 4
     \put(0,0)      {\line(0,1)   {200}}          % bond 4 to 5
     \put(0,200)    {\line(5,3)   {171}}          % bond 5 to 6
     \put(171,303)  {\line(5,-3)  {171}}          % bond 6 to 1
     \ifx#7D                                      % double 1,6
       \put(306,191)   {\line(-5,3)  {126}}
       \else \ifx#7S                              % second subst.
             \else\put(342,200) {\line(5,-3) {128}}
                  \put(475,100) {#7}    \fi  \fi  %  on 1
     \ifx#8D
       \put(178,-67)   {\line(5,3)   {126}}       % double 3,2
       \else\ifx#8S
            \else\put(156,-203) {\line(0,1){100}} % outs. double
                 \put(186,-203) {\line(0,1){100}} %  and subst.
                 \put(135,-283) {#8}    \fi   \fi %  on 3
     \ifx#9D
       \put(26,26)     {\line(0,1)   {148}} \fi   % double 4,5
     \ifx#9C \put(171,100) {\circle{250}}   \fi   % circle for
                                                  %  aromaticity
```

```
    \ifx#1Q                                                   % subst. on 1
      \else\put(342,200)  {\line(5,3){128}}
           \put(475,250)  {#1}              \fi
    \ifx#2Q
      \else\put(342,0)    {\line(5,-3){128}}         % subst. on 2
           \put(475,-100)    {#2}           \fi
   \ifx#3Q
      \else\put(171,-203) {\line(0,1){100}}
           \put(150,-283) {#3}              \fi   % subst. on 3
    \ifx#4Q
       \else\put(0,0)      {\line(-5,-3){128}}    % subst. on 4
            \put(-430,-116){\makebox(300,87)[r]{#4}} \fi
    \ifx#5Q
      \else\put(0,200)    {\line(-5,3){128}}      % subst. on 5
           \put(-430,234) {\makebox(300,87)[r]{#5}}  \fi
    \ifx#6Q
      \else\put(171,303)  {\line(0,1){100}}       % subst. on 6
           \put(150,410)  {#6}              \fi
    \ifx#7D  \ifx#9C \message{Error: ring double bond with
                                   circle} \fi   \fi
   \end{picture}          }              % end sixringa macro


\newcommand{\sixringb}[9]    {
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 % This variation of the sixring can typeset all combina-     %
 % tions of double bonds in the ring through argument 9.      %
 % A para-quinoid structure is also possible.                 %
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 % macros to typeset the ring double bonds:
\newcommand{\di}    {\put(316,174){\line(0,-1){148}}} %double 1-2
\newcommand{\dii}   {\put(178,-67){\line(5,3)  {126}}} %double 2-3
\newcommand{\diii}  {\put(162,-67){\line(-5,3){126}}} %double 3-4
\newcommand{\dfour}{\put(26,26)   {\line(0,1)  {148}}} %double 4-5
\newcommand{\dv}    {\put(36,191)  {\line(5,3)  {126}}} %double 5-6
\newcommand{\dsix}  {\put(306,191){\line(-5,3){126}}} %double 6-1
   \begin{picture}(\pw,\pht)(-\xi,-\yi)
     \put(342,200)   {\line(0,-1)   {200}}            % bond 1 to 2
     \put(342,0)     {\line(-5,-3)  {171}}            % bond 2 to 3
     \put(171,-103)  {\line(-5,3)   {171}}            % bond 3 to 4
     \put(0,0)       {\line(0,1)    {200}}            % bond 4 to 5
     \put(0,200)     {\line(5,3)    {171}}            % bond 5 to 6
     \put(171,303)   {\line(5,-3)   {171}}            % bond 6 to 1
```

```
\ifx#7Q
 \else\put(158,303) {\line(0,1) {100}}       % outs. double
      \put(184,303) {\line(0,1) {100}}       %  bond and
      \put(130,420) {#7}    \fi              %  sub. on 6
\ifx#8Q
 \else\put(156,-203){\line(0,1){100}}        % outs. double
      \put(186,-203){\line(0,1){100}}        %  bond and
      \put(135,-283) {#8}   \fi              %  sub.on 3
% 17 double bond combinations:
\ifcase#9 \put(171,100) {\circle{250}}        % circle
  \or \di \or \dii \or \or diii \or \di \diii % arg 9=1-5
  \or \or \or dfour \or \dfour \di            % arg 9=6-9
  \or \dfour \dii \or \or \or \or \or \or \dv % arg 9=10-16
  \or \dv \di \or \dv \dii \or \or \dv \diii  % arg 9=17-20
  \or \dv \diii \di \or \or \or \or \or \or   % arg 9=21-27
  \or \or \or \or \or \dsix \or \or \dsix \dii %      28-34
  \or \or \dsix \diii \or \or \or             % arg 9=35-39
  \or \dsix \dfour \or \or \dsix \dfour \dii  % arg 9=40-42
\fi
\ifx#1Q                                       % subst. on 1
  \else\put(342,200)  {\line(5,3){128}}
       \put(475,250)  {#1}       \fi
\ifx#2Q
  \else\put(342,0)    {\line(5,-3){128}}     % subst. on 2
       \put(475,-100)   {#2}    \fi
\ifx#3Q
  \else\put(171,-203) {\line(0,1){100}}
       \put(150,-283) {#3}       \fi         % subst. on 3
\ifx#4Q
   \else\put(0,0)     {\line(-5,-3){128}}    % subst. on 4
        \put(-430,-116){\makebox(300,87)[r]{#4}}  \fi
\ifx#5Q
  \else\put(0,200)    {\line(-5,3){128}}     % subst. on 5
       \put(-430,234) {\makebox(300,87)[r]{#5}}   \fi
\ifx#6Q
  \else\put(171,303)  {\line(0,1){100}}      % subst. on 6
       \put(150,410)  {#6}       \fi
\ifx#7D  \ifx#9C \message{Error: ring double bond with
                                 circle}  \fi  \fi
\end{picture}          }              % end sixringb macro
```

