The Asian Journal of TEX, Volume 3, No. 2, December 2009
Article revision 2009/10/30

Articles, Books, and Internet Documents with Structural Formulas Drawn by XyMTEX — Writing, Submission, Publication, and Internet Communication in Chemistry

Shinsaku Fujita

Shonan Institute of Chemoinformatics and Mathematical Chemistry
fujita-sicimc@nifmail.jp http://xymtex.com

KEYWORDS structural formula, XyMTeX, XyM notation, XyMML, markup language, XyMJava, XML, HTML, TeX, LaTeX, PostScript, PDF, steroid

ABSTRACT Preparation methods of chemical documents containing chemical structural formulas have been surveyed by referring to the author's experiences of publishing books, emphasizing differences before and after the adoption of TEX/LATEX-typesetting as well as before and after the development of XyMTEX. The recognition of XyMTEX commands as linear notations has led to the concept of the XyM notation, which has further grown into XyMML (XyM Markup Language) as a markup language for characterizing chemical structural formulas. XML (Extensible Markup Language) documents with XyMML are converted into HTML (Hypertext Markup Language) documents with XyM notations, which are able to display chemical structural formulas in the Internet by means of the XyMJava system developed as a Java applet of an internet browser. On the other hand, the same XML documents with XyMML are converted into LATEX documents with XyMTEX commands (the same as XyM notations), which are able to print out chemical structural formulas of high quality. Functions added by the latest version (4.04) of XyMTEX have enhanced abilities of drawing complicated structures such as steroids. LATEX documents with XyMTEX formulas can be converted into PDF (Portable Document Format) documents directly or via PostScript document. Applications of such PDF documents in online or semi-online submission to scientific journals have been discussed.

1 Introduction

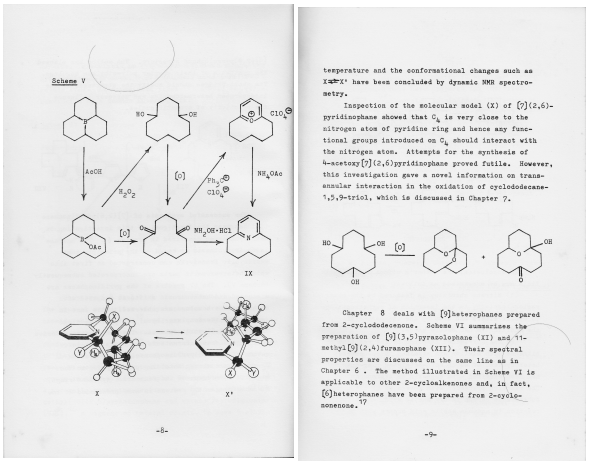
Because chemical structural formulas are essential to organic chemistry [1], they should be drawn by an easy and convenient method and then incorporated into chemical documents (articles, books, and other publications) in an efficient way. During the last 40 years, tools for drawing chemical structures have made remarkable progress, which was mainly provided by the advance of computer technology. Because these years overlapped with the author's career of R&D, his experience in catching up with

Copyright © 2009 by the Korean TEX Society

such remarkable progress of hardware and software would be helpful in realizing what happened during these years.

1.1 Before Release of XeTeX

The author started studies on synthetic organic chemistry in the 1970s, when chemical structural formulas were drawn manually by using a drawing pen (e.g., a rotring pen) and a chemical triangle (so-called Fieser chemist's triangle). In preparing manuscripts of articles and books, they were sticked on pages which were separately prepared by using a manual typewriter. Then, the resulting manuscripts served as camera-ready materials for printing. Figure 1 shows two pages selected from the author's doctoral thesis [3] as a typical example of such a book prepared manually.


Scheme V

temperature and the conformational changes such as X⇌X' have been concluded by dynamic NMR spectrometry.

Inspection of the molecular model (X) of [7](2,6)-pyridinophane showed that C_4 is very close to the nitrogen atom of pyridine ring and hence any functional groups introduced on C_4 should interact with the nitrogen atom. Attempts for the synthesis of 4-acetoxy[7](2,6)pyridinophane proved futile. However, this investigation gave a novel information on transannular interaction in the oxidation of cyclododecane-1,5,9-triol, which is discussed in Chapter 7.

Chapter 8 deals with [9]heterophanes prepared from 2-cyclododecenone. Scheme VI summarizes the preparation of [9](3,5)pyrazolophane (XI) and 11-methyl[9](2,6)furanophane (XII). Their spectral properties are discussed on the same line as in Chapter 6. The method illustrated in Scheme VI is applicable to other 2-cycloalkenones and, in fact, [6]heterophanes have been prepared from 2-cyclononenone.[7]

-8-

-9-

FIGURE 1. Manual typesetting with structural formulas drawn manually in the author's thesis [3].

In the 1980s, explosive advances in computer technology (both hardware and software) provided computer systems for drawing chemical structures, where ChemDraw appeared as a pioneering tool and was followed by ISIS/Draw, ChemWindow, ChemSketch, and so on [26]. Note that early versions of ChemDraw ran only on Macintosh personal computers and that there were no tools for drawing chemical structures of high quality which ran on MS-DOS personal computers at that time.

In the latter half of the 1980s, the author shifted his R&D to interdisciplinary fields between chemistry and mathematics (chemical combinatorial enumerations and mathematical stereochemistry), which required both mathematical equations and chemical structural formulas. In those days, there was only TEX/LATEX as software of typesetting mathematical equations of high quality. Hence, the selection of the TEX/LATEX was inevitable if mathematical aspects were considered. Because the TEX/LATEX program ran only under UNIX operating systems during the early days, it was not suitable to the use of MS-DOS personal computers. Fortunately, however, an MS-DOS version of TEX/LATEX was released from ASCII Co. in 1987. The release motivated the start of TEX/LATEX-typesetting in order to satisfy the mathematical as-

1989	Start of writing articles with Japanese MicroTEX (Ascii Co.) on PC-9800 (NEC).
1991	LATEX-typesetting and publication of "Symmetry and Combinatorial Enumeration in Chemistry" from Springer-Verlag (Heidelberg-Berlin).
1993	LATEX-typesetting and publication of "LATEX for Chemists and Biochemists" (in Japanese) from Tokyo Kagaku Dojin (Tokyo).
1993*	First release of XƆMTEX (version 1.00) for typesetting chemical structural formulas.
1997*	XƆMTEX/LATEX-typesetting and publication of "XƆMTEX—Typesetting Chemical Structural Formulas" from Addison-Wesley Japan (Tokyo).
1999*	Release of XƆMTEX Version 2.00 for supporting the XƆM notation as a linear notation of structural formulas.
2001	LATEX-typesetting and publication of "Computer-Oriented Representation of Organic Reactions" from Yoshioka-Shoten (Kyoto).
2002*	Release of XƆMTEX Version 4.00 for supporting PostScript language
2004	XƆMTEX/LATEX-typesetting and publication of "Organic Chemistry of Photography" from Springer-Verlag (Heidelberg-Berlin).
2007	XƆMTEX/LATEX-typesetting and publication of "Diagrammatical Approach to Molecular Symmetry and Enumeration of Stereoisomers" from University of Kragujevac (Kragujevac).
2009*	Release of XƆMTEX Version 4.04 for supporting complicated formulas such as steroids.

FIGURE 2. Short private history of XƆMTEX with LATEX. Each item with an asterisk is concerned with the development of XƆMTEX itself.

pects of the author's research, as shown in Figure 2.

In the first half of the 1990s, laser-beam printers became available to amateur use, so as to bring about high-quality printing of manuscripts prepared by TEX/LATEX. In 1991, the author published a monograph on chemical combinatorics [4], which was typeset by using TEX/LATEX loaded in an MS-DOS personal computer. The text of the monograph was printed out by a laser-beam printer to give a camera-ready manuscript. Figure 3 shows two pages selected from the monograph, which contain mathematical equations prepared by TEX/LATEX and structural formulas drawn manually. At that time, drawing tools for MS-DOS personal computers were not easily available, because ChemDraw ran only on Macintosh personal computers. Although ChemTEX [22] was reported as a drawing tool based on the LATEX picture environment, it was not adopted because of a limited number of accessible structures. As for the situation of electronic publishing and chemical text processing in those days, see [28].

1.2 Release of XƆMTEX

In 1993, the author published a book on TEX/LATEX-typesetting, where target readers were expected to be chemists and biochemists [5]. Chapter 10 of this book discussed various possibilities of incorporating structural formulas in the process of TEX/LATEX-typesetting. Although ChemTEX [22] was cited as a pioneering tool for drawing chemical structures, the use of ChemTEX revealed several drawbacks to its applicability. In particular, because ChemTEX spends an argument to assign one substituent, a large

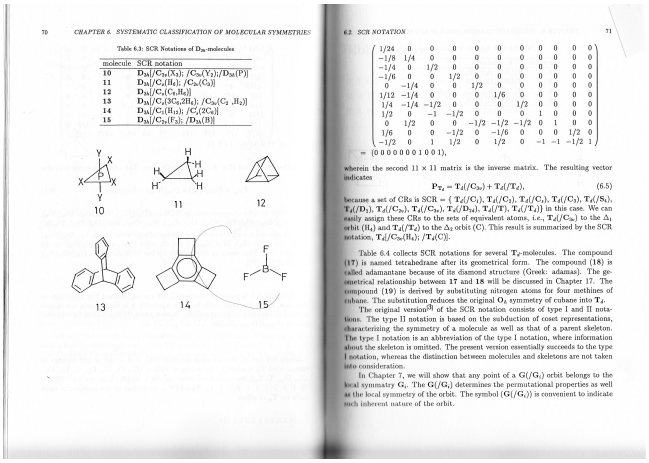


FIGURE 3. LaTeX-typesetting with structural formulas drawn manually [4].

```
\sixheterovi%                          %mother skeleton
[bd]%                                  %bondlist
{1==\downnobond{N}{H};3==N}%           %atomlist
{2D==O;4==NH$_{2}$;5==HOCH$_{2}$OH}%   %substlist
```

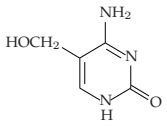


FIGURE 4. XyMTEX command of fundamental level for drawing 5-(hydroxymethyl)cytosine.

number of substituents in complicated structures cannot be assigned systematically. Note that each control sequence (command) of TEX/LATEX is able to take at most 9 arguments. Moreover, ChemTEX has no systematic procedures of assigning substituents, bond multiplicities, inner hetero atoms, etc. as well as of constructing fused rings, spiro rings, etc.

