STRUCTURAL REPRESENTATION
OF POLYMERS

Introduction

Since virtually the dawn of modern chemistry, scientists have found it desirable to represent chemical substances by name, and also, where possible, by a structure that shows how the atoms are connected. Storage of information about a substance, whether on paper or in a computer database, is of value only if a substance can be located, for example by use of its name, molecular formula, structure, or some combination of these; otherwise the substance record, together with any attached data, is essentially lost.

Knowledge of these three key properties of a substance affords versatility in its retrieval. This is relatively easy for substances that are nonpolymeric in nature; although the molecular formula of a substance is usually not unique, its molecular formula together with a unique systematic name, a unique structure, or both, are sufficient to distinguish a nonpolymeric substance from all other substances.

However, for polymers, these statements are not completely true. By its very nature, a polymer is usually not a single, discrete substance, but rather a mixture of somewhat different molecules. Although these have strong similarities to one another, at the molecular level probably no two individual molecules are identical in terms of such properties as chain length (molecular weight), linearity (degree of branching or cross-linking), arrangement of repeating units (alternating, block, random, statistical, etc), or conformation (spatial arrangement of atoms within a chain).

Thus, polymers are inherently different from nonpolymers, and many attempts have been made to define what constitutes a polymer. The International Union of Pure and Applied Chemistry (IUPAC), an international consortium...
composed of professional scientists who propose and recommend standards for chemical nomenclature and structure representation, defines a polymer as “a substance composed of molecules characterized by the multiple repetition of one or more species of atoms or groups of atoms (constitutional units) linked to each other in amounts sufficient to provide a set of properties that do not vary markedly with the addition or removal of one or a few constitutional units” (1).

A polymer may be assigned a unique name, e.g., poly(oxy-1,2-ethanediyl) (structure-based) or poly(oxirane) (source-based), but since such nomenclature (see NOMENCLATURE) incorporates no properties such as molecular weight, these names per se are insufficient to distinguish one sample from another. Likewise, any single complete molecular formula that represents a discrete molecule within the polymer is inadequate to describe the whole polymeric mixture, and to write a molecular formula for every distinct molecule present in a polymer would be hopeless. Therefore, compromises must be made. This article describes the current situation with regard to structural representation of polymers.

General Discussion

Polymer structural representation has two purposes:

1. identification—to distinguish a given polymer from all others
2. description—to record structural details that will enable the polymer to be retrieved in all searches for appropriate classes of polymers.

The ideal representation of a polymer is perhaps one that includes all known structural information; that is, the structure of the polymer itself and the structures of the substance(s) from which it was formed. This is usually impracticable; most polymers are represented in just one of these two ways.

**Structure-based polymer representation:** the polymer is represented in terms of the structure of its constitutional or structural repeating unit. Two common expressions are in use: “constitutional repeating unit” (CRU)—preferred by IUPAC; and “structural repeating unit” (SRU)—preferred by Chemical Abstracts Service (CAS) (2) (see Chemical Abstracts Service Information System); the two are virtually synonymous, and SRU will be used in this article. For example, polyethylene can be represented by the SRU \((\text{CH}_2\text{CH}_2)_n\) (the largest fragment that can be written without repetition) or by the SRU \((\text{CH}_2=\text{CH}_2)_n\) (obtained by opening the double bond in ethylene). (IUPAC recommends that in printed communications the subscript “n” as well as any other letters denoting multiplicities of polymeric sequences be printed in italic type or, in the absence of italics, be underlined. This article, which addresses primarily database records and their retrieval, indicates subscripts in the upright style without underlining.)

**Source-based polymer representation** (also called product-by-process representation): the polymer is represented in terms of the names and structures (where known) of the substance(s) from which the polymer was formed. For example, polyethylene is represented by the expression \((\text{CH}_2=\text{CH}_2)_n\), because the structure of the monomer, ethylene, from which it is made is \(\text{CH}_2=\text{CH}_2\).

Source-based polymer representation preceded structure-based representation because polymers were made before they were structurally characterized. The
introduction of structure-based representation, however, did not render source-based representation obsolete, and currently the two complementary systems coexist. Source-based representation continues to exist for several reasons:

1. For some polymers the method is preferable to structure-based representation. For example, the source materials used to manufacture Bakelite™ or Melmac™ are, in practice, more important than their structures; even if all the SRUs present in these highly cross-linked polymers could be delineated, to represent them thus would serve little purpose.

2. Many source-based names are sufficiently explicit for the structures to be instantly deducible by scientists; therefore it is unnecessary to state them.

3. For condensation polymers prepared from three or more monomers, it is impossible to write a meaningful SRU unless special additional rules are formulated to create an artificial or stylized one.

4. Source-based names often lend themselves more easily to verbal and written communications, eg, “polycaprolactone” versus “poly[oxy(1-oxo-6-hexanediyl)]”.

5. Some problems concerning structure-based representation remain unresolved in terms of their delineation in a computer database. For example, the structure representation of an SRU within a larger SRU still poses software problems for CAS. The CAS Registry System, designed over 30 years ago, cannot handle a structural representation like the one shown in Figure 1a (SRU within a larger SRU). It does provide a source-based representation, as shown in Figure 1b, with two monomers of the type \( (A \cdot B)_x \), one of which is an SRU with two hydroxy end groups and the other a dicarboxylic acid (2).

Thus, CAS can register an SRU within a polymer, eg, \([U-(A)_n-V \cdot B]_x\), and a polymer within a polymer, eg, \([(A)_x \cdot B]_x\), but not an SRU within an SRU, eg, \([(A)_n \cdot B]_n\).

Such complex structure-based representations pose no problems for more recent systems such as MDL’s ISIS™ program with features designed expressly for polymer structure storage and retrieval.

In contrast, for many polymer classes, especially those that result from considerable chemical transformation of starting materials, it is more useful to represent them by structure than in terms of the source materials used to prepare them.