```
\newcommand{\chair}[8]       {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The chair macro typesets the saturated carbon sixring in   %
% its most favorable conformation. Axial and equatorial      %
% substituents can be attached in 4 positions.               %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
  \begin{picture}(\pw,\pht)(-\xi,-\yi)
   \put(0,0)        {\line(3,-4)  {170}}      % bond 1 to 2
   \put(170,-226)   {\line(3,1)   {403}}      %      2 to 3
   \put(573,-91)    {\line(6,-1)  {210}}      %      3 to 4
   \put(783,-126)   {\line(-3,4)  {170}}      %      4 to 5
   \put(613,100)    {\line(-3,-1) {403}}      %      5 to 6
   \put(210,-35)    {\line(-6,1)  {210}}      %      6 to 1
                                          % bonds to subst. :
   \put(0,0)        {\line(0,1)   {100}}      % axial on 1
   \put(170,-226)   {\line(0,-1)  {100}}      % axial on 2
   \put(573,-91)    {\line(0,1)   {100}}      % axial on 3
   \put(783,-126)   {\line(0,-1)  {100}}      % axial on 4
   \put(613,100)    {\line(0,1)   {100}}      % axial on 5
   \put(210,-35)    {\line(0,-1)  {100}}      % axial on 6
   \put(0,0)        {\line(-3,-1) {128}}      % eq. on 1
   \put(170,-226)   {\line(-3,1)  {128}}      % eq. on 2
   \put(573,-91)    {\line(3,-1)  {128}}      % eq. on 3
   \put(783,-126)   {\line(3,1)   {128}}      % eq. on 4
   \put(613,100)    {\line(3,-1)  {128}}      % eq. on 5
   \put(210,-35)    {\line(-3,1)  {128}}      % eq on 6
                                          % variable subst.:
   \ifx#1Q\else\put(-30,110)     {#1}    \fi % axial on 1
   \ifx#2Q\else\put(140,-406)    {#2}    \fi % axial on 2
   \ifx#3Q\else\put(543,9)       {#3}    \fi % axial on 3
   \ifx#4Q\else\put(753,-306)    {#4}    \fi % axial on 4
   \ifx#5Q\else\put(-430,-85) {\makebox(300,87)[r]{#5}}
   \fi                                    % eq. on 1
   \ifx#6Q\else\put(-260,-226){\makebox(300,87)[r]{#6}}
   \fi                                    % eq. on 2
   \ifx#7Q\else\put(415,-230) {\makebox(300,87)[r]{#7}}
   \fi                                    % eq. on 3
   \ifx#8Q\else\put(916,-115)    {#8}    \fi % eq. on 4
  \end{picture}              }         % end chair macro
```

```
\newcommand{\naphth}[9]         {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets the naphthalene ring system.              %
% One or both rings can be saturated. An optional sub-          %
% stituent is possible at each ring position.                   %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
  \begin{picture}(\pw,\pht)(-\xi,-\yi)
   \multiput(0,200)(342,0){2}    {\line(5,3) {171}} % 7,8; 8a,1
   \multiput(171,303)(342,0){2} {\line(5,-3){171}} % 8,8a; 1,2
   \multiput(0,0)(342,0){3}     {\line(0,1) {200}} % 6,7;4a,8a
                                                   % 3,2
   \multiput(0,0)(342,0){2}     {\line(5,-3){171}} % 6,5; 4a,4
   \multiput(171,-103)(342,0){2}{\line(5,3) {171}} % 5,4a; 4,3
   \ifx#1Q
     \else\put(513,303)         {\line(0,1) {100}}  % sub. on 1
          \put(492,410)         {#1}    \fi
   \ifx#2Q
     \else\put(684,200)         {\line(5,3) {128}}  % sub. on 2
          \put(817,250)         {#2}    \fi
   \ifx#3Q
     \else\put(684,0)           {\line(5,-3){128}}  % sub. on 3
          \put(817,-100)        {#3}    \fi
   \ifx#4Q
     \else\put(513,-103)        {\line(0,-1){100}}  % sub. on 4
          \put(492,-283)        {#4}    \fi
   \ifx#5Q
     \else\put(171,-103)        {\line(0,-1){100}}  % sub. on 5
          \put(150,-283)        {#5}    \fi
   \ifx#6Q
     \else\put(0,0)             {\line(-5,-3){128}} % sub. on 6
          \put(-430,-116) {\makebox(300,87)[r]{#6}}    \fi
   \ifx#7Q
     \else\put(0,200)           {\line(-5,3) {128}} % sub. on 7
          \put(-430,234)  {\makebox(300,87)[r]{#7}}    \fi
   \ifx#8Q
     \else\put(171,303)         {\line(0,1)  {100}} %sub. on 8
          \put(150,410)         {#8}    \fi
   \ifx#9S                                          %all single
     \else\put(322,180)         {\line(0,-1) {160}} %double
                                                    %  4a,8a
          \put(164,-76)         {\line(-5,3) {147}} %double 5,6
```

```
        \put(27,193)          {\line(5,3)  {147}} %double 7,8
   \fi
   \ifx#9D\put(517,280)       {\line(5,-3) {147}} %double 1,2
          \put(517,-80)       {\line(5,3)  {147}} %double 4,3
   \fi
 \end{picture}       }                  % end naphthalene macro


\newcommand{\terpene}[9]    {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets the bicyclo(2.2.1)heptane ring                            %
% system with optional methyl groups at the one-carbon                          %
% bridge.                                                                       %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
\begin{picture}(\pw,\pht)(-\xi,-\yi)
 \put(0,0)        {\line(5,1)  {196}}        % bond 5 to 4
 \put(196,39)     {\line(5,-2) {186}}        % bond 4 to 3
 \put(382,-35)    {\line(2,5)  {66}}         % bond 3 to 2
 \put(448,130)    {\line(-5,2) {186}}        % bond 2 to 1
 \put(262,204)    {\line(-5,-1){196}}        % bond 1 to 6
 \put(66,165)     {\line(-2,-5){66}}         % bond 6 to 5
 \put(196,39)     {\line(0,1)  {330}}        % long part of
                                             %  bridge
 \put(262,204)    {\line(-2,5) {66}}         % shorter part
                                             %  of bridge
 \ifx#1Q                                     % subst. on 1
   \else\put(262,204)   {\line(4,3)  {120}}
        \put(387,267)   {#1}            \fi
 \ifx#2Q                                     % subst. on 2
   \else\put(448,130)   {\line(4,3)  {120}}
        \put(573,193)   {#2}            \fi
 \ifx#3Q                                     % subst. on 3
   \else\put(382,-35)   {\line(5,-2) {120}}
        \put(507,-121)  {#3}            \fi
 \ifx#4Q                                     % subst. on 4
   \else\put(196,39)    {\line(2,-5) {56}}
        \put(231,-180)  {#4}            \fi
 \ifx#5Q                                     % subst. on 5
   \else\put(0,0)       {\line(-5,-2){139}}
        \put(-441,-95)  {\makebox(300,87)[r]{#5}}  \fi
 \ifx#6Q                                     % subst. on 6
   \else\put(66,165)    {\line(-4,3) {120}}
```

```
       \put(-362,216)  {\makebox(300,87)[r]{#6}}  \fi
 \ifx#7M\put(196,369)  {\line(4,3)  {120}} % methyl
       \put(196,369)   {\line(-4,3) {120}} %  groups
       \put(321,432)   {\small ${\rm CH_3}$}
       \put(-226,425)                      %  on bridge
       {\makebox(300,87)[r]{\small ${\rm CH_3}$}} \fi
 \ifx#8Q\else
 \ifx#8O\put(443,142)  {\line(5,2)  {130}} % double-
       \put(453,118)   {\line(5,2)  {130}} %  bonded
       \put(583,155)   {$O$}               %  O on 2
   \else\put(448,130)  {\line(5,-2) {120}} % sec. subst.
       \put(573,44)    {#8}  \fi       \fi %  on 2
 \ifx#9Q\else
 \ifx#9D\put(368,-1)   {\line(2,5) {46}}   % double 3 to 2
   \else\put(382,-35)  {\line(4,3) {120}}  % second subst.
       \put(507,28)    {#9}    \fi     \fi %  on 3
\end{picture}    }   % end terpene macro