In the same 1993, the XyMTEX version 1.00 was released to give systematic procedures of assigning substituents (subslist), bond multiplicities (bondlist), inner hetero atoms (atomlist) etc. as well as of constructing fused rings, spiro rings, etc. [6, 7]. The manual of the XyMTEX system was published as a reference book in 1997 [8]. The XyMTEX system was cited in a textbook on LATEX published in 1997 [20].

For example, a structural formula of 5-(hydroxymethyl)cytosine is drawn by writing down the code shown in Figure 4, where the components of the structural formula are divided into a mother skeleton, a bondlist (bond list), an atomlist (atom list), and a subslist (substituent list). After processing by LATEX with XyMTEX, the code produces the corresponding structural formula shown in the bottom of Figure 4.

1.3 Aims of the Present Article

The first aim of the present article is to show that XƒMTEX commands can be regarded as a set of linear notations (the XƒM notation), which can be further extended into XƒMML (the XƒM markup language). The second aim is to show its applicability to internet communication, where XƒMJava is discussed as a tool of incorporating chemical structures into HTML documents. This way will be described under the term XƒM TEXnology, which is newly coined to stress potentialities of the XƒMTEX system and related applications. The third aim is to discuss practical aspects of the XƒMTEX system, where combination with PostScript or with PDF technology provides structural formulas of high quality, which are suitable for publication use.

After the release of the XƒMTEX system, there appeared several approaches for drawing chemical structural formulas in TEX/LATEX, e.g., PPCHTEX [23], OCHEM [25], and so on. Differences in methodologies between these approaches and the XƒMTEX system will become clear by virtue of the following discussions under the term XƒM TEXnology.

2 XƒM TEXnology for Internet Communication

Although LATEX has been discussed as a document repository for the Internet [21], practical methods have not been fully developed to cover all disciplines of science. The XƒMTEX system and related XƒM TEXnology will serve as tools for communicating chemical structural data in the Internet.

2.1 Total Features of Internet Communication by XƒM TEXnology

Although the XƒMTEX system was originally developed as a typesetting system based on the TEX/LATEX technology, it has been extended to support internet communication of chemical structures. Total features of the developed system for internet communication are shown in Figure 5. They are referred to here as "XƒM TEXnology", where XƒMTEX is also regarded as a logo of the "XƒM TEXnology".

2.2 XƒMTEX Commands as the XƒM Notation

By scrutinizing rules of such systematic nomenclatures as the IUPAC nomenclature, each chemical structure is found to be represented by a mother skeleton, a bond list (bondlist) , an atom list (atomlist), a substituent list (subslist), and some other optional lists if necessary [16, 18]. According to this finding, each XƒMTEX command of fundamental level (cf. Fig. 4) is designed to be a linear connection between a mother skeleton and such lists as described above. This means that the set of XƒMTEX commands can be regarded as linear notation (XƒM notation) for representing chemical structures [16, 9]. In this context, the XƒMTEX system is regarded as an implementation of the XƒM notation to print out chemical structures by means of the LATEX system [18]. It follows that other applications based on the XƒM notation can be considered if we aim at other targets different from printing.

When a macro command `\cyctosine` is defined as follows by starting from the code in the top of Figure 4:

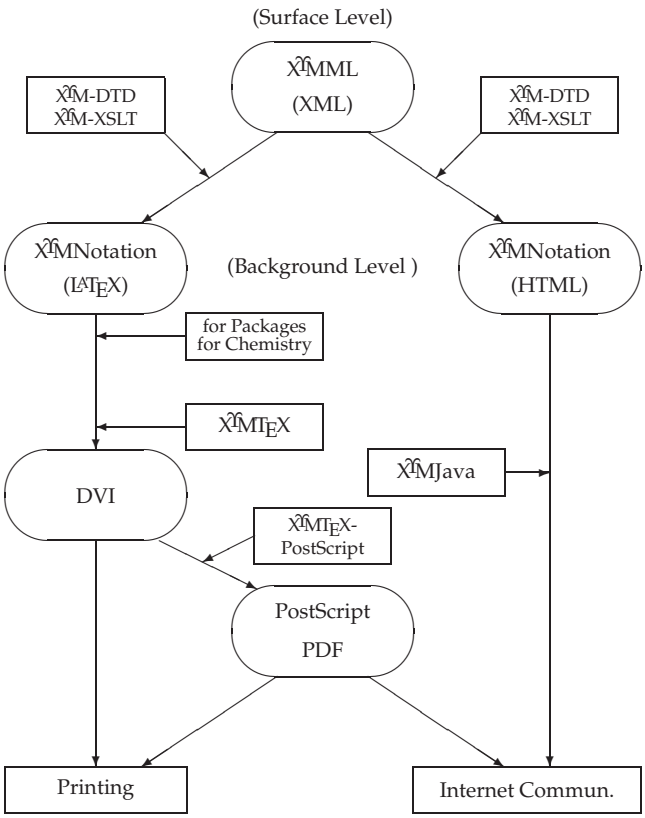


FIGURE 5. Communication system of chemical documents on the basis of XM Technology, which is composed of XMTEX, XM Notation, and XMML.

```
\def\cytosine#1{\sixheterovi[ce]%
{1==\downnobond{N}{H};3==N}{2D==O;4==NH$_{2}$;#1}}
```

we are able to print out the same chemical structure as shown in the bottom of Fig. 4 by writing down the following code:

```
\cytosine{5==HOCH$_{2}$}
```

The command name `\cytosine` is regarded as a new mother skeleton, which is accompanied by a subslist `5==HOCH$_{2}$`. This XMTEX command of higher level corresponds to the IUPAC name, i.e., 5-(hydroxymethyl)cytosine.

2.3 XMJava for Internet Communication

As shown in the right-hand branch of Figure 5, the XMJava system has been developed as a Java applet in order to interpret XM notations (XMTEX commands) embedded in a HTML document and to draw chemical structural formulas on a CRT display [29]. The object-oriented technique, especially the design-pattern approach, was applied to parse a XM notation in the XMJava system, where a chemical model was introduced to

encapsulate information on chemical structural formulas. The XyMJava system works well as a Java applet in internet browsers such as Internet Explorer. It follows that the XyMJava system is regarded as another implementation of the XyM notation to display chemical structures by means of the Java language.

2.4 XyMML Induced by XyM Notation

Because internet communication uses documents marked with tags based on markup languages, XyMML for marking up chemical structures has been developed by starting from XyMTEX, as shown in the top part of Fig. 2. The XyMML was originally developed on the basis of SGML (Standard Generalized Markup Language) [17, 18], which aims mainly at printing purposes. Because SGML was improved into XML for supporting internet communication, the XyMML has been extended to be based on XML [30]. As examples of the descriptions of the XyM-DTD (xymml.dtd for XyMML), the definitions of the XyM element and the subslist element are shown as follows:

```
<!ELEMENT XyM
 (XyMname,XyMnmsfx?,XyMbody?,
 skbndlst?,(bondlist|bndpatrn)?,
 atomlist?,subslist,omitlist?)>

<!ELEMENT subslist         (subst*)>
<!ELEMENT subst
 (substid,bondmod?,bndsbmod?,substval)>
<!ELEMENT substid         (#PCDATA)>
<!ELEMENT bondmod         (#PCDATA)>
<!ELEMENT bndsbmod         (#PCDATA)>
<!ELEMENT substval
((%gn0.el;)*|XyMyl|ryl|lyl)*>
```

To convert an XML document containing XyMML codes into an HTML document containing XyM notations, a translator program (XyM-XSLT, xymml.xsl) due to XSLT (Extensible Stylesheet Language Transformations) has been developed, as shown in the right-hand branch of Figure 2 [30]. The resulting HTML document is browsed by virtue of an internet browser loaded with the XyMJava system, where chemical structural formulas (due to XyM notations derived from XyMML codes) are displayed on a CRT display. On the other hand, another translator program (xymtex.xsl) due to XSLT has been also developed to convert the same XyMML document into a XyMTEX document. The document is processed by the TEX/LATEX typesetting system to give a printed matter containing chemical structural formulas due to XyMTEX. As a result, the XyM TEXnology (Figure 2) has been shown to have the potentiality of serving as a kernel for integrating internet communication, electronic publishing, and conventional publishing in chemistry.

The following code for drawing 2-fluoropyridine shows an example of a XyMML description in a XML file, in which xymml.dtd (the definition file for XyM-DTD) and xymml.xsl (the definition file for XyM-XSLT) are incorporated.

```
<xymml:XyM id="A3" x="300" y="300">
<xymml:XyMname>pyridine</xymml:XyMname>
<xymml:XyMnmsfx>vi</xymml:XyMnmsfx>
```

```
<xymml:subslist>
        <xymml:subst>
        <xymml:substid>2</xymml:substid>
        <xymml:substval>F</xymml:substval>
        </xymml:subst>
</xymml:subslist>
</xymml:XyM>
```

When the XML file is browsed by using an internet browser (e.g., Internet Explorer) supporting XML and XSLT as well as the XyMJava applet, it displays a structural formula of 2-fluoropyridine, as shown in the right part of Figure 6. Inside the internet browser (cf. the right-hand branch of Figure 5), the XyMML description shown above is translated into the HTML code shown in the left part of Figure 6, where the string `\pyridinevi{2==F}` for substituting for the value parameter represents a XyM notation produced automatically. The HTML code also contains several parameters for starting the XyMJava system, which are output as the attributes of the applet tag. Thereby, the string `\pyridinevi{2==F}` is parsed by the XyMJava system to display the formula of 2-fluoropyridine.

```
<table><tr><td align="center">
<a name="#A3"></a></td></tr>
<tr><td align="center">
<applet code="XyMApplet.class"
name="XyMJava" width="300" height="300">
<param name="xymnot"
value="\pyridinevi{2==F}">
</applet></td></tr>
<tr><td align="center">
<a href="#RA3"><b>4 </b>
</a></td></tr></table>
```

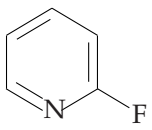


FIGURE 6. HTML code generated by XyMXSLT and the output of a structural formula by using XyMJava.

On a similar line, a XyMML description in a XML file is converted into the corresponding XyM notation in a LaTeX file by using a XML parser supporting XSLT [30], as shown in the left-hand branch of Figure 5. The resulting LaTeX file is treated by TeX/LaTeX with the XyMTeX so as to produce a printed document containing chemical structural formulas.