---

**Fig. 1.** Representations of polyester from hexanedioic (adipic) acid and poly(oxy-1,2-ethanediyl) [poly(oxyethylene)]; asterisks at ends of SRU denote unspecified end groups.
The structure of a polymer is generally displayed by representing it as a parenthesized expression with a subscript letter outside the closing parenthesis. Two such expressions are commonly used:

1. For structure-based polymer representation, the expression used is $(-A-)_n$, where A represents the SRU of the polymer and the subscript n indicates that A is repeated an indefinite number of times. An SRU is an atom-by-atom or group-by-group representation of the smallest unit that, when repeated, in theory—and sometimes in reality—represents the polymer. Thus, poly(oxymethylene) can be expressed as $(-\text{O-CH}_2-)_n$.

2. For source-based polymer representation, the expression used is $(A \cdot B \cdot \ldots)_x$, where A, B, etc, represent the atom-by-atom or group-by-group representation(s) of polymerizable substance(s) used to prepare the polymer and the subscript x indicates that A, B, etc, are repeated an indefinite number of times. Thus, nylon-6,6 can be expressed as $(\text{H}_2\text{N-}(\text{CH}_2)_6\text{-NH}_2\cdot\text{HO}_2\text{C-(CH}_2)_4\text{-CO}_2\text{H})_x$.

Expressions such as $(-\text{O-CH}_2-)_n$ and $(\text{H}_2\text{N-}(\text{CH}_2)_6\text{-NH}_2\cdot\text{HO}_2\text{C-(CH}_2)_4\text{-CO}_2\text{H})_x$ depict actual structures, not molecular formulae. Although expressions such as $(\text{CH}_2\text{O})_n$ and $(\text{C}_6\text{H}_{15}\text{N}_2\cdot\text{C}_6\text{H}_{10}\text{O}_4)_x$ are used by CAS to represent the molecular formulae of poly(oxymethylene) and nylon-6,6 respectively, these expressions are, per se, neither structure-based nor source-based representations of these two polymers. There is a link, however, between the full structural representation and its molecular formula. The subscript n is used consistently for both an SRU and its molecular formula; similarly, the subscript x is used consistently for both a homopolymer (copolymer) structure expressed in terms of monomer(s) and its corresponding molecular formula.

These two types of representation, both of which are more theoretical than actual (although they can be actual), have advantages and disadvantages.

For structure-based polymer representation the advantages and disadvantages are as follows:

1. Advantages:
   a. The theoretical (which is sometimes the actual) repeating unit of the polymer is depicted precisely and unambiguously.
   b. If structural irregularities in the polymer are ignored, there is presumably no doubt about the exact structure.
   c. End groups may be indicated by modification of the expression $(-A-)_n$ to a format such as $U(-A-)_n-\text{V}$, where U and V are structures of end groups; thus, $\alpha$-methacryloyl-$\omega$-methoxypoly(oxymethylene) is represented by the expression $\text{CH}_2\text{=C(C}_3\text{H}_5\text{)}\text{-CO-}(\text{-O-CH}_2\text{-CH}_2-)_n\text{-OCH}_3$. The corresponding molecular formula is written in the format $(\text{A})_n\text{W}$, where A is the molecular formula of the SRU (exclusive of end groups) and W is the summation of molecular formulae of end groups U and V; thus, W (atom count) = U (atom count) + V (atom count). In the
latter case, however, without a name or structure, it is sometimes im-
possible to determine which atoms belong to U and which to V. Thus, the
two SRUs $\alpha$-methacryloyl-$\omega$-methoxypoly(oxyethylene) and $\alpha$-acyl-
loyl-$\omega$-ethoxypoly(oxyethylene) are represented by the same molecular formula
expression $(C_2H_4O)_nC_5H_8O_2$.

(2) **Disadvantages:**

a. The method fails to indicate the monomers from which the polymer was
prepared; for a polymer chemist this is frequently an important consid-
eration.

b. For polymers prepared from asymmetrical monomers that give rise to a
product containing several SRUs, the pedantry involved in naming every
possibility can become tedious.

(3) For source-based polymer representation the advantage and disadvantages
are as follows.

(4) **Advantage:**

a. A polymer can always be represented in terms of the monomers used
to synthesize it, even if the structures of some of them are imprecisely
defined (eg, castor oil). It must be pointed out, however, that monomers
can never, per se, show the structure of a polymer containing them; the
reader must use other knowledge to deduce it.

(5) **Disadvantages:**

a. The atom-by-atom structure of the polymer is not depicted, even when it
can be deduced with some degree of certainty.

b. Scattering of information may occur from either or both of the follow-
ing sources: (a) the same polymer may be synthesized from different
monomers; (b) under different reaction conditions the same monomer(s)
may give rise to more than one polymer structure.

Some polymers are traditionally represented by the structure-based method,
some by the source-based method, and others by both methods, but always as
separate expressions, rather than one all-encompassing expression.

Provided (1) both structure-based and source-based types of representation
are stored, (2) all practical representational possibilities for both types are stored,
and (3) the storage/retrieval system is capable of retrieving all stored representa-
tions when only one is specified, a key advantage of using separate expressions for
representations of a polymer is that searchers can find a polymer by searching any
possible representation. Currently, condition (1) is sometimes met, whereas con-
ditions (2) and (3) are frequently not, which results in scattered data and frequent
failure to retrieve critical information.

Seemingly random use of structure-based and source-based representations
with no cross-referencing between them has resulted in decades of confusion.
Fig. 2. “Phase-shift” versions of PET represented as an SRU. While it is possible to write seven structures, as far as CAS searching is concerned, both the human searcher and the CAS Messenger computer program recognize that (a) = (e), (b) = (d), and (f) = (g).

For even the simplest of polymers, a searcher is faced with three fundamental questions: (1) is the desired polymer represented as structure-based or source-based; (2) if it is structure-based, how is the structure drawn; (3) if it is source-based, what monomers were used? To take a polyester example, a question such as whether parent acids or simple derivatives (e.g., methyl esters or phenyl esters or acid chlorides) were used is often far more important to a polymer scientist than the structure of the final polymer.