\newcommand{\steroid}[9]        {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets the steroid skeleton. Optional double  %
% bonds and substituents make it possible to print the       %
% structures of common steroids.                             %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
   \begin{picture}(\pw,\pht)(-\xi,-\yi)
     \multiput(0,0)(342,0){3}    {\line(0,1) {200}} %3-2,
                                                    % 5-10,7-8
     \multiput(513,303)(342,0){3}{\line(0,1) {200}} %9-11,14-13,
                                                    %  15-16
     \multiput(0,200)(342,0){3}  {\line(5,3) {171}} %2-1,
                                                    % 10-9,
                                                    % 8-14
     \multiput(0,0)(342,0){2}    {\line(5,-3){171}} %3-4, 5-6
     \multiput(171,-103)(342,0){2{\line(5,3) {171}} %4-5, 6-7
     \multiput(171,303)(342,0){2}{\line(5,-3){171}} %1-10, 9-8
     \multiput(513,503)(342,0){2}{\line(5,3) {171}} %11-12,
                                                    % 13-17
     \multiput(684,606)(342,0){2}{\line(5,-3){171}} %12-13,
                                                    % 17-16
     \put(855,303)               {\line(1,0) {342}} %14-15
     \put(855,503)               {\line(0,1) {128}} % methyl 18
```

```
\put(795,638)                {\small ${\rm CH_3}$}
\ifx#1D\put(36,191)          {\line(5,3)  {126}}% double
 \else\ifx#1Q                                    %  1 to 2
       \else\put(520,514)    {\line(-5,3)  {128}}
            \put(506,492)    {\line(-5,3)  {128}}
            \put(83,547)     {\makebox(300,87)[r]{#1}} \fi
\fi                             % outside double & subst. 11
\ifx#2D\put(162,-67)         {\line(-5,3) {126}} % double
 \else\ifx#2Q                                     %  4 to 3
       \else\put(-7,11)      {\line(-5,-3) {128}}
            \put(7,-11)      {\line(-5,-3) {128}}
            \put(-430,-116) {\makebox(300,87)[r]{#2}} \fi
\fi                             % outside double & subst. 3
\ifx#3Q
 \else\put(0,0)        {\line(-5,-3){128}}           % subst.
      \put(-430,-116) {\makebox(300,87)[r]{#3}} \fi %  on 3
\ifx#4D\put(178,-67)   {\line(5,3)  {126}}     \fi % double
                                                   %  4 to 5
\ifx#5D\put(378,9)     {\line(5,-3) {126}}         % double
                                                   %  6 to 5
 \else\ifx#5Q
        \else\multiput(1011,606)(30,0){2} {\line(0,1) {100}}
             \put(985,713)          {#5}          \fi
\fi                                % outside double & subst. 17
\ifx#6D\put(316,174)   {\line(0,-1) {148}}     \fi % double
                                                   % 5 to 10
\ifx#6M\put(342,200)   {\line(0,1)  {128}}         % methyl
                                                   %  19
       \put(282,335)   {\small ${\rm CH_3}$}     \fi
\ifx#7Q
 \else\put(1026,606)   {\line(0,1)  {100}}         % lower
                                                   %  subst.
      \put(995,713)    {#7}                    \fi %  part
                                                   %  on 17
\ifx#8Q
 \else\put(1026,791)   {\line(0,1)  {100}}         % upper
                                                   %  subst.
      \put(995,900)    {#8}                    \fi %  part
                                                   %  on 17
\ifx#9Q
 \else\put(1026,606)   {\line(1,0)  {128}}         % the other
      \put(1160,575)   {#9}                    \fi %  subst.
```

```
                                                        %  on 17
   \end{picture}       }                          % end steroid macro


\newcommand{\hetthree}[8]       {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The hetthree macro draws a three-membered ring with one     %
% hetero atom.                                                %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(170,-170)  {#8}                          % hetero atom
  \put(0,0)       {C}                           % C-2
  \put(360,0)     {C}                           % C-3
  \put(80,30)     {\line(1,0)  {270}}           % bond 2-3
  \put(70,-10)    {\line(1,-1) {100}}           % bond 2-1
  \put(350,-10)   {\line(-1,-1){100}}           % bond 3-1
  \ifx#1Q                                       % substituent
    \else\put(210,-180) {\line(0,-1){80}}       %  on het. atom
         \put(180,-340) {#1}                \fi
  \ifx#6S\put(-10,30)   {\line(-1,0){140}}      % subst. on C-2
         \put(-460,-10) {\makebox(300,87)[r]{#2}} % with hor.
                                                % bond
    \else\ifx#6H\put(-90,30)   {\line(-1,0) {140}}% bond and H
                \put(-70,0)    {H}              %  on C-2
                \put(-540,-10) {\makebox(300,87)[r]{#2}}
         \else  \put(-310,-10) {\makebox(300,87)[r]{#2}} \fi \fi
                                                % no bond on 2
  \ifx#7S\put(440,30)   {\line(1,0) {140}}      % subst. on C-3
         \put(590,0)    {#3}                   %  with hor. bond
    \else\ifx#7H\put(430,0)    {H}              %  bond with H
                \put(510,30)   {\line(1,0)  {140}}
                \put(660,0)    {#3}
         \else  \put(445,0)    {#3}   \fi \fi % no bond on 3
  \ifx#4Q
    \else\put(40,80)    {\line(0,1) {140}}      % second subst.
         \put(-225,220) {\makebox(300,87)[r]{#4}} \fi % on C-2
  \ifx#5Q
    \else\put(400,80)   {\line(0,1) {140}}      % second subst.
         \put(360,230)  {#5}              \fi %  on C-3
 \end{picture}   }      % end hetthree macro
```