It should be noted here that the methodology of XyMML [17, 18] is quite different from that of ChemML [27]. Thus, XyMML has adopted markup formats derived from XyM notations (XyMTeX commands), where the concept of mother skeletons with substituents assures the correspondence of XyMML to the IUPAC nomenclature. For example, the XyMML code of 2-fluoropyridine contains a mother skeleton

```
<xymml:XyMname>pyridine</xymml:XyMname>
```

and a substituent represented by

`<xymml:substid>2</xymml:substid>` and `<xymml:substval>F</xymml:substval>`,

so that the XƒMML code meets the IUPAC name "2-fluoropyridine".

In contrast, ChemML has adopted formats equivalent to connection tables [2], which lack the concept of mother skeletons as well as the concept of substituents. For example, a primitive form of such a connection table for representing 2-fluoropyridine may be as follows:

```
1 N      1-2 s  1-6 d
2 C      2-3 d          2-7 s
3 CH     3-4 s
4 CH     4-5 d
5 CH     5-6 s
6 CH
7 F
```

where `s` denotes a single bond and `t` denotes a double bond. Because the ChemML code for 2-fluoropyridine is essentially equivalent to this connection table, a mother skeleton "pyridine" as well as a substituent "F" (in a different situation from "N") are not explicitly described so as to be recognized only by using some additional algorithm. This means that such information as "a pyridine derivative" obtained once by chemists is discarded or concealed during the process of ChemML coding.

3 XƒMTEX Typesetting of Printing Quality

3.1 Size Reduction

The XƒMTEX system has been improved to satisfy practical requirements of printing, e.g., easy handling and high quality. As an improvement in easy handling, the XƒMTEX version 3.00 has been designed to support the size reduction of chemical structural formulas [19]. This function works within the LATEX picture environment combined with a revised version [19] of the epic package [32]. Because the use of the epic package means that short bonds are drawn by multiple output of dots or short lines, the quality of small structural formulas is still insufficient.

The XƒMTEX system now supports two methods of size-reduction, as shown in Figure 7. The `\scalebox` command is used after including the graphicx package. Note that wedged bonds appearing in Figure 7 can be drawn only by XƒMTEX 4.00 or later versions.

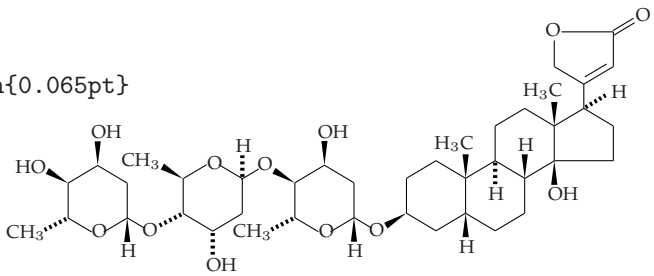
3.2 PostScript-Compatible Version of XƒMTEX

As for printing of high quality, the XƒMTEX version 4.00 has been released to provide chemical documents as PostScript files of high quality [11]. Because PostScript files containing chemical structures due to XƒMTEX can be converted into EPS (Encapsulated PostScript) files, which can be incorporated into other PostScript files. Such PostScript files can be converted into PDF files, which serve as a key to cover both traditional printing and internet communication, as shown in the bottom part of Figure 5.

The PostScript-compatible version of XƒMTEX (version 4.00 or later versions) serves as a versatile tool of preparing digital data for printing, which contain chemical structural formulas of high quality. The author published a monograph on organic com-

```
\def\digitoxin{%
\begin{XyMcompd}(3600,1500)(-2000,50){}{}
\steroid{%
3B==\lyl(3==O){3==\sixheterovi{1==O}{2==(yl);2GB==H;4B==OH;6A==CH$_{3}$;%
5B==\lyl(5==O){5==\sixheterov{1==O}{2==(yl);2FA==H;4A==OH;6B==CH$_{3}$;%
5A==\lyl(3==O){3==\sixheterovi{1==O}{2==(yl);2GB==H;4B==OH;%
6A==CH$_{3}$;5B==HO}}}}};%
5B==H;8B==H;9A==H;%
{10}B==\lmoiety{H$_{3}$C};{13}B==\lmoiety{H$_{3}$C};%
{14}B==OH;{17}GA==H;%
{{17}}==\fiveheterov[a]{4==O}%
{1==(yl);3D==O}}
\end{XyMcompd}}
```

```
{\changeunitlength{0.065pt}
\digitoxin}
```



```
\scalebox{0.65}{\digitoxin}
```

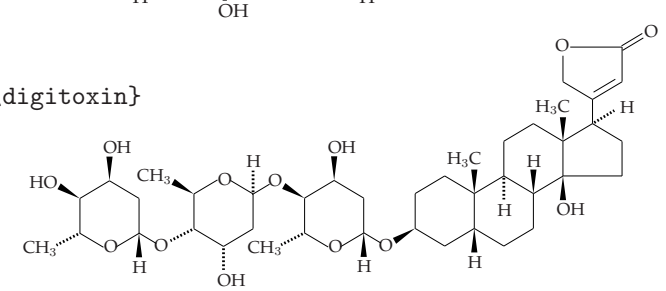


FIGURE 7. Two methods of size-reduction in XƎMTEX typesetting.

pounds for color photography [10] by using the XƎMTEX system. Figure 8 shows two pages selected from the monograph, which contain structural formulas drawn by XƎMTEX. A book review [31] on this book referred to the quality of involved structural formulas as follows:

> "Readers will be especially impressed by the structural formulas in the text, *created by software that the author has developed himself*; as well as showing the molecular structure very clearly, they also give complete information about functional groups, substituents, counterions, etc." [Emphasis has been added by the present author.]

Digital data generated by XƎMTEX under TEX/LATEX processing are capable of containing mathematical equations due to the original function of TEX/LATEX as well as chemical structural formulas due to the function of XƎMTEX. The author published a monograph on mathematical organic stereochemistry [12] by using the XƎMTEX system. In this monograph, the concept of mandala has been proposed to enumerate

In particular, the presence of a 2-methoxyethoxy group has been found to enhance the dye-releasing efficiency. This enhancement has been presumed to stem from an intramolecular chelating effect, as shown in Fig. 19.15 [9,10]. The 1,2-attack of a hydroxide ion to the sulfonimido group of **44** gives an adduct (**45**), where the 2-methoxyethoxy group chelates the hydrogen atom of the hydroxy group in terms of hydrogen bonding. This chelation results in the formation of an eight-membered ring of the intermediate (**45**), as designated by ⑧. The releasing of the sulfonamide (**39**) requires a further intermediate (**46**), which contains a six-membered ring designated by ⑥. The six-membered ring is presumed to be more plausible than the eight-membered ring.

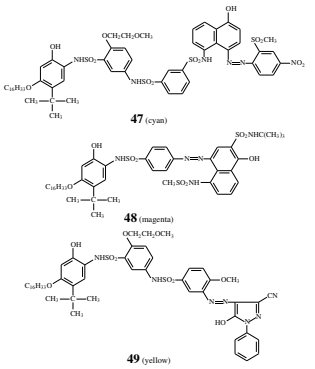


Figure 19.16. An improved set of cyan, magenta, and yellow *o*-sulfonamidophenol dye releasers for color reproduction in instant color photography [22].

The design of *o*-sulfonamidophenol moieties by using the model compounds described above has arrived at an improved set of dye releasers, as shown in Fig. 19.16 [22]. The specification of this patent contains experimental comparisons between dye releasers having a *t*-butyl group and the ones having a methyl group by using practical photographic films. Moreover, the effect of 2-alkoxyethoxy groups has been also incorporated in the dye releasers shown in Fig. 19.16 [23]. This type of *o*-sulfonamidophenol dye releasers has been used in the FI-10 film for the FOTORAMA system marketed by Fuji Photo Film in 1981 [1].

Syntheses

As collected in Table 19.2, various benzoxazole intermediates (**34a**–**34g**) for preparing 2-amino-4-alkyl-5-alkoxyphenols have been easily obtained in high yields. In particular, **34c** serves as a starting material for preparing the dye releasers listed in Fig. 19.16 [16,22].

Another versatile method for preparing 2-amino-4-alkyl-5-alkoxyphenols is the Friedel-Crafts alkylation of 2-acetamido-5-alkoxyphenol, as shown in Fig. 19.17 [24]. Thus, the key step is the conversion of **51** into **52**, where isobutene is used as a carbon source and Amberlite 15 (a synthetic ion-exchange resin from Rohm & Haas Co.) is used as a catalyst.

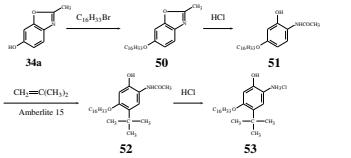


Figure 19.17. Synthesis of 2-amino-4-*t*-butyl-5-hexadecyloxyphenol by a Friedel-Crafts alkylation [24]. Homologs having other *t*-butyl groups have also been reported. This is a key for preparing *o*-sulfonamidophenol dye releasers.

Since the presence of a 2-alkoxyethoxy group has been found to enhance the dye-releasing efficiency in the model experiments described above [9,10] as well as in practical usage [22,23], 2-(2-alkoxyethoxy)benzenesulfonic acids having a nitrogen function have become important as key intermediates for preparing dye releasers of high efficiency. The introduction of such a 2-alkoxyethoxy group at the ortho position of a sulfonyl function is illustrated in Fig. 19.18 [25]. The first method described in one of the patents [23] has used sodium 2-methoxyethylate ($Na^+ \ ^-OCH_2CH_2OCH_3$) prepared by adding sodium hydride

FIGURE 8. LaTeX-typesetting with structural formulas drawn by XyMTeX [10].