Any polymer representation used for identification must be predictable, i.e., easily arrived at by a searcher applying a set of rational and invariable rules. Without knowledge of rules for determining the order of an atom-by-atom representation of an SRU, a searcher seeking information on poly(ethylene terephthalate) (PET), for example, would theoretically have to search for all seven structures shown in Figure 2.

Each of these would have a unique name, and so a search by name, vice one by structure, would not narrow the number of possibilities.

CAS (2) and IUPAC (3) have devised rules for the orientation of SRUs that, when followed, reduce a searcher’s choices to one of the above seven; the preferred structure for PET is shown as Figure 2a.

Until September 2002, searchers faced an essentially insolvable problem in a search for source-based representations for PET. A searcher might think of searching only for the two representations shown as Figures 3a and 3b; the representation shown as Figure 3c, which represents PET synthesized by a modern, industrially significant route, would thus not be retrieved.

A search for source-based representations for a polymer therefore depended upon a searcher’s ability to surmise how it might be represented; in other words, it was a guessing game.

The three above-listed questions arise for searches involving only simple polymers; for more complex polymers, further questions arise such as how many

\[
\text{(a)} \quad \text{HO} - \text{CH}_2 - \text{CH}_2 - \text{OH} \cdot \text{MeO}_2\text{C} - p\cdot\text{C}_6\text{H}_4 - \text{CO}_2\text{Me}, \\
\text{(b)} \quad \text{HO} - \text{CH}_2 - \text{CH}_2 - \text{OH} \cdot \text{ClCO} - p\cdot\text{C}_6\text{H}_4 - \text{COCl}, \\
\text{(c)} \quad \text{HO} - \text{CH}_2 - \text{CH}_2 - \text{OH} \cdot \text{HO}_2\text{C} - p\cdot\text{C}_6\text{H}_4 - \text{CO}_2\text{H},
\]

Fig. 3. Different expressions for source-based representations of PET.
comonomers were used to prepare them, whether end groups need to be considered, and whether the polymers are aftertreated (post-treated), block, branched, comb, cross-linked, dendritic, graft, star, etc.

To say that one polymer is or is not different from another will always require some arbitrary judgment of what is or is not significant in polymer identification. For example, are the two polymers represented by the source-based expressions shown as Figures 3a and 3c sufficiently “different” to be worthy of two separate representations? In terms of source-based polymer registration, CAS regards them as two separate polymers. Some industrial consumers of PET may find that the presence of carboxy end groups rather than methoxycarbonyl ones makes no difference for their particular application, while others may find that end group composition makes an appreciable difference to product properties; thus, the answer may depend upon many factors.

It must also be pointed out that neither structure-based nor source-based representation currently takes into account any structural irregularities. Real polymers, as synthesized, are seldom structurally perfect. Depending on the type of polymer, various kinds of structural irregularities may arise during a polymerization process.

Some examples are given below:

(1) PET may contain irregularities known as “ether defect,” a colloquial name indicating the presence of some \( \text{O} - \text{CH}_2 - \text{CH}_2 - \text{O} - \text{CH}_2 - \text{CH}_2 - \text{O} \) moieties.

(2) Depending upon polymerization concentration, difunctional acrylic compounds (eg, ethylene diacylate) may give polymers that are hyperbranched, hyper-cross-linked, or both.

(3) Low density polyethylene is irregularly branched.

(4) Polymers formed between simple diamines (eg, 1,5-pentanediamine) and simple dihalo compounds (eg, 1,5-dichloropentane) are frequently branched, cross-linked, or both.

Such irregularities, which may vary from batch to batch, can affect polymer properties appreciably.

Thus polymer structures represented on paper or stored in a database as ideal or theoretical representations are frequently not true representations of actual polymer structures. Before September 2002 all these problems and questions posed significant difficulties for the average polymer searcher. The following sections discuss what has been achieved to date in terms of organizing polymer structures and nomenclature, and of facilitating polymer searching.

In September 2002, CAS introduced a new on-line search-tool called POLYLINK, which is designed to solve the above problems regarding “phase-shift” and retrieval of structure-based representations and their source-based counterparts (4).

**Organizations and Databases**

Over the last 50 years, two major organizations, CAS and IUPAC, have tackled the daunting task of how best to represent polymer structures. Their achievements
are impressive, and the two organizations are in virtually complete agreement on structure-based and source-based representations of polymers, both on paper and in searchable on-line databases such as the CAS Registry File. However, much still remains to be done; some aspects of copolymer representation need enhancement, and such newer areas as aftertreated (post-treated), comb, dendritic, and star polymers are handled inadequately.

In addition to these organizations, computer software companies such as Daylight and MDL Information Systems, Inc. (MDL) offer for sale computer programs that are suitable for either mainframe or personal computers, and many companies use them to create their own proprietary databases containing agricultural chemicals, pharmaceuticals, polymers, etc. Some computer programs accommodate chemical structures only, while others accommodate textual data only. MDL’s ISIS™/Host is a computer program that enables users to combine chemical structures from one file with textual data from another to produce an integrated product. MDL’s ISIS program enables users to register a wide variety of polymer structures.

Storage and Retrieval of Polymer Structures

Storage of Polymer Structures. CAS has developed comprehensive sets of nomenclature rules and structure conventions that have become standards in the world of chemical literature (see Chemical Abstracts Service Information System). CAS selects polymers for indexing according to a set of rules (5). The Registry File records are created from chemical information published in books, patents, dissertations, scientific journal articles, etc.

MDL’s ISIS program permits registration of polymer structure representations as source-based, structure-based, or both. Users who create their own databases with such a program are therefore free to store polymer representations in any way they wish within program limitations. The optionally available “duplicate detection” feature alerts a person registering a substance that the structure is already in the database. For example, the three possible representations of poly(oxy-1,2-ethanediyl) shown as Figures 4a–4c are “perceived” by MDL’s ISIS program as being “phase-shift” versions of the same polymer; use of the duplicate detection feature during registration of any one of these three representations can alert a registrar that it is already registered as 4a, 4b, and 4c.