```
\newcommand{\hetifive}[9]        {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The hetifive macro typesets a five-membered ring with one     %
% hetero atom.                                                   %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(342,0)         {\line(-5,-3) {140}}         % bond 2,1
  \put(342,0)         {\line(0,1)   {200}}         % bond 2,3
  \put(342,200)       {\line(-1,0)  {342}}         % bond 3,4
  \put(0,200)         {\line(0,-1)  {200}}         % bond 4,5
  \put(0,0)           {\line(5,-3)  {140}}         % bond 5,1
  \ifx#1Q
   \else\put(171,-137) {\line(0,-1)  {83}}         % subst.on
        \put(135,-283) {#1}                    \fi % het.atom
  \ifx#2Q
   \else\ifx#2O\put(349,11) {\line(5,-3) {128}}    % outside
              \put(335,-11){\line(5,-3) {128}}    %  double O
              \put(475,-120)            {O}       %  on C-2
        \else\put(342,0)     {\line(5,-3) {128}}  % single
             \put(475,-100) {#2}               \fi %  subst.
  \fi                                              %  on C-2
  \ifx#3Q
   \else\put(342,200)  {\line(5,3)   {128}}        % subst. on
        \put(475,250)  {#3}                    \fi %  C-3
  \ifx#4Q
   \else\put(0,200)    {\line(-5,3)  {128}}        % subst. on
        \put(-430,234) {\makebox(300,87)[r]{#4}}\fi %  C-4
  \ifx#5Q
   \else\ifx#5O\put(-7,11)  {\line(-5,-3) {128}}   % outside
              \put(7,-11)  {\line(-5,-3) {128}}   %  double O
              \put(-200,-130)            {O}      %  on C-5
        \else\put(0,0)      {\line(-5,-3) {128}}  % single
             \put(-430,-116){\makebox(300,87)[r]{#5}} \fi % sub.
  \fi                                              %  on C-5
  \ifx#6D\put(316,26)       {\line(0,1)  {148}} \fi % double 2,3
  \ifx#7D\put(316,174)      {\line(-1,0) {290}} \fi % double 3,4
  \ifx#8D\put(26,174)       {\line(0,-1) {148}} \fi % double 4,5
  \put(135,-130)            {#9}                   % hetero atom
 \end{picture}              }     % end one-hetero fivering macro


\newcommand{\heticifive}[9]      {
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The heticifive macro typesets a five-membered ring with      %
% two hetero atoms separated by a carbon atom.                 %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(342,0)          {\line(-5,-3) {140}}          % bond 2,1
  \put(342,0)          {\line(0,1)   {160}}          % bond 2,3
  \put(0,200)          {\line(1,0)   {300}}          % bond 4,3
  \put(0,200)          {\line(0,-1)  {200}}          % bond 4,5
  \put(0,0)            {\line(5,-3)  {140}}          % bond 5,1
  \ifx#1Q
   \else\put(171,-137) {\line(0,-1)  {83}}           % subst. on
        \put(135,-283) {#1}                      \fi % het-1
  \ifx#2Q
   \else\ifx#20\put(349,11)   {\line(5,-3) {128}}    % outside
               \put(335,-11)  {\line(5,-3) {128}}    % double 0
               \put(475,-120) {0}                    % on C-2
        \else\put(342,0)      {\line(5,-3) {128}}    % single sub.
             \put(475,-100)   {#2}               \fi % on C-2
  \fi
  \ifx#3Q
   \else\put(370,217)  {\line(5,3)    {100}}         % subst. on
        \put(475,250)  {#3}                      \fi % on het-3
  \ifx#4Q
   \else\ifx#40\put(-7,189)   {\line(-5,3) {128}}    % outside
               \put(7,211)    {\line(-5,3) {128}}    % double 0
               \put(-200,250) {0}                    % on C-4
         \else\put(0,200)     {\line(-5,3) {128}}    % single sub.
              \put(-430,234)  {\makebox(300,87)[r]{#4}} \fi
  \fi                                                % on C-4
  \ifx#5Q
   \else\put(0,0)       {\line(-5,-3)  {128}}        % 1. single
        \put(-430,-116){\makebox(300,87)[r]{#5}}\fi % subst.
                                                     % on C-5
  \ifx#6D\put(316,26)  {\line(0,1)    {130}}    \fi % double 2,3
  \ifx#7Q
   \else\ifx#7D\put(26,174)   {\line(0,-1) {148}}    % double 4,5
        \else\put(0,0)        {\line(-5,3) {128}}    % 2. subst.
             \put(-430,34)    {\makebox(300,87)[r]{#7}}  \fi
  \fi                                                % on C-5
  \put(135,-130)             {#8}                    % het.atom 1
```

```
  \put(310,170)                {#9}                    % het.atom 3
 \end{picture}                 }     % end heticifive macro


\newcommand{\pyrazole}[8]      {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets the pyrazole ring with optional                   %
% substituents and double bonds inside and outside the                  %
% ring.                                                                 %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(200,-84)          {\line(5,3)    {110}}          % bond 1,2
  \put(342,200)          {\line(0,-1)   {140}}          % bond 3,2
  \put(342,200)          {\line(-1,0)   {342}}          % bond 3,4
  \put(0,200)            {\line(0,-1)   {200}}          % bond 4,5
  \put(0,0)              {\line(5,-3)   {140}}          % bond 5,1
  \put(135,-130)         {N}                            % ring N-1
  \put(310,-30)          {N}                            % ring N-2
  \ifx#1Q
   \else\put(171,-137) {\line(0,-1)   {83}}            % subst. on
        \put(150,-283) {#1}                       \fi % on N-1
  \ifx#2Q
   \else\put(370,-17)  {\line(5,-3)   {100}}           % subst. on
        \put(475,-100) {#2}                       \fi % N-2
  \ifx#3Q
   \else\ifx#30\put(335,211)  {\line(5,3)  {128}}   % outside
               \put(349,189)  {\line(5,3)  {128}}   % double 0
               \put(475,250)  {0}                   % on C-3
        \else\put(342,200)    {\line(5,3)  {128}}   % single sub.
             \put(475,250)    {#3}              \fi % on C-3
  \fi
  \ifx#4Q
   \else\put(0,200)     {\line(-5,3)   {128}}          % single sub.
        \put(-430,234) {\makebox(300,87)[r]{#4}}\fi % on C-4
  \ifx#5Q
   \else\ifx#50\put(-7,11)     {\line(-5,-3){128}}   % outside
               \put(7,-11)     {\line(-5,-3){128}}   % double 0
               \put(-200,-130){0}                    % on C-5
        \else\put(0,0)         {\line(-5,-3){128}}   % single sub.
             \put(-430,-116)   {\makebox(300,87)[r]{#5}}  \fi
  \fi                                                 % on C-5
  \ifx#6D\put(316,174) {\line(0,-1)   {114}}    \fi % double 3,2
```

```
  \ifx#7D\put(316,174) {\line(-1,0)    {290}}    \fi % double 3,4
  \ifx#8D\put(26,174)  {\line(0,-1)    {148}}    \fi % double 4,5
\end{picture}          }          % end pyrazole macro