Table 7.4. Subduction Table Based on LCRs for $\mathbf{D}_{2d}$ [2]

	$\downarrow\mathbf{C}_1$	$\downarrow\mathbf{C}_2$	$\downarrow\mathbf{C}_s$	$\downarrow\mathbf{C}_i$	$\downarrow\mathbf{S}_4$	$\downarrow\mathbf{C}_{2v}$	$\downarrow\mathbf{D}_2$	$\downarrow\mathbf{D}_{2d}$
$\mathbf{D}_{2d}(/\mathbf{C}_1)$	$8\mathbf{C}_1(/\mathbf{C}_1)$	$4\mathbf{C}_2(/\mathbf{C}_1)$	$4\mathbf{C}_s(/\mathbf{C}_1)$	$4\mathbf{C}_i(/\mathbf{C}_1)$	$2\mathbf{S}_4(/\mathbf{C}_1)$	$2\mathbf{C}_{2v}(/\mathbf{C}_1)$	$2\mathbf{D}_2(/\mathbf{C}_1)$	$\mathbf{D}_{2d}(/\mathbf{C}_1)$
$\mathbf{D}_{2d}(/\mathbf{C}_2)$*	$4\mathbf{C}_1(/\mathbf{C}_1)$	$4\mathbf{C}_2(/\mathbf{C}_2)$	$2\mathbf{C}_s(/\mathbf{C}_1)$	$2\mathbf{C}_i(/\mathbf{C}_1)$	$2\mathbf{S}_4(/\mathbf{C}_2)$	$2\mathbf{C}_{2v}(/\mathbf{C}_2)$	$2\mathbf{D}_2(/\mathbf{C}_2)$	$\mathbf{D}_{2d}(/\mathbf{C}_2)$
$\mathbf{D}_{2d}(/\mathbf{C}_s)$	$4\mathbf{C}_1(/\mathbf{C}_1)$	$2\mathbf{C}_2(/\mathbf{C}_1)$ +$2\mathbf{C}_s(/\mathbf{C}_s)$	$\mathbf{C}_s(/\mathbf{C}_s)$	$2\mathbf{C}_i(/\mathbf{C}_1)$	$\mathbf{S}_4(/\mathbf{C}_1)$	$\mathbf{C}_{2v}(/\mathbf{C}_s)$	$\mathbf{D}_2(/\mathbf{C}_2)$ +$\mathbf{D}_2(/\mathbf{C}_2)$	$\mathbf{D}_{2d}(/\mathbf{C}_s)$
$\mathbf{D}_{2d}(/\mathbf{C}_s)$	$4\mathbf{C}_1(/\mathbf{C}_1)$	$2\mathbf{C}_2(/\mathbf{C}_1)$	$2\mathbf{C}_s(/\mathbf{C}_1)$ +$2\mathbf{C}_s(/\mathbf{C}_s)$	$\mathbf{C}_s(/\mathbf{C}_s)$	$\mathbf{S}_4(/\mathbf{C}_1)$	$\mathbf{C}_{2v}(/\mathbf{C}_s)$ +$\mathbf{C}_{2v}(/\mathbf{C}_s')$	$\mathbf{D}_2(/\mathbf{C}_1)$	$\mathbf{D}_{2d}(/\mathbf{C}_s)$
$\mathbf{D}_{2d}(/\mathbf{S}_4)$*	$2\mathbf{C}_1(/\mathbf{C}_1)$	$2\mathbf{C}_2(/\mathbf{C}_2)$	$\mathbf{C}_s(/\mathbf{C}_1)$	$\mathbf{C}_s(/\mathbf{C}_1)$	$2\mathbf{S}_4(/\mathbf{S}_4)$	$\mathbf{C}_{2v}(/\mathbf{C}_2)$	$\mathbf{D}_2(/\mathbf{C}_2)$	$\mathbf{D}_{2d}(/\mathbf{S}_4)$
$\mathbf{D}_{2d}(/\mathbf{C}_{2v})$	$2\mathbf{C}_1(/\mathbf{C}_1)$	$2\mathbf{C}_2(/\mathbf{C}_2)$	$\mathbf{C}_s(/\mathbf{C}_1)$	$2\mathbf{C}_i(/\mathbf{C}_s)$	$\mathbf{S}_4(/\mathbf{C}_2)$	$2\mathbf{C}_{2v}(/\mathbf{C}_{2v})$	$\mathbf{D}_2(/\mathbf{C}_2)$	$\mathbf{D}_{2d}(/\mathbf{C}_{2v})$
$\mathbf{D}_{2d}(/\mathbf{D}_2)$*	$2\mathbf{C}_1(/\mathbf{C}_1)$	$2\mathbf{C}_2(/\mathbf{C}_2)$	$2\mathbf{C}_s(/\mathbf{C}_s)$	$\mathbf{C}_i(/\mathbf{C}_1)$	$\mathbf{S}_4(/\mathbf{C}_2)$	$\mathbf{C}_{2v}(/\mathbf{C}_2)$	$2\mathbf{D}_2(/\mathbf{D}_2)$	$\mathbf{D}_{2d}(/\mathbf{D}_2)$
$\mathbf{D}_{2d}(/\mathbf{D}_{2d})$	$\mathbf{C}_1(/\mathbf{C}_1)$	$\mathbf{C}_2(/\mathbf{C}_2)$	$\mathbf{C}_s(/\mathbf{C}_s)$	$\mathbf{C}_i(/\mathbf{C}_s)$	$\mathbf{S}_4(/\mathbf{S}_4)$	$\mathbf{C}_{2v}(/\mathbf{C}_{2v})$	$\mathbf{D}_2(/\mathbf{D}_2)$	$\mathbf{D}_{2d}(/\mathbf{D}_{2d})$

* Forbidden CR. See Chapter 7 of Fujita's book [2].

$\mathcal{A}^{\dagger}$, the two permutation diagrams corresponding to $\mathbf{C}_s = \{I, \sigma_{d(1)}\}$ are selected so as to give Fig. 7.10 (for I and $\sigma_{d(1)}$), which corresponds to Fig. 7.9. By careful comparison between the first box of the top row and the counterpart of the bottom row in Fig. 7.9, we could find that the eight vertices of **1** are divided into four orbits (i.e., {1, 2}, {3, 8}, {4, 7}, and {5, 6}).[l] In contrast, this division can be found more easily by Fig. 7.10, because one of such orbits (i.e., {1, 2}) is marked with solid circles.[m] Hence, the fixation of the first box of the top row during the action of $\mathbf{C}_s = \{I, \sigma_{d(1)}\}$ is more clearly demonstrated so as to give the first box of the bottom row in Fig. 7.10.

Thus, by comparing the two diagrams of Fig. 7.10, one can find easily that the $\{\mathcal{A}_1^{\dagger}\}$ and the $\{\mathcal{A}_3^{\dagger}\}$ are respectively immobile (fixed or stabilized), while the $\mathcal{A}_2^{\dagger}$ and the $\mathcal{A}_4^{\dagger}$ are interchanged into each other.[n] As a result, the symmetry restriction from $\mathbf{D}_{2d}$ to $\mathbf{C}_s$ divides the four-membered orbit of assemblies ($\mathcal{A}^{\dagger}$) into two one-membered orbits ($\{\mathcal{A}_1^{\dagger}\}$ and $\{\mathcal{A}_3^{\dagger}\}$) and a two-membered orbit($\{\mathcal{A}_2^{\dagger}, \mathcal{A}_4^{\dagger}\}$).

Because the resulting one-membered orbit $\{\mathcal{A}_1^{\dagger}\}$ (or $\{\mathcal{A}_3^{\dagger}\}$) is fixed by $\mathbf{C}_s$, the orbit is concluded to be governed by the LCR $\mathbf{C}_s(/\mathbf{C}_s)$. Because the two-membered orbit of assemblies $\{\mathcal{A}_2^{\dagger}, \mathcal{A}_4^{\dagger}\}$) is fixed by $\mathbf{C}_1$ and the two assemblies $\mathcal{A}_2^{\dagger}$ and $\mathcal{A}_4^{\dagger}$ are permuted by the action of the $\sigma_{d(1)}$-operation, the orbit is concluded to be governed by the LCR $\mathbf{C}_s(/\mathbf{C}_1)$. Thereby, the subduction represented by Fig. 7.10 is summarized into eq. 7.46. This procedure shows that Fig. 7.10 is capable of deriving results equivalent to those derived by Fig. 7.9. Thus, the same subduction table as Table 7.4 can be alternatively obtained. Then,

[l] Strictly speaking, this division of the eight vertices is controlled by the subduction of the RRR $(\mathbf{C}_1\backslash\mathbf{D}_{2d} \downarrow \underline{\mathbf{C}_s}$ (cf. eq. 6.58 in Chapter 6), where the underlined $\mathbf{C}_s$ is the local symmetry of the LCR $\mathbf{D}_{2d}(/\underline{\mathbf{C}_s})$, which governs the orbit of assemblies $\mathcal{A}^{\dagger} = \{\mathcal{A}_1^{\dagger}, \mathcal{A}_2^{\dagger}, \mathcal{A}_3^{\dagger}, \mathcal{A}_4^{\dagger}\}$. In general, the symmetrical behavior of a **K**-molecule derived from a regular body of **G** is described by the LCR $\mathbf{G}(/\underline{\mathbf{K}})$, while the division of |**G**| vertices in the regular body of **G** is controlled by the subduction $(\mathbf{C}_1\backslash\mathbf{G} \downarrow \underline{\mathbf{K}}$. Note that the underlined **K**'s are selected to be common. This feature is the basis of the concept of mandalas, as described in the next chapter.

[m] Although the formulation using $\mathcal{A}^{\dagger}$ is more understandable than the formulation using $\mathcal{A}^{*}$, the latter is adopted as a primary formulation because its generality is superior to the former. For example, one can select {3, 8}, {1, 2}/{3, 8}, or {1, 2}/{4, 7} as vertices to be marked with solid circles. The formulation using $\mathcal{A}^{*}$ connotes these alternatives as well as the special case shown in Fig. 7.10 (i.e., the selection of {1, 2}).

[n] Do not confuse an orbit of vertices with an orbit of assemblies. The discussion described here is concerned with the subduction of the orbit of assemblies, i.e., $\mathcal{A}^{\dagger} = \{\mathcal{A}_1^{\dagger}, \mathcal{A}_2^{\dagger}, \mathcal{A}_3^{\dagger}, \mathcal{A}_4^{\dagger}\}$, which is governed by the LCR $\mathbf{D}_{2d}(/\mathbf{C}_s)$.