Retrieval of Polymer Structures. It is axiomatic that to conduct an efficient search, searchers need to know the rules of the database or file being searched. Although this article is not intended to be a teaching manual on how to search polymer files or databases, some basic differences between searching the CAS Registry File and searching an MDL database by means of the ISIS program merit discussion.

![Fig. 4. “Phase-shift” versions of poly(oxy-1,2-ethanediyl).]
Searchers who choose to search a publicly available file such as the CAS Registry File must learn the rules imposed on them by CAS (2). The CAS Registry File is remotely searchable via the Scientific and Technical Information Network (STN) in Columbus, Ohio. Anyone with an STN account and a computer equipped with the necessary hardware and software may access this file. As of September 2002, the Registry file contains names, structures, or both for over 42 million substances, of which nearly 934,000 are polymers.

Those who choose to search a private database such as one created with MDL’s ISIS must learn a different set of search techniques. ISIS™/Draw, the ISIS chemical drawing package, is a presentation package for polymer structures. ISIS/Draw allows scientists to draw polymer structures with ease and speed, and then to search and share them. Researchers also have the enhanced capability of accessing 2D and 3D databases, with a wide range of search options and a link to global data. ISIS/Draw includes full annotation capabilities, which enable researchers to label collections of atoms or bonds with related function or process information (6).

To perform a polymer structure search using ISIS, the searcher uses ISIS/Draw to draw the desired structure, eg, a complete polymer or a fragment. Polymer structure queries may include any of the standard ISIS query properties, eg, atom lists, ring counts, substitution counts, or bond types. The searcher transfers the drawn structure to ISIS™/Base and initiates a search. Search modes are

1. Current Mol (looks for the specific polymer only)
2. Flexmatch (allows some flexibility in what is considered an “exact” match)
3. SSS (“substructure search”; looks for all polymers that include the query structure)
4. “Search by Form” (combines SSS with any other data provided in the query form, eg, physical properties).

MDL’s ISIS program has an extremely valuable search feature: during a search for any given polymer, use of MDL’s “flexmatch” program results in retrieval of all representations stored—both structure-based (regardless of “phase-shift”) and source-based structures. Thus, for PET, assuming that the seven SRU structures shown as Figures 2a–2g and the three source-based structures shown as Figures 3a–3c were registered in a database, a search for any one of the 10 structures would retrieve all 10. The CAS Registry File added this feature in September, 2002 (4).

The rest of this article discusses structure representation for the types of polymers most commonly encountered. First, the principal CAS and IUPAC guidelines governing regular, single-strand, organic polymers are presented. This is followed by an overview of more complex substances—aftertreated (post-treated), alternating, block, comb, cross-linked, dendritic, graft, hyperbranched, hyper-cross-linked, star, and star-block polymers. There is also a section that summarizes stereochemistry in polymers, followed by another that reviews inorganic and quasi-inorganic polymers, double-strand (ladder) polymers, siloxanes and silicones, and oligomers and telomers.
Regular, Single-Strand, Organic Polymers

The word “regular” in the heading is used here to mean “ordinary”, or, in the case of copolymers, random or statistical types with no special attributes such as “nonpolymeric spacer groups” (called junction units by IUPAC). As mentioned above, these are drawn usually as either structure-based, i.e., \(-(-A-)\_n-\), or source-based, i.e., \(-(A:B:\ldots)_x\).

**Structure-based versus Source-based Representation.** There are no absolute rules that specify which types of polymers shall be represented as structure-based, which as source-based, and which by both methods. IUPAC makes no definitive statement about whether a polymer shall be represented as source-based or structure-based; thus, polymers may be represented both ways, and many examples of this can be found in both IUPAC’s publications (1,6–13) and the CAS Registry File.

CAS states (2) that “specific polymers are named on the basis of the monomers from which they are formed and/or on the basis of their structure, as represented by an SRU.” Since original documents do not always provide sufficient structural information to allow generation of the SRU name, the method most frequently used for describing polymeric substances is by citation of the component monomers. A few commercial polymers, each of which accounts for a large number of index entries, are indexed only at the SRU-based systematic polymer name. Systematic (SRU) nomenclature for polymers has been adopted from the system developed by the Committee on Nomenclature of the Division of Polymer Chemistry of the American Chemical Society (ACS) (14). Note: the IUPAC recommendations (*Pure Appl. Chem.* 48, 373–385 (1976)) are in full agreement with CAS practice. Names derived by this system, in addition to monomer-based entries, are cited for polymers whose structural repeating units are well-documented or can confidently be assumed.

The “few commercial polymers” to which CAS refers (2,15) are the following:

1. nylon-6
2. nylon-6,6
3. PET
4. \(\alpha\)-hydro-\(\omega\)-hydroxypoly(oxy-1,2-ethanediyl) (polyethylene glycol)
5. \(\alpha\)-hydro-\(\omega\)-hydroxypoly[oxy[methyl-1,2-ethanediyl]] (polypropylene glycol)

In accordance with CAS’s policy (2,15), CAS assumes SRUs for such two-component condensation polymers as

1. polyamides derived from a dibasic acid and a diamine
2. polyesters derived from a dibasic acid and a dihydric alcohol
3. polycarbonates from carbonic acid and a dihydric alcohol.

This is not limited to common polymers; as long as the criteria are met (15), uncommon polymers are also registered with both structure-based and source-based representations.
In contrast, CAS traditionally represents the following polymer types by only one method. These examples are illustrative only; the list is not meant to be exhaustive.

(1) **Source-based:**

a. Acetylenic  
b. Acrylic and methacrylic  
c. Alkyd  
d. Ethylenic  
e. Polyamic acids, polyamides, polyesters, polyimides, polyurethanes, and combinations of these types that are prepared from more than two monomers  
f. Thermoset  
g. Vinyl  
h. Any other polymer for which the structure is unknown  

Polymers in these eight groups are usually assigned source-based names, but these groups do not necessarily correspond to specific CAS polymer classes. A CAS publication (16) cites 42 specific polymer classes; an up-to-date list may be viewed by entering CAS Registry File online and executing the command: expand A/PCT.