\newcommand{\hetisix}[9]       {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The hetisix macro typesets a six-membered ring with        %
% one hetero atom.                                           %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
  \begin{picture}(\pw,\pht)(-\xi,-\yi)
    \put(342,0)     {\line(-5,-3)  {140}}    % bond 2 to 1
    \put(0,0)       {\line(5,-3)   {140}}    % bond 6 to 1
    \put(342,0)     {\line(0,1)    {200}}    % bond 2 to 3
    \put(342,200)   {\line(-5,3)   {171}}    % bond 3 to 4
    \put(171,303)   {\line(-5,-3)  {171}}    % bond 4 to 5
    \put(0,200)     {\line(0,-1)   {200}}    % bond 5 to 6
    \ifx#7D
      \put(316,26)  {\line(0,1)    {148}}    % double 2 to 3
      \else\ifx#7Q
           \else\put(349,11)  {\line(5,-3){128}} % outside
                \put(335,-11) {\line(5,-3){128}} %  double sub.
                \put(475,-120){#7}         \fi   %  on C-2
    \fi
    \ifx#8D
      \put(36,191)  {\line(5,3)    {126}} \fi  % double 5 to 4
    \ifx#1D
      \put(36,9)     {\line(5,-3)  {110}}      % double 6 to 1
      \else\ifx#1Q
           \else\put(171,-220) {\line(0,1) {83}}
                \put(150,-283) {#1}       \fi  % subst. on het.
    \fi
    \put(135,-130)  {#9}                       % the het.atom
    \ifx#2Q
      \else\put(342,0)     {\line(5,-3)  {128}} % subst. on 2
           \put(475,-100)   {#2}
    \fi
    \ifx#3Q
      \else\put(342,200)  {\line(5,3)   {128}} % subst. on 3
           \put(475,250)  {#3}
    \fi
    \ifx#4Q
```

```
      \else\put(171,303)  {\line(0,1)   {100}} % subst. on 4
           \put(150,410)  {#4}
    \fi
    \ifx#5Q
      \else\put(0,200)    {\line(-5,3)  {128}} % subst. on 5
           \put(-430,234) {\makebox(300,87)[r]{#5}}
    \fi
    \ifx#6Q
      \else\put(0,0)      {\line(-5,-3) {128}} % subst. on 6
           \put(-430,-116){\makebox(300,87)[r]{#6}}
    \fi                                 % end sixring macro with
  \end{picture}     }                   %  one hetero atom


\newcommand{\pyrimidine}[9]         {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets the pyrimidine ring with optional                        %
% substituents and double bonds inside and outside the                         %
% ring.                                                                        %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
  \begin{picture}(\pw,\pht)(-\xi,-\yi)
    \put(342,0)     {\line(-5,-3)  {140}}   % from 2 to 1
    \put(0,0)       {\line(5,-3)   {140}}   % from 6 to 1
    \put(342,0)     {\line(0,1)    {160}}   % from 2 to 3
    \put(171,303)   {\line(5,-3)   {140}}   % from 4 to 3
    \put(171,303)   {\line(-5,-3)  {171}}   % from 4 to 5
    \put(0,200)     {\line(0,-1)   {200}}   % from 5 to 6
    \put(135,-130)  {N}
    \put(310,170)   {N}
    \ifx#1Q
     \else\put(171,-137)  {\line(0,-1)  {83}} % sub. on N-1
          \put(150,-283)  {#1}               \fi
    \ifx#2Q
     \else\ifx#20\put(349,11)  {\line(5,-3) {128}}  % outside
                 \put(335,-11) {\line(5,-3) {128}}  %  double O
                 \put(475,-120){O}                  %  on C-2
          \else\put(342,0)     {\line(5,-3) {128}}  % single
               \put(475,-100)  {#2}           \fi   %  subst.
    \fi
    \ifx#3Q
     \else\put(370,217)   {\line(5,3)   {100}}      % subst.
          \put(475,250)   {#3}                \fi   %  on N-3
```

```
     \ifx#4Q
      \else\ifx#4O\put(158,303) {\line(0,1)  {100}}  % outside
                 \put(184,303) {\line(0,1)  {100}}  %  double O
                 \put(130,410) {O}                  %  on C-4
           \else\put(171,303)   {\line(0,1)  {100}}  % single
               \put(150,410)     {#4}          \fi  %  subst.
     \fi                                            %  on C-4
     \ifx#5Q
      \else\put(0,200)     {\line(-5,3)   {128}}    % 1. subst.
           \put(-430,234)  {\makebox(300,87)[r]{#5}} \fi
                                                    %  on C-5
     \ifx#6Q
      \else\ifx#6O\put(-7,11)    {\line(-5,-3){128}}  % outside
                 \put(7,-11)    {\line(-5,-3){128}}  %  double O
                 \put(-210,-130){O}                  %  on C-6
           \else\put(0,0)       {\line(-5,-3){128}}  % single s.
               \put(-430,-116) {\makebox(300,87)[r]{#6}} \fi
     \fi                                            %  on C-6
     \ifx#7D\put(306,9)     {\line(-5,-3)  {120}} \fi % 2,1 doub.
     \ifx#8D\put(180,267)  {\line(5,-3)    {120}}     % 4,3 doub.
      \else\ifx#8Q
           \else\put(0,200){\line(-5,-3)  {128}}      % 2. subst.
               \put(-430,84) {\makebox(300,87)[r]{#8}}   \fi
     \fi                                            %  on C-5
     \ifx#9D\put(26,174)    {\line(0,-1)   {148}} \fi % 5,6 doub.
  \end{picture}  }                           % end pyrimidine macro


\newcommand{\pyranose}[9]    {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets monosaccharides with a pyran ring      %
% system.                                                    %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
  \begin{picture}(\pw,\pht)(-\xi,-\yi)
   \put(588,0)          {\line(-1,-1) {159}}         % 1,2 bond
   \put(429,-159)       {\line(-1,0)  {270}}         % 2,3 bond
   \put(159,-159)       {\line(-1,1)  {159}}         % 3,4 bond
   \put(0,0)            {\line(1,1)   {159}}         % 4,5 bond
   \put(159,159)        {\line(1,0)   {225}}         % C-5 to O
   \put(394,130)        {\small O}                   % ring O
   \put(460,130)        {\line(1,-1)  {128}}         % O to C-1
   \ifx#1Q
```

```
    \else\put(588,0)    {\line(1,1)    {100}}             % beta sub.
         \put(700,75)   {#1}                         \fi  %  on C-1
   \ifx#2Q
    \else\put(588,0)    {\line(1,-1)   {100}}             % alpha sub.
         \put(700,-120){#2}                          \fi  %  on C-1
   \ifx#3Q
    \else\put(429,-159){\line(0,1)    {85}}              % up subst.
         \put(225,-75)  {\makebox(250,87)[r]{#3}} \fi    %  on C-2
   \ifx#4Q
    \else\put(429,-159){\line(0,-1)   {85}}              % down sub.
         \put(400,-315){#4}                          \fi %  on C-2
   \ifx#5Q
    \else\put(159,-159){\line(0,1)    {85}}              % up subst.
         \put(130,-73)  {#5}                         \fi  %  on C-3
   \ifx#6Q
    \else\put(159,-159){\line(0,-1)   {85}}              % down sub.
         \put(130,-315){#6}                          \fi  %  on C-3
   \ifx#7Q
    \else\put(0,0)      {\line(0,1)    {85}}              % up subst.
         \put(-270,84)  {\makebox(300,87)[r]{#7}} \fi    %  on C-4
   \ifx#8Q
    \else\put(0,0)      {\line(0,-1)   {85}}              % down sub.
         \put(-270,-160){\makebox(300,87)[r]{#8}}\fi     %  on C-4
   \put(159,159)        {\line(0,1)    {85}}              % C-6
   \put(130,250)        {\small ${\rm CH_{2}}$}           % sub. on
   \put(-370,245)       {\makebox(500,87)[r]{#9}}         %  C-6
 \end{picture}          }                        % end pyranose macro