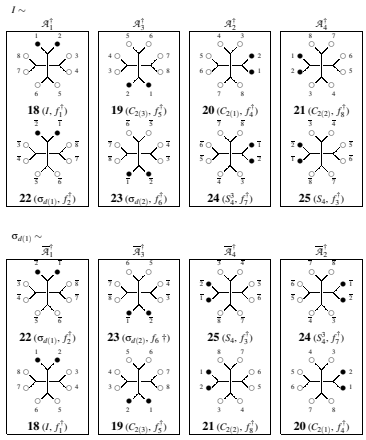


Figure 7.10. The action of I and $\sigma_{d(1)}$ on the $\mathbf{C}_s$-molecule (the four $\mathbf{C}_s$-assemblies) listed in Fig. 7.8. The alignment shown in this diagram corresponds to an ordered set, $\mathcal{A}_0^{\dagger} = \{\mathcal{A}_1^{\dagger}, \mathcal{A}_3^{\dagger}, \mathcal{A}_4^{\dagger}, \mathcal{A}_2^{\dagger}\}$.

the other tables for the LCRs (the USCI-CF table, the USCI table and the mark table) are obtained similarly, where they are equivalent to those for the RCRs described in Chapter 6 (Tables 6.8, 6.9, and 6.10).

Exercise 7.13. Derive eqs. 7.42–7.49 and Table 7.4 diagrammatically by following the procedure given above for Fig. 7.10. Compare the derivation with the one described for obtaining eqs. 6.61–6.68 in Chapter 6.

7.4 Mandalas as Nested Regular Bodies

The discussions described in Section 7.2 have essentially followed Chapters 13 and 15 of Fujita's book [2], although a more diagrammatical approach has been adopted by following partly the treatment reported recently [4,5]. Because the discussions have required

FIGURE 9. LaTeX-typesetting with structural formulas drawn by XyMTeX as well as mathematical expressions [12].

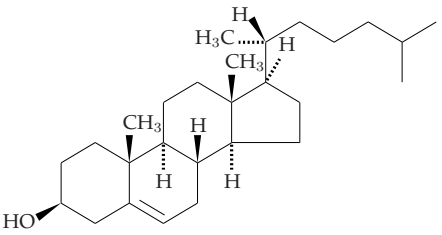
three-dimensional structural formulas, where both graphical expressions (mandalas and structural formulas) by XƎMTEX and mathematical expressions (equations for combinatorial enumeration) by TEX/LATEX are essential components. Figure 9 shows two pages selected from the monograph, which contain structural formulas drawn by XƎMTEX as well as mathematical expressions by the original function of TEX/LATEX.

3.3 XƎMTEX Version 4.04

One of remaining targets in the extension of the XƎMTEX system is to expand a repertoire of commands of higher level (e.g., `\cytosine` described above), most of which can be derived from commands of fundamental level (cf. Fig. 4). Such commands of higher level simplify XƎMTEX codes for drawing structural formulas, because most parts of each formula are already specified in the form of a mother skeleton. Following this guideline, the latest version 4.04 of XƎMTEX supports complicated structural formulas of steroid derivatives [33, 34].

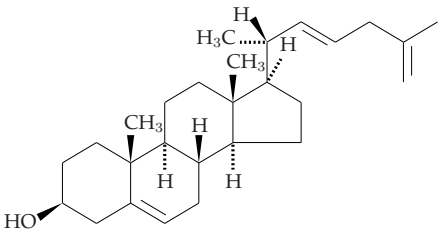
Figure 10 shows steroid derivatives drawn by XƎMTEX version 4.04 [33]. Because the command `\cholestane` gives a steroid skeleton having substituents necessary to compose a cholestane derivative, a minimum set of additional substituents is required to complete such complicated structural formulas, as found in the codes cited at the top of Figure 10.

```
\cholestane[e]{3B==HO}
```



cholesterol

```
\cholestane[e{Zc}{Zg}]{3B==HO}
```



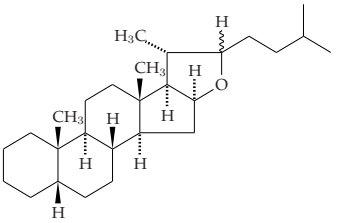
$(22E)$-cholesta-5,22,25-trien-3β-ol

FIGURE 10. Steroid derivatives drawn by XƎMTEX 4.04.

Stereochemical representations of bonds (wedged bonds, bold broken bonds, and wavy bonds) are supported as shown in Figure 11, which shows furostan derivatives drawn by XƎMTEX 4.04 [33]. The structural formula of (22ξ)-5β-furostan is easily drawn by using the command `\furostanbeta`, where the descriptor `{22}U==H` outputs a wavy bond of a 22ξ-hydrogen, whose configuration is unspecified.

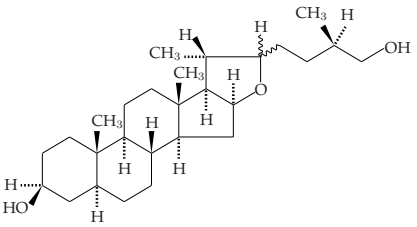
If both of the two bonds at the C-22 are desired to be drawn as wavy bonds as found in the second example of Figure 11, a lower-level command `\steroid` is used in combination with the command `\tetramethylene` and `\fivefusev`, where `3U` (a designator for a wavy bond) is specified in the subslist of the command `\fivefusev`.

(22ξ)-5β-furostan



```
\begin{XyMcompd}(1900,1250)(300,200){}{}
\furostanbeta{{22}U==H}
\end{XyMcompd}
```

(22ξ, 25R)-5α-furostan-3β,26-diol (dihydropseudotigogenin)



```
\begin{XyMcompd}(2400,1300)(50,150){}{}
\steroid[{s{\fivefusev{2==O}
{4SA==CH$_{3}$;4SB==H;5FA==H;3FU==H;%
3U==\tetramethylenei{}{1==(yl);%
3SB==CH$_{3}$;3SA==H;4W==OH}}{e}}}]
{3SB==HO;3SA==H;5A==H;%
8B==H;9A==H;{14}A==H;%
{10}B==CH$_{3}$;%
{13}B==CH$_{3}$;%
{20}S==CH$_{3}$;{16}FA==H}
\end{XyMcompd}
```

FIGURE 11. Furostan derivatives drawn by XyMTEX 4.04.

4 XyMTEX for Submitting Articles to Journals

4.1 Online Systems of Submission

In the Internet, PDF has been wide-spread as a de facto standard format, because Adobe Reader (as an add-on software of an internet browser) is freely available to read such PDF documents. Hence, many scientific journals have recently adopted PDF documents in their online submission/peer-reviewing processes, where electronic data (PDF files) are used in place of printout matters. This trend would grow steadily, because PDF was authorized as an ISO standard (ISO 32000-1) in 2008.

Because dvi files prepared from TEX/LATEX-documents can be converted into PDF files (directly or via PostScript files), the trend described in the preceding paragraph allows us to use TEX/LATEX as software for preparing articles as electronic data of PDF.

1. **Semi-online submission:** A journal published from the Chemical Society of Japan (*Bull. Chem. Soc. Jpn.*) receives PDF documents submitted via the online system of J-STAGE. After acceptance, the corresponding tex file with EPS files are sent by mail if the author uses TEX/LATEX. These files are converted into a file of in-house TEX-format, which is processed by 3B2 (Advent) to give a printed matter or SGML (XML) data [24].
2. **Online submission:** Journals published from Elsevier (e.g., *Tetrahedron*) and from Springer (e.g., *J. Math. Chem.*) adopt online submission systems based on Editorial Manager (Aries). On the other hand, journals from American Chemical

Society (e.g., *J. Chem. Inf. Model.*) and from Wiley (e.g., *Angew. Chem.*) adopt online submission systems based on Manuscript Central (ScholorOne). These online systems accept LaTeX2$_\varepsilon$ documents, where, after each tex file with EPS files are sent on line, they are processed on line to produce the corresponding PDF file automatically. An online manual common to systems based on Manuscript Central is available from [35].

4.2 Workflow for Submitting Articles to Journals

Figure 12 shows a workflow for semi-online and online submission of articles containing structural formulas drawn by XyMTeX.

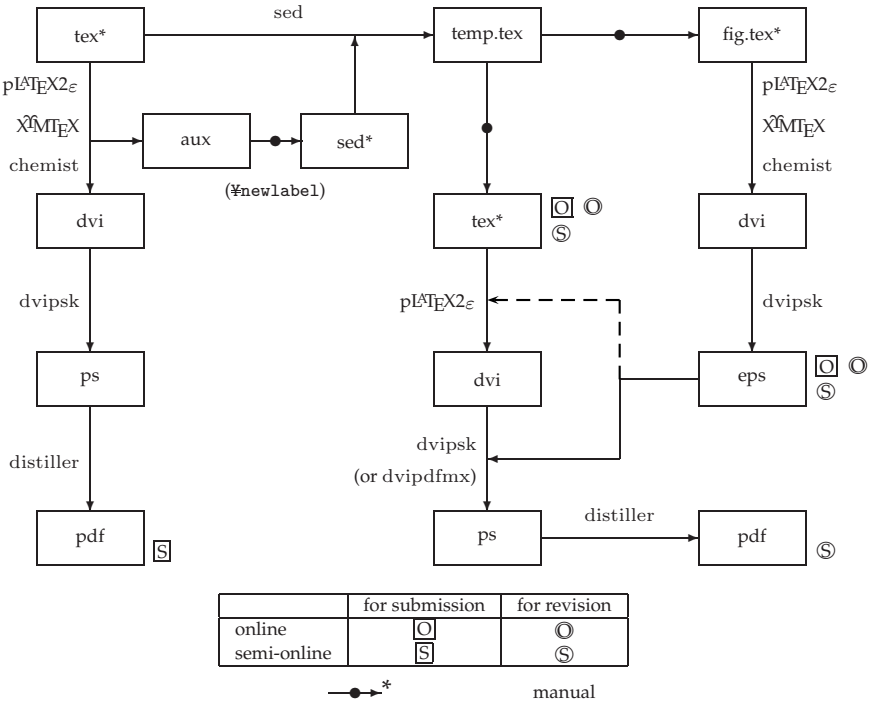


FIGURE 12. Online and semi-online submission of articles containing structural formulas drawn by XyMTeX. The present article was prepared by the semi-online procedure.