(2) **Structure-based:**

a. Double-strand (ladder)

To take one example, the CAS principles—for both registration and searching—determine that a polymer such as poly(oxy-1,2-ethanediyl) is registered and named only as structure a in Figure 4; registration of structure 4b is forbidden by ordering rules. Structure 4c is the mirror image of structure 4a and is equated to structure 4a in the process of registration and naming.

It might therefore be assumed that retrieval of structures 4b and/or 4c during a search denotes incorrect indexing. However, owing to the way in which CAS's Messenger software operates, a search for structure 4c also retrieves the polymer with structure 4a. Some displayed images appear with the head atom on the left; others are mirror image when compared with the name. Regardless of which way round the on-screen image appears, retrieval of the desired polymer was historically likely to fail unless CAS's principles were followed. This search problem has been largely solved by the introduction in September 2002 of the POLYLINK search tool (4).

A comprehensive review of structure-based and source-based representations for regular, single-strand, organic polymers has been published (17).

**Homopolymers.**

**Structure-based Method.** The CAS principles (2,14) and IUPAC (3) recommendations for structure-based polymers are as follows. The descending order
of priority of citation (and of structuring the SRU) is (1) heterocyclic rings, (2) acyclic heteroatoms in the order O, S, Se, Te, N, P, As, Sb, Bi, Si, Ge, Sn, Pb, B, Hg, (3) carbocyclic rings, (4) acyclic carbon chains. If substituents are present, otherwise identical parent radicals in the SRU are chosen by the principles, in turn, of maximum number, lowest locants, and earliest alphabetic order of substituents. The shortest path (smallest number of atoms) is taken from the most preferred multivalent radical to another occurrence of the same radical (if present) within the SRU, then to the next most preferred radical, and so on. For SRUs containing more than one heterocyclic ring, there are further rules for determining heterocyclic ring seniority (18).

The complete registration process may be summarized as follows:

(1) Step 1: The atom sequence that defines the SRU is identified.
(2) Step 2: The “head” atom or senior bivalent radical of the sequence is defined.
(3) Step 3: The SRU is oriented, i.e., a decision is made on which way to proceed along the atom sequence.
(4) Step 4: The SRU is named.

As an example, consider the 12-backbone-atom SRU sequence:

\[
\text{Row A} \quad \text{Row B} \\
\begin{align*}
\text{H} & \quad \text{H} & \quad \text{H} & \quad \text{H} \\
\text{6} & \quad \text{7} & \quad \text{8} & \quad \text{5} & \quad \text{3} & \quad \text{2} & \quad \text{1} \\
\text{8} & \quad \text{7} & \quad \text{6} & \quad \text{5} & \quad \text{4} & \quad \text{3} & \quad \text{2} & \quad \text{1} \\
\text{1} & \quad \text{2} & \quad \text{3} & \quad \text{4} & \quad \text{5} & \quad \text{6} & \quad \text{7} & \quad \text{8} \\
\end{align*}
\]

Determination of the head atom, and subunit citation order, is as follows: O is senior to S, and S is senior to N; therefore, one of the two oxygen atoms must be the head atom. The SRU must therefore be either row A, atoms 1 to 12, or row B, atoms 1 to 12. The row A sequence places the N atom in position 6, whereas the row B sequence places it in position 5; the latter is therefore correct, and the final SRU is \((\text{O---O---O---C--N--C--C---S--C--N--C--O---C---O---C---N---C---S---C---N---C---C---})_n\).

Essentially the same steps should be followed by searchers in order to conduct an effective search for structure-based polymers.

For polymers in SCION, a DuPont proprietary database, the same four basic steps are used for SRU identification and orientation. However, the rules that control these steps are different; this results in names and structures for SRUs that are sometimes different from SRUs as defined and oriented by CAS principles (19).

**Source-based Method.** For source-based representations, CAS structures homopolymers by an expression such as \((A)_n\), where A is a monomer. A typical CAS source-based homopolymer representation is shown as Figure 5.

IUPAC makes no recommendations for formal source-based, one-component homopolymer nomenclature or structural representation. However, IUPAC recognizes that for common idealized structural representations such as polyethylene, polypropylene, and polybutadiene, these semisystematic or trivial names are approved for use in scientific work. Where the semisystematic name is an obvious
source-based name, the polymer referred to is derived from the indicated source (see Appendix A in Ref. 3).

**Copolymers.**

Structure-based Method. CAS does not generally index unterminated copolymers of SRUs by the structure-based method. Formats such as poly(oxyethylene)/poly(oxytetramethylene) or poly[(oxyethylene)/ (oxytetramethylene)] are not used. Instead, a source-based representation is used (see Fig. 9 below).

IUPAC states (7) that a comprehensive structure-based system of naming copolymers would be desirable, but that such a system presupposes a knowledge of the structural identity of all the SRUs as well as their sequential arrangements within the polymer. They suggest that when this information is available, the structure-based nomenclature used for single-strand polymers (3,13) can be extended to copolymers. A separate publication (8) depicts many structure-based copolymer representations.

Source-based Method. The source-based structural representation of a copolymer is in the format (A·B···x, where A, B, etc, are monomers. Figures 6–8 illustrate the method.

Although poly(oxyethylene) and poly(oxytetramethylene) are represented as structure-based homopolymers, the poly(oxyethylene/oxytetramethylene) copolymer is represented as a source-based polymer (see Fig. 9).

CAS does not use words such as “random” or “statistical” to describe random or statistical copolymers; copolymers without special attributes such as “block” are assumed to be random.

A comprehensive IUPAC publication (8) discusses nomenclature for source-based copolymers.
Complex Polymers

Space limitations preclude more than an overview of these types: aftertreated (post-treated), alternating, block, comb and graft, cross-linked, dendritic, hyperbranched, hyper-cross-linked, star, and star-block. Comprehensive reviews of structure-based and source-based representations for them have been published (20,21).