\newcommand{\furanose}[8]    {
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 % This macro typesets monosaccharides with a furan ring      %
 % system.                                                    %
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 %
  \begin{picture}(\pw,\pht)(-\xi,-\yi)
   \put(448,0)            {\line(-1,-2) {89}}        % bond 1,2
   \put(359,-179)         {\line(-1,0)  {270}}       % bond 2,3
   \put(89,-179)          {\line(-1,2)  {89}}        % bond 3,4
   \put(0,0)              {\line(5,3)   {188}}       % C-4 to O
   \put(192,110)          {\small O}                 % ring O
   \put(260,113)          {\line(5,-3)  {188}}       % O to C-1
   \ifx#1Q
```

```
   \else\ifx#1N\put(448,0) {\line(0,1)   {380}}     % long bond
                                                    %  for nucl.
        \else\put(448,0)   {\line(5,3)   {105}}     % beta sub.
             \put(558,50)  {#1}                \fi %  on C-1
  \fi
  \ifx#2Q
   \else\put(448,0)        {\line(5,-3)  {105}}     % alpha sub.
        \put(558,-90)      {#2}                \fi  %   on C-1
  \ifx#3Q
   \else\put(359,-179)     {\line(0,1)   {85}}      % up subst.
        \put(155,-95)      {\makebox(250,87)[r]{#3}}\fi %on C-2
  \ifx#4Q
   \else\put(359,-179)     {\line(0,-1)  {85}}      % down sub.
        \put(330,-335)     {#4}                \fi  %  on C-2
  \ifx#5Q
   \else\put(89,-179)      {\line(0,1)   {85}}      % up sub.
        \put(60,-93)       {#5}                \fi  %  on C-3
  \ifx#6Q
   \else\put(89,-179)      {\line(0,-1)  {85}}      % down sub.
        \put(60,-335)      {#6}                \fi  %   on C-3
  \ifx#7Q
   \else\put(0,0)          {\line(0,-1)  {85}}      % down sub.
        \put(-270,-160)    {\makebox(300,87)[r]{#7}}\fi %on C-4
  \put(0,0)                {\line(0,1)   {85}}      % C-5
  \put(-30,90)             {\small ${\rm CH_{2}}$}
  \put(-530,80)            {\makebox(500,87)[r]{#8}}
 \end{picture}        }          % end furanose macro


\newcommand{\purine}[9]          {
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 % This macro typesets the purine ring system with optional   %
 % double bonds and substituents.                             %
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 %
   \begin{picture}(\pw,\pht)(-\xi,-\yi)
     \put(0,0)       {\line(0,1)   {160}}          % bond 2 to 1
     \put(0,0)       {\line(5,-3)  {140}}          % from 2 to 3
     \put(342,0)     {\line(-5,-3) {140}}          % from 4 to 3
     \put(342,0)     {\line(0,1)   {200}}          % from 4 to 5
     \put(322,20)    {\line(0,1)   {160}}          % double 4 to 5
     \put(342,200)   {\line(-5,3)  {171}}          % from 5 to 6
     \put(171,303)   {\line(-5,-3) {140}}          % from 6 to 1
```

```
\put(342,200)  {\line(1,0)   {300}}          % from 5 to 7
\put(684,0)    {\line(0,1)   {160}}          % from 8 to 7
\put(684,0)    {\line(-5,-3) {140}}          % from 8 to 9
\put(342,0)    {\line(5,-3)  {140}}          % from 4 to 9
\put(-32,170)  {\small N}                    % N at 1
\put(135,-130) {\small N}                    % N at 3
\put(652,170)  {\small N}                    % N at 7
\put(477,-130) {\small N}                    % N at 9
\ifx#1Q
  \else\put(-128,277) {\line(5,-3) {100}}    % subst. on 1
       \put(-430,234) {\makebox(300,87)[r]{#1}} \fi
\ifx#2D\put(27,7)     {\line(5,-3) {115}}    % double 2 to 3
  \else\multiput(0,-15)(0,30){2} {\line(-5,-3) {128}}
       \put(-430,-116){\makebox(300,87)[r]{#2}}
\fi                                          %outside double
                                             %  bond on 2
\ifx#3Q                                      % subst. on 3
  \else\put(171,-130) {\line(0,-1) {73}}
       \put(135,-283) {#3}
\fi
\ifx#4D\put(170,281)  {\line(-5,-3){115}} \fi % d. bond 6,1
\ifx#5Q
  \else\multiput(156,303)(30,0){2} {\line(0,1) {100}}
       \put(135,410)  {#5}                   % outside d.-bond
\fi                                          % and subst. on 6
\ifx#6Q                                      % single-bonded
  \else\put(171,303)  {\line(0,1)  {100}} % subst. on 6
       \put(150,410)  {#6}         \fi
\ifx#7Q                                      % subst. on 7
  \else\put(812,277)  {\line(-5,-3){100}}
       \put(817,250)  {#7}
\fi
\ifx#8D\put(664,20)   {\line(0,1)  {130}} % d.-bond 8 to 7
  \else\multiput(684,-15)(0,30){2} {\line(5,-3) {128}}
       \put(817,-100) {#8}                   % outside d.-bond
\fi                                          % and subst. on 8
\ifx#9Q                                      % subst. on 9
  \else\put(513,-130) {\line(0,-1) {73}}
       \put(492,-283) {#9}
\fi
\end{picture}      }            % end purine macro
```