Pdf files containing structural formulas directly produced by XyMTeX A draft for online submission or a first manuscript for semi-online submission is generated by the process shown in the leftmost branch of Figure 12, where structural formulas are directly produced by XyMTeX codes incorporated in the text. For example, the code for Figure 10 in the tex* file (named "`example1.tex`" tentatively) on the leftmost branch of Figure 12 contains XyMTeX codes for drawing steroid derivatives as follows:

```
%example1.tex
\documentclass{article}
```

```
\usepackage{xymtexps}
\usepackage{graphicx}
\begin{document}
\begin{figure}
\begin{center}
\begin{tabular}{cc}
\verb/\cholestane[e]{3B==HO}/ & \verb/\cholestane[e{Zc}{Zg}]{3B==HO}/ \\
\scalebox{0.75}{\cholestane[e]{3B==HO}} &
\scalebox{0.75}{\cholestane[e{Zc}{Zg}]{3B==HO}} \\
cholestrol & $(22E)$-cholesta-5,22,25-trien-3$\beta$-ol \\
\end{tabular}
\end{center}
\caption[]{Steroid derivatives drawn by \XyMTeX{}~4.04.}
\label{ff:B2}
\end{figure}
\end{document}
```

where the command `\scalebox` is used after including the graphicx package and several style files of the PStricks package are included automatically. Note that the declaration `\usepackage{xymtexps}` allows the inclusion of all the style files of the XyMTEX system.

When XyMTEX is used in the PostScript-compatible mode, a dvi file generated from a tex file is converted into a ps (PostScript) file by means of dvipsk (or dvips). The resulting ps file is further converted into a pdf file by using an appropriate converter such as Adobe Distiller. For the purpose of semi-online submission, the pdf file (designated by [S]) is submitted via e-mail to the editorial office of a journal.

Eps files of figures To submit a first manuscript in the case of online submission or to submit a final (revised) manuscript via semi-online, the data of figures due to XyMTEX are converted into eps files, as shown in the rightmost branch of Figure 12. A single eps file corresponding to each figure is generated by starting from a fig.tex* file, which is obtained from the original tex file via a temp.tex file. The generation of the intermediate temp.tex file is conducted by means of a sed program using a sed file for adjusting reference keys. The latter sed file is obtained manually from an aux file generated automatically during the LaTeX2$_\varepsilon$ processing of the original tex file.

For example, a fig.tex* file (named "`example1-Fig.tex`") is obtained from "`example1.tex`" via the corresponding temp.tex file. The sed program has replaced the temporary label `\label{ff:B2}` by a systematic one `\label{f10}`, which corresponds to Figure 10. Note that `\pagestyle{empty}` is added manually to avoid page numbering and that the figure environment is commented out.

```
%example1-Fig.tex
\documentclass{article}
\usepackage{xymtexps}
\usepackage{graphicx}
\pagestyle{empty}
\begin{document}
%\begin{figure}
\begin{center}
\begin{tabular}{cc}
\verb/\cholestane[e]{3B==HO}/ & \verb/\cholestane[e{Zc}{Zg}]{3B==HO}/ \\
\scalebox{0.75}{\cholestane[e]{3B==HO}} &
```

```
\scalebox{0.75}{\cholestane[e{Zc}{Zg}]{3B==HO}} \\
cholestrol & $(22E)$-cholesta-5,22,25-trien-3$\beta$-ol \\
\end{tabular}
\end{center}
%\caption[]{Steroid derivatives drawn by \XyMTeX{}~4.04.}
%\label{f10}
%\end{figure}
\end{document}
```

As shown in the rightmost branch of Figure 12, a dvi file (`example1-Fig.dvi`) generated from a fig.tex* file (`example1-Fig.tex`) is converted into an eps file per figure by means of dvipsk. For example, the fig.tex* file generates an eps file named "`example1-Fig10.eps`" tentatively.

Manuscripts with eps files A manuscript to be sent to the editorial office of a journal is generated from a tex* file which includes such eps files by using the command `\includegraphcs` after inputting the graphicx package. Such a tex* file is generated manually from a temp.tex file, as shown in the middle branch of Figure 12. For example, a tex* file named "`example1-final.tex`" is manually generated as follows by starting from the temp file derived from "`example1.tex`". The declaration `\usepackage{xymtexps}` in its preamble may be commented out, because the inclusion of the XyMTeX system is no longer necessary to draw structural formulas. In this case, however, the logo XyMTeX in the caption of the figure requires the inclusion of the XyMTeX system.

```
%example1-final.tex
\documentclass{article}
\usepackage{xymtexps}
\usepackage{graphicx}
\begin{document}
\begin{figure}
\begin{center}
\includegraphics[scale=1]{example1-Fig10.eps}
\end{center}
\caption[]{Steroid derivatives drawn by \XyMTeX{}~4.04.}
\label{f10}
\end{figure}
\end{document}
```

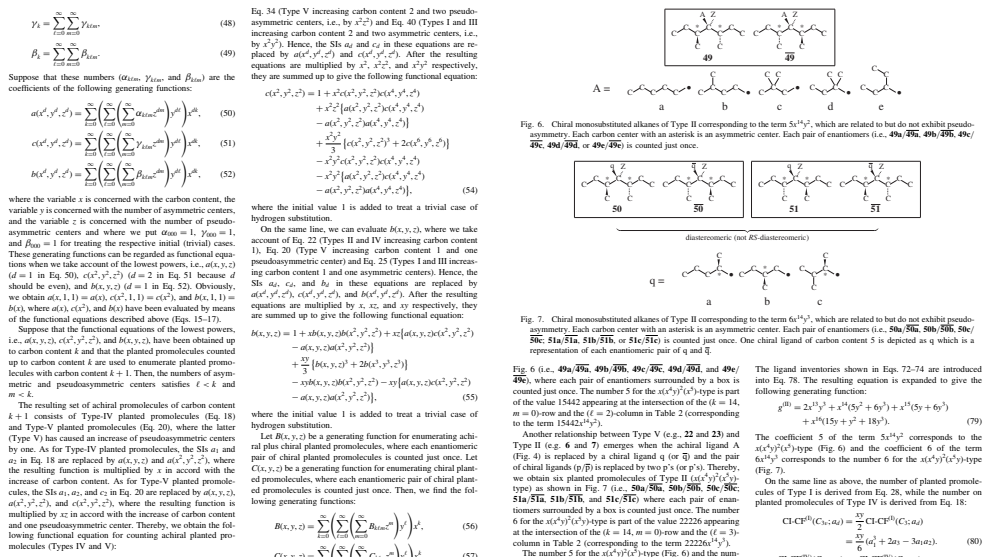
The tex* file ("`example1-final.tex`") is processed by LaTeX2$_\varepsilon$ to give a dvi file, in which `\special` commands store the names of eps files to be incorporated. Then, the dvi file is converted into a ps file by using dvipsk and further converted into a pdf file. Direct conversion systems of the dvi file into a pdf file (e.g., dvipdfmx) are available.

Manuscripts to be submitted For semi-online submission of a first manuscript, the pdf file (designated by Ⓢ) is sent via e-mail to the editorial office of a journal. After acceptance, a pdf file, a tex* file, and eps files, which are marked by Ⓢ, are sent to the office by e-mail or mail. On the other hand, the process of online submission requires a set of a tex* file and eps files, which are marked by Ⓞ. They are submitted by using an online submission system. After acceptance, the revised set of a tex* file and eps files, which are marked by Ⓞ, is sent by using the same online submission system.

As found by these items, online and semi-online submission deeply depend on the PDF technology. Because structural formulas drawn by TeX/LaTeX–XyMTeX can

be easily converted into PDF, the present method based on TEX/LATEX–XƎMTEX turns out to be one of predominant routes to cope with online and semi-online submission. This issue is confirmed by the author's recent experience in online and semi-online submission as follows.

1. Figure 13 shows two pages of the author's contribution as an example article [13], which contain complicated mathematical equations (left) as well as chemical structural formulas generated by the XƎMTEX system (right). Because the article was awarded the BCSJ Award Article Vol. 81 Feb., the pdf file of its full text is available without fee from the home page of the BCSJ [36].


200 Bull. Chem. Soc. Jpn. Vol. 81, No. 2 (2008) BCSJ AWARD ARTICLE

$$\gamma_k = \sum_{\ell=0}^{\infty}\sum_{m=0}^{\infty}\gamma_{k\ell m}, \tag{48}$$

$$\beta_k = \sum_{\ell=0}^{\infty}\sum_{m=0}^{\infty}\beta_{k\ell m}. \tag{49}$$

Suppose that these numbers ($\alpha_{k\ell m}$, $\gamma_{k\ell m}$, and $\beta_{k\ell m}$) are the coefficients of the following generating functions:

$$a(x^d, y^d, z^d) = \sum_{k=0}^{\infty}\left(\sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty}\alpha_{k\ell m}z^{dm}\right)y^{d\ell}\right)x^{dk}, \tag{50}$$

$$c(x^d, y^d, z^d) = \sum_{k=0}^{\infty}\left(\sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty}\gamma_{k\ell m}z^{dm}\right)y^{d\ell}\right)x^{dk}, \tag{51}$$

$$b(x^d, y^d, z^d) = \sum_{k=0}^{\infty}\left(\sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty}\beta_{k\ell m}z^{dm}\right)y^{d\ell}\right)x^{dk}, \tag{52}$$

where the variable x is concerned with the carbon content, the variable y is concerned with the number of asymmetric centers, and the variable z is concerned with the number of pseudo-asymmetric centers and where we put $\alpha_{000} = 1$, $\gamma_{000} = 1$, and $\beta_{000} = 1$ for treating the respective initial (trivial) cases. These generating functions can be regarded as functional equations when we take account of the lowest powers, i.e., $a(x, y, z)$ ($d = 1$ in Eq. 50), $c(x^2, y^2, z^2)$ ($d = 2$ in Eq. 51 because d should be even), and $b(x, y, z)$ ($d = 1$ in Eq. 52). Obviously, we obtain $a(x, 1, 1) = a(x)$, $c(x^2, 1, 1) = c(x^2)$, and $b(x, 1, 1) = b(x)$, where $a(x)$, $c(x^2)$, and $b(x)$ have been evaluated by means of the functional equations described above (Eqs. 15–17).

Suppose that the functional equations of the lowest powers, i.e., $a(x, y, z)$, $c(x^2, y^2, z^2)$, and $b(x, y, z)$, have been obtained up to carbon content k and that the planted promolecules counted up to carbon content k are used to enumerate planted promolecules with carbon content $k + 1$. Then, the numbers of asymmetric and pseudoasymmetric centers satisfies $\ell < k$ and $m < k$.