Aftertreated (Post-treated) Polymers. CAS describes post-treated polymers as those polymers that are modified, either partially or completely, after the basic polymer backbone or skeleton has been formed. These modifications may result from a reaction or series of reactions of the main chain, end groups, and/or pendant functional groups of the polymer. Post-treatment is indicated by modification terms after the polymer name that describe the nature of the post-treatment. CAS divides post-treated polymers into two types: structurable and registrable; and nonregistrable (15). A third category, registrable and referred to as “nameable but not structurable” (15), is now provided with structures, and the backfile is being updated on an ongoing basis with structure-based access. (A description of the indexing enhancements is also available from CAS Customer Service and is offered on the world wide web (WWW) at http://www.cas.org/EO/enhanc.html.)

IUPAC discusses post-treated polymers in a 1994 publication (12), and several examples of structure-based representations are given. As of September 2002, IUPAC has made no recommendations regarding nomenclature or structural representation for post-treated source-based polymers.
Alternating (and Other Periodic) Polymers. Alternating and other periodic copolymers are viewed somewhat differently by CAS and IUPAC.

Neither “alternating” nor “periodic” is used as a modifier by CAS for alternating or periodic polymers named on the basis of their structure. The text descriptors “alternating” and “periodic” are not used with SRUs because any alternation or periodicity would be obvious in the structure and would not need to be differentiated from another similar structure by means of a text descriptor.

Registration by CAS of source-based representation of alternating copolymers with the term “alternating” as an integral part of the Registry record began with the 12CI (1987) (2). Alternating copolymers are distinguished in CA Registry File from random copolymers by including the term “alternating” in the Registry record of a given copolymer with its source-based name (see Fig. 10).

IUPAC recommends that an alternating polymer be named as a copolymer comprising two species of monomeric units (or monomers) distributed in alternating sequence (7). The arrangement ABABAB—or (AB)n thus represents an alternating copolymer. As far as possible, alternating and periodic polymers are treated as regular polymers (8). An alternating sequence arrangement of monomers is represented by (A-alt-B), and the corresponding polymer is named poly(A-alt-B).

For source-based representations, IUPAC recommends that alternating and periodic copolymer names contain the connectives -alt- and -per-, respectively.
Table 1 lists the source-based connectives recommended by IUPAC for alternating, periodic, and several other polymer types. According to the above IUPAC recommendations, the source-based name for PET and the trivial name nylon-6,6 represent alternating copolymers.

**Block Polymers.** This section discusses structure-based and source-based nomenclature for ordered, segmented, and unordered block polymers, and polymers with “nonpolymeric spacer units” or “junction units.” Star-block polymers are discussed under star polymers.

When CAS began to register polymers in 1967, no distinction was made between block and other copolymers. The information might have been part of the descriptive “index modification,” but it was not part of the Registry record. Registration of block copolymers with the term “block” as an integral part of the Registry record began with the 12CI (1987).

CAS makes no distinction among ordered, segmented, and unordered block polymers. For source-based representation, block polymers are distinguished from random polymers by indexing as copolymers at the monomer names. The term “block” is cited in a special modification after all other structural information (2).

IUPAC has published recommendations for structure-based and source-based representations of block polymers (7,8,12).
Comb and Graft Polymers. These two types of complex polymers are discussed together because there is overlapping technology, and it is not always possible to distinguish between them.

CAS does not recognize comb polymers as a type. Many (but not all) comb polymers are classified as graft polymers; for example, the comb polymer from...
Table 1. IUPAC Source-based Copolymer Classification

<table>
<thead>
<tr>
<th>Type</th>
<th>Connective</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unspecified</td>
<td>-co-</td>
<td>poly(A-co-B)</td>
</tr>
<tr>
<td>Statistical</td>
<td>-stat-</td>
<td>poly(A-stat-B)</td>
</tr>
<tr>
<td>Random</td>
<td>-ran-</td>
<td>poly(A-ran-B)</td>
</tr>
<tr>
<td>Alternating</td>
<td>-alt-</td>
<td>poly(A-alt-B)</td>
</tr>
<tr>
<td>Periodic</td>
<td>-per-</td>
<td>poly(A-per-B-per-C)</td>
</tr>
<tr>
<td>Block</td>
<td>-block-</td>
<td>polyA-block-polyB</td>
</tr>
<tr>
<td>Graft</td>
<td>-graft-</td>
<td>polyA-graft-polyB</td>
</tr>
</tbody>
</table>

Acrylonitrile, polyethylene glycol monomethacrylate, and styrene is classified as a graft polymer, whereas comb polymers from homopolymers of macromonomers are not.

Registration of graft copolymers with the term “graft” as an integral part of the Registry record began with the 12CI (1987). When CAS began to register source-based representations of polymers in 1967, no distinction was made between graft and other copolymers. The information might have been part of the descriptive “index modification,” i.e., the word “graft” might appear somewhere in the text, but it was not part of the Registry record. Comb and graft polymers comprising an SRU backbone and an SRU side chain are represented as graft polymers with source-based names (see Fig. 11).

Differentiation between the polymeric substrate and the applied monomer is not made. For example, whether 1-hexene is grafted onto 1-heptene homopolymer or vice versa, the preferred name is “1-heptene, polymer with 1-hexene, graft” with an additional name “1-hexene, polymer with 1-heptene, graft” (2). Thus, CAS makes no distinction between poly(A-graft-B) and poly(B-graft-A).

Application of the text descriptor 8:PM, GRAFT to graft polymers began with the 12CI (1987); earlier records are not so designated, but this is of minor importance to searchers because this descriptor is currently displayable but not searchable. The term “graft”, which is searchable, is also part of the name.