```
\newcommand{\fuseiv}[9]       {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a four-carbon fragment designed to     %
% be connected to another ring at two places. As the         %
% result, a sixring is fused linearly to another ring        %
% system.                                                    %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(0,200)      {\line(5,3)  {171}}                    % NE bond
  \put(171,303)    {\line(5,-3) {171}}                    % SE bond
  \put(342,200)    {\line(0,-1) {200}}                    % S  bond
  \put(342,0)      {\line(-5,-3){171}}                    % SW bond
  \put(171,-103)   {\line(-5,3) {171}}                    % NW bond
  \ifx#1Q
   \else\put(171,303)    {\line(0,1)    {100}}            % subst.
        \put(150,410)    {#1}                      \fi %  on top
  \ifx#2Q
   \else\put(342,200)    {\line(5,3)    {128}}            % subst.
        \put(475,250)    {#2}                      \fi %  top rt.
  \ifx#3Q
   \else\put(342,0)      {\line(5,-3)   {128}}            % subst.
        \put(475,-100)   {#3}                      \fi %  low rt.
  \ifx#4Q
   \else\put(171,-103)   {\line(0,-1)   {100}}            % bottom
        \put(150,-283)   {#4}                      \fi %  subst.
  \ifx#5D \put(36,191)   {\line(5,3)    {126}}     \fi % NE double
  \ifx#6D \put(180,267)  {\line(5,-3)   {126}}         % SE double
   \else\ifx#6Q
         \else\put(342,200)   {\line(5,-3)   {128}}    % 2. subst.
              \put(475,100)   {#6}                 \fi %  top rt.
  \fi
  \ifx#7D \put(316,174)  {\line(0,-1)   {148}}     \fi % S  double
  \ifx#8D \put(306,9)    {\line(-5,-3)  {126}}         % SW double
   \else\ifx#8Q
         \else\put(342,0)     {\line(5,3)    {128}}    % 2. subst.
              \put(475,50)    {#8}                 \fi % lower rt.
  \fi
  \ifx#9D \put(162,-67)  {\line(-5,3)        {126}}\fi % NW double
 \end{picture}        }                  % end fuseiv macro


\newcommand{\fuseup}[9]    {
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a four-carbon fragment designed to     %
% connect to another ring at two places. As the result,      %
% a sixring is fused angularly to the original ring.         %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(-171,303)        {\line(0,1)   {200}}          % N  bond
  \put(-171,503)        {\line(5,3)   {171}}          % NE bond
  \put(0,606)           {\line(5,-3)  {171}}          % SE bond
  \put(171,503)         {\line(0,-1)  {200}}          % S  bond
  \put(171,303)         {\line(-5,-3) {171}}          % SW bond
  \ifx#1Q
   \else\put(-171,503)  {\line(-5,3)  {128}}       % upper left
        \put(-600,537)  {\makebox(300,87)[r]{#1}}\fi %  subst.
  \ifx#2Q
   \else\put(0,606)     {\line(0,1)   {100}}          % top sub.
        \put(-19,713)   {#2}                    \fi
  \ifx#3Q
   \else\put(171,503)   {\line(5,3)   {128}}          % top rt.
        \put(304,553)   {#3}                    \fi %  subst.
  \ifx#4Q
   \else\put(171,303)   {\line(5,-3)  {128}}          % lower rt.
        \put(304,203)   {#4}                    \fi %  subst.
  \ifx#5D\put(-145,329) {\line(0,1)   {148}}    \fi % N double
  \ifx#6D\put(-135,494) {\line(5,3)   {126}}    \fi % NE double
  \ifx#7D\put(9,570)    {\line(5,-3)  {126}}    \fi % SE double
  \ifx#8D\put(145,477)  {\line(0,-1)  {148}}    \fi % S  double
  \ifx#9D\put(135,312)  {\line(-5,-3) {126}}    \fi % SW double
 \end{picture}       }                 % end fuseup macro


\newcommand{\fuseiii}[6]     {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a three-carbon fragment designed to    %
% be connected to another ring at two places. As the         %
% result, a fivering is fused linearly to the original       %
% ring system.                                               %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \begin{picture}(\pw,\pht)(-\xi,-\yi)
  \put(0,200)       {\line(1,0)   {342}}              % E  bond
  \put(342,200)     {\line(0,-1)  {200}}              % S  bond
```

```
  \put(342,0)      {\line(-5,-3) {171}}                % SW bond
  \put(171,-103)   {\line(-5,3)  {171}}                % NW bond
  \ifx#1Q
   \else\put(342,200)    {\line(5,3)    {128}}         % upper rt.
        \put(475,250)    {#1}                       \fi %  subst.
  \ifx#2Q
   \else\put(342,0)      {\line(5,-3)   {128}}         % lower rt.
        \put(475,-100)   {#2}                       \fi %  subst.
  \ifx#3Q
   \else\put(171,-103)   {\line(0,-1)   {100}}         % bottom
        \put(150,-283)   {#3}                       \fi %  subst.
  \ifx#4Q
   \else\put(342,200)    {\line(5,-3)   {128}}         % 2. upper
        \put(475,100)    {#4}                       \fi %  rt. sub.
  \ifx#5Q
   \else\put(342,0)      {\line(5,3)    {128}}         % 2. lower
        \put(475,50)     {#5}                       \fi %  rt. sub.
  \ifx#6D \put(316,174) {\line(0,-1)   {148}}     \fi % S double
\end{picture}      }                  % end fuseiii macro


\newcommand{\cto}[3]   {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets an arrow for a chemical equation and   %
% puts text above and below it.                              %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
 \len=50   \multiply \len by #3                  % calc. arrow
 \advance  \len by 120                           % length
   \begin{picture}(\pw,\pht)(-\xi,-\yi)
    \put(60,50)      {\vector(1,0)   {\len}}    % draw arrow
    \put(90,70)      {\makebox(\len,70)[l]
                      {\scriptsize ${\rm #1}$}}  % text on top
    \put(90,-40)     {\makebox(\len,70)[l]
                      {\scriptsize ${\rm #2}$}}  % text below
  \end{picture}      }


\newcommand{\sbond}[1]    {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a horizontal single bond of user-      %
% specified length.                                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%
\xbox=#1  \advance \xbox by 2
\hspace{1.5pt} \parbox{\xbox pt} {\rule{#1 pt}{0.4pt} }
\xbox=50     }  % end sbond macro


\newcommand{\dbond}[2]    {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a horizontal double bond of user-      %
% specified length.                                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
\xbox=#1  \advance \xbox by 2
\hspace{1.5pt}\parbox{\xbox pt} {\rule{#1 pt}{0.4pt}
                                 \vspace{-#2 pt}\\
                                 \rule{#1 pt}{0.4pt}  }
\xbox=50     }  % end dbond macro


\newcommand{\tbond}[2]   {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro typesets a horizontal triple bond of user-      %
% specified length.                                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
\xbox=#1   \advance  \xbox by 2
\hspace{1.5pt} \parbox{\xbox pt} {\rule{#1 pt}{0.4pt}
                                  \vspace{-#2 pt}\\
                                 \rule{#1 pt}{0.4pt}
                                  \vspace{-#2 pt}\\
                                 \rule{#1 pt}{0.4pt}   }
\xbox=50     }   % end tbond macro


  \begin{picture}(900,900)(-300,-300)
   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
   % This is not a macro. It is the LaTeX code that generated   %
   % figure 3.5 on p. 20. It is included here to illustrate     %
   % the coding of the isoprenoid zig-zag bond lines.           %
   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
   %
   \put(342,200)    {\line(0,-1)  {200}}          % The ring:
   \put(342,0)      {\line(-5,-3) {171}}
```