The resulting set of achiral promolecules of carbon content $k + 1$ consists of Type-IV planted promolecules (Eq. 18) and Type-V planted promolecules (Eq. 20), where the latter (Type V) has caused an increase of pseudoasymmetric centers by one. As for Type-IV planted promolecules, the SIs a_1 and a_2 in Eq. 18 are replaced by $a(x, y, z)$ and $a(x^2, y^2, z^2)$, where the resulting function is multiplied by x in accord with the increase of carbon content. As for Type-V planted promolecules, the SIs a_1, a_2, and c_2 in Eq. 20 are replaced by $a(x, y, z)$, $a(x^2, y^2, z^2)$, and $c(x^2, y^2, z^2)$, where the resulting function is multiplied by xz in accord with the increase of carbon content and one pseudoasymmetric center. Thereby, we obtain the following functional equation for counting achiral planted promolecules (Types IV and V):

$$\begin{aligned} a(x, y, z) = 1 &+ xa(x, y, z)a(x^2, y^2, z^2) \\ &+ xz\left\{a(x, y, z)c(x^2, y^2, z^2)\right. \\ &\left. - a(x, y, z)a(x^2, y^2, z^2)\right\}, \end{aligned} \tag{53}$$

where the initial value 1 is added to treat a trivial case of hydrogen substitution.

To evaluate $c(x^2, y^2, z^2)$, we should take account of Eq. 36 (Types II and IV increasing carbon content 2, i.e., by x^2), Eq. 34 (Type V increasing carbon content 2 and two pseudo-asymmetric centers, i.e., by x^2z^2) and Eq. 40 (Types I and III increasing carbon content 2 and two asymmetric centers, i.e., by x^2y^2). Hence, the SIs a_d and c_d in these equations are replaced by $a(x^d, y^d, z^d)$ and $c(x^d, y^d, z^d)$. After the resulting equations are multiplied by x^2, x^2z^2, and x^2y^2 respectively, they are summed up to give the following functional equation:

$$\begin{aligned} c(x^2, y^2, z^2) = 1 &+ x^2c(x^2, y^2, z^2)c(x^4, y^4, z^4) \\ &+ x^2z^2\left\{a(x^2, y^2, z^2)c(x^4, y^4, z^4)\right. \\ &\left. - a(x^2, y^2, z^2)a(x^4, y^4, z^4)\right\} \\ &+ \frac{x^2y^2}{3}\left\{c(x^2, y^2, z^2)^3 + 2c(x^6, y^6, z^6)\right\} \\ &- x^2y^2c(x^2, y^2, z^2)c(x^4, y^4, z^4) \\ &- x^2y^2\left\{a(x^2, y^2, z^2)c(x^4, y^4, z^4)\right. \\ &\left. - a(x^2, y^2, z^2)a(x^4, y^4, z^4)\right\}, \end{aligned} \tag{54}$$

where the initial value 1 is added to treat a trivial case of hydrogen substitution.

On the same line, we can evaluate $b(x, y, z)$, where we take account of Eq. 22 (Types II and IV increasing carbon content 1), Eq. 20 (Type V increasing carbon content 1 and one pseudoasymmetric center) and Eq. 25 (Types I and III increasing carbon content 1 and one asymmetric centers). Hence, the SIs a_d, c_d, and b_d in these equations are replaced by $a(x^d, y^d, z^d)$, $c(x^d, y^d, z^d)$, and $b(x^d, y^d, z^d)$. After the resulting equations are multiplied by x, xz, and xy respectively, they are summed up to give the following functional equation:

$$\begin{aligned} b(x, y, z) = 1 &+ xb(x, y, z)b(x^2, y^2, z^2) + xz\left\{a(x, y, z)c(x^2, y^2, z^2)\right. \\ &\left. - a(x, y, z)a(x^2, y^2, z^2)\right\} \\ &+ \frac{xy}{3}\left\{b(x, y, z)^3 + 2b(x^3, y^3, z^3)\right\} \\ &- xyb(x, y, z)b(x^2, y^2, z^2) - xy\left\{a(x, y, z)c(x^2, y^2, z^2)\right. \\ &\left. - a(x, y, z)a(x^2, y^2, z^2)\right\}, \end{aligned} \tag{55}$$

where the initial value 1 is added to treat a trivial case of hydrogen substitution.

Let $B(x, y, z)$ be a generating function for enumerating achiral plus chiral planted promolecules, where each enantiomeric pair of chiral planted promolecules is counted just once. Let $C(x, y, z)$ be a generating function for enumerating chiral planted promolecules, where each enantiomeric pair of chiral planted promolecules is counted just once. Then, we find the following generating functions:

$$B(x, y, z) = \sum_{\ell=0}^{\infty}\left(\sum_{l=0}^{\infty}\left(\sum_{m=0}^{\infty}B_{l\ell m}z^m\right)y^l\right)x^\ell, \tag{56}$$

$$C(x, y, z) = \sum_{\ell=0}^{\infty}\left(\sum_{l=0}^{\infty}\left(\sum_{m=0}^{\infty}C_{l\ell m}z^m\right)y^l\right)x^\ell, \tag{57}$$

where we place $B_{000} = 1$ and $C_{000} = 1$ for the trivial cases and $B_{l\ell m}$ and $C_{l\ell m}$ represent the numbers of planted promolecules at issue, which have the molecular formula x^ℓ have ℓ asymmetric centers and m pseudoasymmetric centers.

$$B(x, y, z) = \frac{1}{2}\left\{a(x, y, z) + b(x, y, z)\right\} \tag{58}$$

Fig. 6. Chiral monosubstituted alkanes of Type II corresponding to the term $5x^{14}y^5$, which are related to but do not exhibit pseudo-asymmetry. Each carbon center with an asterisk is an asymmetric center. Each pair of enantiomers (i.e., **49a**/**$\overline{\text{49a}}$**, **49b**/**$\overline{\text{49b}}$**, **49c**/**$\overline{\text{49c}}$**, **49d**/**$\overline{\text{49d}}$**, or **49e**/**$\overline{\text{49e}}$**) is counted just once.

Fig. 7. Chiral monosubstituted alkanes of Type II corresponding to the term $6x^{14}y^5$, which are related to but do not exhibit pseudo-asymmetry. Each carbon center with an asterisk is an asymmetric center. Each pair of enantiomers (i.e., **50a**/**$\overline{\text{50a}}$**, **50b**/**$\overline{\text{50b}}$**, **50c**/**$\overline{\text{50c}}$**, **51a**/**$\overline{\text{51a}}$**, **51b**/**$\overline{\text{51b}}$**, or **51c**/**$\overline{\text{51c}}$**) is counted just once. One chiral ligand of carbon content 5 is depicted as q which is a representation of each enantiomeric pair of q and $\overline{\text{q}}$.

Fig. 6 (i.e., **49a**/**$\overline{\text{49a}}$**, **49b**/**$\overline{\text{49b}}$**, **49c**/**$\overline{\text{49c}}$**, **49d**/**$\overline{\text{49d}}$**, and **49e**/**$\overline{\text{49e}}$**), where each pair of enantiomers surrounded by a box is counted just once. The number 5 for the $x(x^4y^1)^2(x^2)$-type is part of the value 15442 appearing at the intersection of the ($k = 14$, $m = 0$)-row and the ($\ell = 2$)-column in Table 2 (corresponding to the term $15442x^{14}y^5$).

Another relationship between Type V (e.g., **22** and **23**) and Type II (e.g. **6** and **7**) emerges when the achiral ligand A (Fig. 4) is replaced by a chiral ligand q (or $\overline{\text{q}}$) and the pair of chiral ligands (q/$\overline{\text{q}}$) is replaced by two p's (or p's). Thereby, we obtain six planted promolecules of Type II ($x(x^4y^1)^2(x^2y)$-type) as shown in Fig. 7 (i.e., **50a**/**$\overline{\text{50a}}$**, **50b**/**$\overline{\text{50b}}$**, **50c**/**$\overline{\text{50c}}$**, **51a**/**$\overline{\text{51a}}$**, **51b**/**$\overline{\text{51b}}$**, and **51c**/**$\overline{\text{51c}}$**) where each pair of enantiomers surrounded by a box is counted just once. The number 6 for the $x(x^4y^1)^2(x^2y)$-type is part of the value 22226 appearing at the intersection of the ($k = 14$, $m = 0$)-row and the ($\ell = 3$)-column in Table 2 (corresponding to the term $22226x^{14}y^5$).

The number 5 for the $x(x^4y^1)^2(x^2)$-type (Fig. 6) and the number 6 for the $x(x^4y^1)^2(x^2y)$-type (Fig. 7) are more clearly demonstrated by starting from Eq. 23. Because Eq. 23 counts two enantiomers of each pair separately, each pair can be counted just once by calculating the half of Eq. 23 as follows:

$$\begin{aligned} \text{CI-CF}^{\text{III}}(C_{3v}; a_d, b_d) &= \frac{1}{2}\text{CI-CF}^{\text{III}}(C_s; a_d, b_d) \\ &= \frac{1}{2}(b_1b_2 - a_1a_2). \end{aligned} \tag{78}$$

The ligand inventories shown in Eqs. 72–74 are introduced into Eq. 78. The resulting equation is expanded to give the following generating function:

$$\begin{aligned} g^{(\text{II})} = 2x^{13}y^3 &+ x^{14}(5y^2 + 6y^3) + x^{15}(5y + 6y^3) \\ &+ x^{16}(15y + y^2 + 18y^3). \end{aligned} \tag{79}$$

The coefficient 5 of the term $5x^{14}y^2$ corresponds to the $x(x^4y^1)(x^2)$-type (Fig. 6) and the coefficient 6 of the term $6x^{14}y^3$ corresponds to the number 6 for the $x(x^4y^1)^2(x^2y)$-type (Fig. 7).

On the same line as above, the number of planted promolecules of Type I is derived from Eq. 18, while the number on planted promolecules of Type IV is derived from Eq. 18:

$$\begin{aligned} \text{CI-CF}^{(\text{I})}(C_{3v}; a_d) &= \frac{xy}{2}\text{CI-CF}^{(\text{II})}(C_s; a_d) \\ &= \frac{xy}{6}(a_1^2 + 2a_3 - 3a_1a_2). \end{aligned} \tag{80}$$

$$\begin{aligned} \text{CI-CF}^{(\text{IV})}(C_{3v}; a_d) &= x\text{CI-CF}^{(\text{IV})}(C_s; a_d) \\ &= xa_1a_2. \end{aligned} \tag{81}$$

The ligand inventories shown in Eqs. 72–74 are introduced into Eqs. 80 and 81. The resulting equations are expanded to give the following generating functions:

$$g^{(\text{I})} = 10x^{16}y, \tag{82}$$

$$g^{(\text{IV})} = 25x^{16}. \tag{83}$$

FIGURE 13. An example article with structural formulas drawn by XƎMTEX as well as mathematical expressions by TEX/LATEX [13].