IUPAC has proposed that graft polymers be named in the same way as regular polymers (4) or irregular polymers (12) but with minor modifications. IUPAC also gives recommendations for irregular polymers in which polymer chains are attached to the main chain through a constitutional unit (linking unit) different from the constitutional unit of the side chain (8). Graphic representations of graft polymers have been published (8).

An IUPAC publication (11) discusses both structure-based and source-based representations of comb polymers; a comb polymer is described as a macromolecule comprising a main chain with trifunctional points from which linear side chains emanate. If the subchains between the branch points of the main chain and the terminal subchains of the main chain are identical with respect to constitution and degree of polymerization, and the side chains are identical with respect to constitution and degree of polymerization, the macromolecule is termed a regular comb macromolecule. If at least some of the branch points are of functionalities greater than three, the macromolecule may be termed a brush macromolecule.
Cross-linked Polymers. As the name indicates, these have two or more separate polymeric chains connected by a cross-link; it is implied that the cross-link is generally (1) relatively short (and not itself polymeric) and (2) at irregular intervals along the chains, but there seem to be no firm rules on these points. Ladder polymers comprising two linear strands cross-linked at regular intervals are discussed under Double-Strand (Ladder) Polymers.

The Index Guide (2) gives no information on how cross-linked polymers are indexed, because CAS treats them as copolymers. IUPAC has published no specific recommendations on nomenclature or structural representations for cross-linked structure-based polymers. For source-based representations of polymers, IUPAC defines a cross-link as a small region in a macromolecule from which at least four chains emanate, which is formed by reactions involving sites or groups on existing macromolecules or by interactions between existing macromolecules (22). The small region may be an atom, a group of atoms, or a number of branch points connected by bonds, groups of atoms, or oligomeric chains. In the majority of cases a cross-link is a covalent structure but the term is also used to describe sites of weaker chemical interactions, portions of crystallites, and even physical interactions and entanglements.

Cross-links, the branch units of star macromolecules and other junction units, are optionally specified by their source-based names after the name of the
macromolecule, with the Greek connective ν separated by hyphens. When the content of the cross-linking monomer is high, the macromolecule is treated as a copolymer molecule (11).

**Dendritic, Hyperbranched, Hyper-cross-linked, Star, and Star-Block Polymers.** These groups of polymers are discussed together because there is overlapping technology that makes it difficult to discuss them separately. In the last decade there has been a sharp rise in interest of these types of polymers.

Tomalia and co-workers (23) describe a dendrimer as having three distinguishing architectural features: an initiator core, interior layers (generations) composed of repeating units radially attached to the initiator core, and an exterior or surface of terminal functionality attached to the outermost generation.

These dendrimers differ from classical monomers or oligomers by their extraordinary symmetry, high branching, and maximized (telechelic) terminal functionality. Dendrimer synthesis is accomplished by a variety of strategies involving time-sequenced propagation techniques. The resulting dendrimers grow in a geometrically progressive fashion. Chemical bridging of these dendrimers (24) leads to Starburst™ [Dow Chemical Co. (25)] polymers. Tomalia and Durst have written a detailed account of Starburst/cascade dendrimers and hyperbranched structures (26).

Turner (27) considers hyperbranched polymers to be highly branched, non-cross-linked polymers prepared by the step-growth polymerization of A,B (or ABₓ) polymers. Thus, A₂B means that for every two type A reactive groups there is one type B reactive group, while AB₂ means that for every one type A reactive group there are two type B reactive groups. AB₂ and AB₃ (or A₂B and A₃B) are the commonest; ABₓ (or AₓB) types where x > 3 tend to be so sterically crowded that polymerization is difficult unless the reactive groups are well separated.

Voit (28) portrays the essential differences between dendrimers and hyperbranched polymers as follows:

1. Dendrimers are prepared by a stepwise approach; they are structurally perfect or nearly perfect (i.e., the branching ideality approaches 100%), their size is limited (i.e., they cannot grow forever), and their molecular weight is dependent upon the number of generations.
2. Hyperbranched polymers may be prepared by a one-step synthesis under conventional polycondensation conditions; their molecular weight is dependent upon the condensation reaction, not upon the number of generations, and the branching ideality is typically between 50 and 75%.

The terms “star polymer” and “star-shaped polymer” are used somewhat imprecisely. The same term is used to describe at least two physically different polymeric “clusters.” A star polymer may be a polymer comprising a precise center of known structure from which radiate a known number of arms. The arms, which may contain one or more SRUs or polymerized monomers, may be random, statistical, or blocky; polymers containing the latter are usually called star-block polymers.

A star polymer may also comprise an unknown number of arms connected to the center of the cluster by addition to the reaction mass of a second, multifunctional monomer or SRU that ties all the linear growing polymer chains together.
into a cross-linked (microgel) core. These are sometimes called star-branch polymers (29). Cloutet and co-workers (30) however, describe a star-branched polymer as a “dendritic block polymer.” A third type of star polymer comprises a dendritic core from which emanate a known number of arms (31). In any of these types of star polymers, the arms may be all the same, or different arms may have different compositions; the latter known variously as heterostar (32), miktoarm star (33), or variegated star polymers (11).

CAS handles dendrimers as regular organic compounds or polymers, depending on the degree of polymerization. The number of nonhydrogen atoms is also a factor. Low generation dendrimers having 10 or fewer repeating units and a nonhydrogen atom count less than 253 are structured and named systematically as specific nonpolymeric compounds. Higher generation dendrimers having more than 10 repeating units or a nonhydrogen atom count greater than 253 (which thus forces a manual registration at a lower degree of polymerization) are structured and named as polymers.

Polymers having dendritic or starburst structure are indexed at the controlled term DENDRITIC POLYMERS, which was added as a new index heading starting with the CA 13th Collective Index Period in 1992. Additional index entries are made for specific dendrimers. The approach for preparing divergent or convergent dendrimers may be indicated in the text modification when so specified by the author of the source document. [Convergent and divergent refer to the two basic methods of forming dendrimers; as the names imply, convergent dendrimers (typified by the work of Fréchet and co-workers (34)) are built “inward,” i.e., from the periphery toward the core, whereas divergent dendrimers (typified by the work of Tomalia and co-workers (23,24)) are formed by building outward from the core.] Currently CAS has no specific nomenclature for polymeric dendrimers; they are named according to regular polymer nomenclature rules. The polyfunctional core used in either the divergent or convergent approach is not included in the polymer registration. CAS’s policy on the indexing of dendrimers is described in more detail elsewhere (21).