```
  \put(171,-103)  {\line(-5,3)  {171}}
  \put(0,0)       {\line(0,1)   {200}}
  \put(0,200)     {\line(5,3)   {171}}
  \put(171,303)   {\line(5,-3)  {171}}
  \put(322,180)   {\line(0,-1)  {160}}
  \put(342,0)     {\line(5,-3)  {128}}
  \put(171,303)   {\line(5,3)   {128}}
  \put(171,303)   {\line(-5,3)  {128}}
  \multiput(342,200)(342,0){5}{\line(5,3){171}}  % zig-zag
  \multiput(513,303)(342,0){4}{\line(5,-3){171}} %  lines:
  \multiput(527,270)(342,0){4}{\line(5,-3){135}}
  \multiput(855,303)(684,0){2}{\line(0,1){160}}
  \put(1881,275){=0}
\end{picture}


\newcommand{\branch}[6]    {
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This macro is not listed in chapter VI. It is similar to   %
% cbranch and tbranch. It will produce good-looking          %
% diagrams in single-spaced documents only.                  %
% The macro listing is included here because it uses plain   %
% TeX only, no LaTeX features.                               %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
  \vbox{\halign
    {$##$&$##$\cr
    \ifx#1Q          \ &\ \cr   % upper substituent
      \else          \ &#1\hidewidth\cr\noalign{\vskip-2pt}
    \fi
    \ifx#2S          \ &\thinspace\vert\cr
      \else\ifx#2D   \ &\hspace{1pt}\vert\hspace{.3pt}\vert\cr
              \else  \ &\ \cr
           \fi       \fi
    #3&#4\cr\noalign{\vskip-2pt} % left and right substituents
    \ifx#5S          \ &\thinspace\vert\cr
       \else\ifx#5D  \ &\hspace{1pt}\vert\hspace{.3pt}\vert\cr
               \else \ &\ \cr
            \fi      \fi
    \ifx#6Q          \ &\ \cr    % lower substituent
       \else         \ &#6\hidewidth\cr
    \fi      }       } }   % end branch macro
```

APPENDIX B

COORDINATES OF POINTS OF ATTACHMENT

The tables in this appendix list coordinates of points of attachment that will probably be used most frequently with the techniques described in chapter V. Table B.1 lists the coordinates of the six corners of the carbon sixring and the coordinates at the end of the six bonds extending from the sixring.

Table B.1: Points of attachment in the sixring

Sixring Position	Ring Corner	End of Bond
1	(342,200)	(470,277)
2	(342,0)	(470,-77)
3	(171,-103)	(171,-203)
4	(0,0)	(-128,-77)
5	(0,200)	(-128,277)
6	(171,303)	(171,403)

The ring structures typeset by the macros \fivering, \naphth, \steroid, \hetifive, \heticifive, \pyrazole, \hetisix, \pyrimidine, and \purine have identical coordinates at positions that are equivalent to the sixring positions in regard to the printed diagram. (The position numbers are not necessarily the same as those of the sixring.)

Table B.2 lists the coordinates of other points of attachment in structures typeset by various macros.

Table B.2: Points of attachment in various structures

Macro	Position Description	Coordinates
\cbranch	begin of left bond	(-150,33)
\cbranch	end of right bond	(230,33)
\cright	begin of left bond	(-150,33)
\cleft	end of right bond	(230,33)
\chemup	end of vertical bond	(33,-150)
\cdown	end of vertical bond	(33,220)
\threering	ring position 1	(300,0)
\threering	end of bond on 1	(428,77)
\naphth	ring position 1	(513,303)
\naphth	end of bond on 1	(513,403)
\naphth	ring position 2	(684,200)
\naphth	end of bond on 2	(812,277)
\steroid	end of vert. bond on 17	(1026,706)
\hetthree	end of bond on 3	(580,30)
\pyranose	end of alpha bond	(688,100)
\pyranose	end of beta bond	(688,-100)
\furanose	end of alpha bond	(553,-63)
\furanose	end of beta bond	(553,63)
\furanose	end of vert. long bond	(448,380)
\purine	point below N(9)	(513,-130)
\fuseiv	upper point of attachment	(0,200)
\fuseiv	lower point of attachment	(0,0)
\fuseup	upper point of attachment	(-171,303)
\fuseup	lower point of attachment	(0,200)
\fuseiii	upper point of attachment	(0,200)
\fuseiii	lower point of attachment	(0,0)

APPENDIX C

SLOPES OF LINES DRAWN WITH LATEX

Table C.1 lists the first quadrant slopes of lines that can be typeset with LaTeX, together with the corresponding degrees of angle. The integers x_s and y_s represent the slope in LaTeX's line-drawing statement

$$\backslash put(x,y)\ \{\backslash line(x_s,y_s)\ \{length\}\}.$$

Corresponding angles in the other quadrants can be generated by preceding x_s and/or y_s with a minus sign.

Table C.1: Slopes of lines possible with LaTeX

x_s, y_s	$\tan\theta(y_s/x_s)$	θ(degrees)
1,0	0.00	0.0
6,1	0.17	9.5
5,1	0.20	11.3
4,1	0.25	14.0
3,1	0.33	18.5
5,2	0.40	21.8
2,1	0.50	26.5
5,3	0.60	31.0
3,2	0.67	33.7
4,3	0.75	36.8
5,4	0.80	38.7
6,5	0.83	39.8
1,1	1.00	45.0
5,6	1.20	50.2
4,5	1.25	51.3
3,4	1.33	53.2
2,3	1.50	56.3
3,5	1.67	59.0
1,2	2.00	63.5
2,5	2.50	68.2
1,3	3.00	71.5
1,4	4.00	76.0
1,5	5.00	78.7
1,6	6.00	80.5
0,1	∞	90.0

VITA

Roswitha T. Haas was born in Bremen, Germany, the eldest daughter of Enno and Lore Huchting. After graduation from high school there she entered the University of Münster in Westphalia, where she completed the study of chemistry with the degree of Doctor of Natural Sciences in May 1960. After a one-year postdoctoral fellowship at the University of Münster she became a staff writer at the Beilstein Institute for Organic Chemistry in Frankfurt.

She left to join her husband in the United States in March 1963. Later that year she accepted another one-year postdoctoral appointment, at the University of Tennessee, Knoxville, where her husband had become a member of the History Department. From 1964 to the present Dr. Haas has been Assistant Professor and then Associate Professor of Chemistry at Knoxville College. Since 1977 she also has been a Research Associate at the Environmental Mutagen Information Center , Oak Ridge National Laboratories.

Granted a leave of absence by Knoxville College in the Fall of 1982, she enrolled in the Graduate School of the University of Tennessee. She received the Master of Science degree with a major in Computer Science in August 1986.

The author is a member of the American Chemical Society and the Association for Computing Machinery and has published articles concerned with organic chemistry and information analysis in science.