Although the pages cited in Figure 13 have been typeset by the printing agency for *Bull. Chem. Soc. Jpn.*, the original manuscript sent to the editorial office via semi-online submission has almost equivalent qualities with respect to mathematical equations and chemical structural formulas. Figure 14 shows two mathematical equations due to TEX/LATEX (eqs. 56 and 57 cited in Figure 13) and chemical structural formulas drawn by XƎMTEX (Figure 6 cited in Figure 13), which have appeared in the original manuscript corresponding to the pages of Figure 13.

2. If structural formulas drawn by XƎMTEX are separately converted into respective eps files (the rightmost branch of Figure 12) and if a tex file coupled with such eps files is verified to be properly processed by using the LATEX system (the mid-

$$B(x,y,z)=\sum_{k=0}^{\infty}\left(\sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty}B_{k\ell m}z^m\right)y^\ell\right)x^k \tag{56}$$

$$C(x,y,z)=\sum_{k=0}^{\infty}\left(\sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty}C_{k\ell m}z^m\right)y^\ell\right)x^k, \tag{57}$$

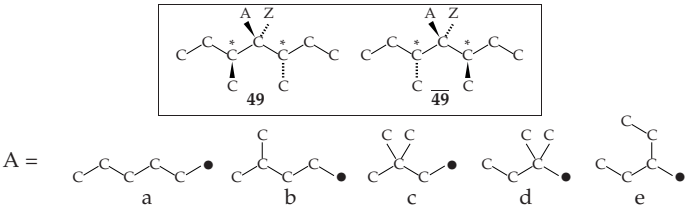


FIGURE 14. Two mathematical equations due to TEX/LATEX and chemical structural formulas drawn by XƒMTEX, which appear in the original manuscript corresponding to the pages of Figure 13.

dle branch of Figure 12), the verified tex file and the eps files can be submitted to the editorial office of a journal which accepts TEX/LATEX files through its own online system of submission. They are processed in the online system so as to be automatically converted into a pdf file (via a dvi file), which is subjected to peer-review (or is passed to publication after acceptance). The author's contributions to *Tetrahedron* [14] and to *J. Math. Chem.* [15] are examples of such online submission, where the accepted articles have been disclosed in the Internet before publication of printed journals.

Moreover, the present article itself has been typeset by using TEX/LATEX–XƒMTEX and submitted to this journal (*The Asian Journal of TEX*) according to the semi-online procedure. Thereby, it can be concluded that the combination of TEX/LATEX and XƒMTEX is versatile to accomplish document production covering mathematical equations and chemical structural formulas.

5 Conclusion

The XƒMTEX system for typesetting chemical structural formulas has been extended to XƒM TEXnology which covers internet communication of chemical documents with chemical structural formulas in addition to the original purpose, i.e., printing of high quality. XƒMTEX commands are recognized as XƒM notation (linear notation), which is extended into XƒMML (markup language) for characterizing chemical structural formulas. An internet browser loaded with XML, XƒMML, XƒMXSLT, and XƒMJava converts XML documents with XƒMML into HTML documents with XƒM notations. The resulting HTML documents are interpreted by the XƒMJava system, which has been developed as a Java applet of an internet browser, so as to display chemical structural formulas in the Internet. In addition, the same XML documents with XƒMML are converted into LATEX documents with XƒMTEX commands (the same as XƒM notations), which are capable of printing out chemical structural formulas of high quality. The

latest version (4.04) of XyMTEX supports additional functions for drawing complicated structures such as steroids. Such LaTeX documents with XyMTEX formulas can be converted into PDF documents directly or via PostScript documents. Applications of such PDF documents in online or semi-online submission to scientific journals have been discussed.

References

1. A. M. Coghill and L. R. Garson (eds.), *The ACS Style Guide: Effective Communication of Scientific Information*, 3rd ed., American Chemical Society, Washington, D.C., 2006.
2. C. H. Davis and J. E. Rush, *Information Retrieval and Documentation in Chemistry*, Greenwood, Westport, 1974.
3. S. Fujita, *Contributions to the Chemistry of Nitrogen-Containing Reactive Species and Strained Rings—Nitrene, Aziridines and Heterophanes*, Ph.D. thesis, Kyoto Univ., 1972.
4. _____, *Symmetry and Combinatorial Enumeration in Chemistry*, Springer-Verlag, Berlin-Heidelberg, 1991.
5. _____, *LaTeX for Chemists and Biochemists*, Tokyo Kagaku Dojin, Tokyo, 1993.
6. _____, *Typesetting Structural Formulae with the Text Formatter TEX/LaTeX*, Comput. Chem. **18** (1994), 109–116.
7. _____, *XyMTEX for Drawing Chemical Structural Formulas*, TUGboat **16** (1995), no. 1, 80–88.
8. _____, *XyMTEX—Typesetting Chemical Structural Formulas*, Addison-Wesley Japan, Tokyo, 1997.
9. _____, *Revival of Linear Notations. XyM Notation as a Linear Notation*, CICSJ Bulletin **22** (2004), 110–115.
10. _____, *Organic Chemistry of Photography*, Springer-Verlag, Berlin-Heidelberg, 2004.
11. _____, *Development of XyMTEX2PS for PostScript Typesetting of Chemical Documents Containing Structural Formulas*, J. Comput. Chem. Jpn. **4** (2005), 69–78.
12. _____, *Diagrammatical Approach to Molecular Symmetry and Enumeration of Stereoisomers*, University of Kragujevac, Faculty of Science, Kragujevac, 2007.
13. _____, *Numbers of Achiral and Chiral Monosubstituted Alkanes Having a Given Carbon Content and Given Numbers of Asymmetric and Pseudoasymmetric Centers*, Bull. Chem. Soc. Jpn. **81** (2008), 193–219.
14. _____, *A New Scheme for Investigating Geometric and Stereoisomeric Features in Stereochemistry*, Tetrahedron **65** (2009), 1581–1592. `http://dx.doi.org/10.1016/j.tet.2008.12.066`.
15. _____, *Correlation Diagrams of Stereoisograms for Characterizing Stereoisomers of Cyclobutane Derivatives*, J. Math. Chem. **47** (2010), no. 1, 145–166. `http://dx.doi.org/10.1007/s10910-009-9539-z`
16. S. Fujita and N. Tanaka, *XyM Notations for Electronic Communication of Organic Chemical Structures*, J. Chem. Inf. Comput. Sci. **39** (1999), 903–914.
17. _____, *XyM Markup Language (XyMML) for Electronic Communication of Chemical Documents Containing Structural Formulas and Reaction Schemes*, J. Chem. Inf. Comput. Sci. **39** (1999), 915–927.

18. _____, *XymTEX (Version 2.00) as Implementation of the XyM Notation and XyM Markup Language*, TUGboat **21** (2000), no. 1, 7–14.

19. _____, *Size Reduction of Chemical Structural Formulas in XymTEX (version 3.00)*, TUGboat **22** (2001), no. 4, 285–289.

20. M. Goossens, F. Mittelbach, S. Rahtz, D. Roegel, and H. Voß, *The LaTeX Graphics Companion*, 2nd ed., Addison Wesley, UpperSaddle River, 2008.

21. M. Goossens and S. Rahtz, *The LaTeX Web Companion: Integrating TeX, HTML, and XML*, Addison Wesley Longman, Reading, 1999.

22. R. T. Haas and K. C. O'Kane, *Typesetting Chemical Structure Formulas with the Text Formatter TeX/LaTeX*, Comput. Chem. **11** (1989), 251–271.

23. J. Hagen and A. F. Otten, *PPCHTEX: Typesetting Chemical Formulas in TeX*, TUGboat **17** (1996), no. 1, 54–66.

24. K. Hayashi, T. Nakaya, and T. Ohta, *Study for E-Journal Production in Japan and E-Publishing with NLM-DTD XML*, JOHO KANRI **51** (2009), 902–913.

25. I. Klöckl, OCHEM, April 2001. `CTAN:support/ochem/`

26. Z. Li, H. Wan, Y. Shi, and P. Ouyang, *Personal Experience with Four Kinds of Chemical Structure Drawing Software. Review on ChemDraw, ChemWindow, ISIS/Draw, and ChemSketch*, J. Chem. Inf. Comput. Sci. **44** (2004), 1886–1890.

27. P. Murray-Rust and H. S. Rzepa, *Chemical Markup Language and XML. Part I. Basic Principles*, J. Chem. Inf. Comput. Sci. **39** (1999), 928–942.

28. A. C. Norris and A. L. Oakley, *Electronic Publishing and Chemical Text Processing*, in: M. Clark (ed.), TeX—Applications, Uses, Methods, Ellis Horwood, Chichester, 1990, Chapter 26, pp. 207–225.

29. N. Tanaka and S. Fujita, *XyMJava System for Word Wide Web Communication of Organic Chemical Structures*, J. Computer Aided Chem. **3** (2002), 37–47.

30. N. Tanaka, T. Ishimaru, and S. Fujita, *WWW (World Wide Web) Communication and Publishing of Structural Formulas by XyMML (XyM Markup Language)*, J. Computer Aided Chem. **3** (2002), 81–89.

31. M. W. Tausch, *Organic Chemistry of Photography (Book Review)*, Angew. Chem. Int. Ed. **44** (2005), 2629.

32. For epic macros, see S. Podar, "Enhancements of the Picture Environment of LaTeX", Manual for Version 1.2 dated July 14, 1986.

33. For an online manual, see S. Fujita, "XymTEX (Version 4.04) for Typesetting Chemical Structural Formulas. An Extension for Drawing Steroid Derivatives." (June 15, 2009), which is available from `http://xymtex.com/fujitas3/xymtex/xym404/xympdf/xymtx404.pdf`.

34. XymTEX version 4.04 is available from the author's homepage: `http://xymtex.com/`.

35. Online manual: `http://mcv3help.manuscriptcentral.com/tutorial/Author.pdf`.

36. The full-text PDF file of the BCSJ Award Article, Vol. 81 Feb is available from `http://www.jstage.jst.go.jp/article/bcsj/81/2/81_193/_article`.