IUPAC has addressed star and star-block polymer representations (8,11). As of September 2002, no specific recommendations on nomenclature or structure representation for dendrimers, hyperbranched, or hyper-cross-linked polymers have been published; an IUPAC document on this topic is close to completion. An IUPAC paper on source-based nomenclature (11) describes a star polymer as a macromolecule containing a single branch point from which linear chains (arms) emanate. In a regular star polymer all arms are identical with respect to constitution and degree of polymerization. A variegated star has different arms composed of different monomers.

Comprehensive reviews of nomenclature for structure-based and source-based representations for dendritic, hyperbranched, and hyper-cross-linked polymers (21,35–38), and for star and star-block polymers (21), have been published.

**Stereochemistry in Polymers**

CAS represents stereochemistry in polymers by text descriptor terms when the necessary information is reported; such terms include isotactic, syndiotactic,
three-diisotactic, erythro-diisotactic, and disyndiotactic (2). The term “atactic” (for a random configuration) is not employed by CAS in indexing specific polymers. In addition to the special terms above, polymer stereochemistry is defined, when appropriate, by the stereo descriptors E, Z, R, S, R*, and S*.

The first extensive IUPAC publications on stereochemistry in high polymers were published in the early 1960s, and subsequently in 1966 as a single article (39). In addition to more conventional polymer names, eg, polyethylidene and polypropylene, the “-amer” nomenclature was introduced in 1952 (10). IUPAC basic definitions relating to stereochemistry, eg, tacticity and isotactic polymer, were published in 1974 (1).

A major 1981 IUPAC document (9) addresses polymer stereochemistry in depth. Key chapters of this document discuss basic definitions (configurational unit, configurational base and repeating units, stereorepeating unit, different types of tacticity, tactic block polymers, and stereoblock polymers); sequences; conformations; and supplementary definitions. Most of the illustrations use the rotated Fischer projections, but some three-dimensional representations are included.

**Miscellaneous Polymer Types**

These include inorganic and quasi-inorganic polymers, double-strand (ladder) polymers, siloxanes and silicones, and oligomers and telomers. A comprehensive review of structure-based and source-based representations for these types of polymers has been published (40).

**Inorganic and Quasi-inorganic Polymers.** A polymer need not be completely inorganic to be classified as an inorganic polymer; an inorganic backbone is sufficient, even if there are organic moieties attached to a backbone atom.

A regular linear polymer that can be described by a preferred SRU in which only one terminal constituent subunit is connected through a single atom to another identical SRU or to an end group is a quasi-single-strand polymer, ie, it does not fit the definition of a regular single-strand polymer, but it can be named in the same manner (13).

CAS represents single-strand inorganic polymers by both structure-based and source-based methods. Figure 12 shows a typical structure-based record;

---

**Fig. 12.** Structure-based record; poly[imino(aminophosphinylidene)].
nitrogen is senior to phosphorus and might be expected to appear on the left but (as explained under Structure-based versus Source-based Representation) this may not always occur.

Source-based representations of polymers are typically recorded as shown in Figure 13.

Some references on phosphazenes (41) show the structures with nitrogen mostly on the left, although some are drawn with phosphorus on the left.

When SRUs are bridged only by metals, systematic polymer nomenclature is not used; instead, the substance is indexed either at the monomeric salt name or by coordination nomenclature, with a modification phrase, in either case, such as “homopolymer” or “polymer with...” (2).

IUPAC (13) presents a system of nomenclature for regular single-strand and quasi-single-strand inorganic and coordination polymers. The system is based on the selection and naming of an SRU defined as the smallest structural unit the repetition of which describes the polymer structure. The name of the polymer is the name of this repeating unit prefixed by the terms “poly,” “catena”, or other structural indicator, and designations for end groups, if desired. IUPAC has created sets of rules for identifying, orienting, and naming the preferred SRU, and subsequently naming

**Double-Strand (Ladder) Polymers.** Double-strand ladder and spiro polymers contain no open-chain bonds in the backbone. Polymers containing some open-chain backbone bonds (called partial ladder, imperfect ladder, step-ladder, or block ladder polymers) are not considered to be true ladder or spiro polymers.

The CAS and IUPAC nomenclature and structure representation systems are very similar. CAS's rules for ladder polymers are published in the *Index Guide* (2). IUPAC defines a ladder or spiro polymer as a double-strand polymer, the molecules of which are formed by an uninterrupted sequence of rings with adjacent rings having one atom in common (spiro polymer) or two or more atoms in common (ladder polymer). A comprehensive system is presented for identifying, orienting, and naming the preferred SRU, and subsequently naming
the polymer (42). The nomenclature and structure representation system used in SCION, a DuPont proprietary database, is novel and radically different (43).

Other prominent articles on double-strand polymers include those by Overberger and Moore (44), DeWinter (45), and Bailey (46); the latter reference contains a short “proposed nomenclature” section that is essentially a reprint of part of an ACS publication (14).

Siloxanes and Silicones. Although IUPAC classifies polysiloxanes as inorganic polymers (13), there is sufficient interest and diversity in them as a group for them to warrant discussion separate from other inorganic polymers.

Polysiloxanes occupy a virtually unique position in the world of polymers. Many polysiloxanes comprise a partly or wholly inorganic backbone built from alternating silicon and oxygen atoms; in contrast, pendent groups or side chains attached to silicon atoms are usually organic. Types discussed in this section include regular single-strand, hyperbranched, ladder, and spiro polysiloxanes.

Prior to 1994, polysiloxanes prepared by hydrolytic polymerization of alkoxysilanes or halosilanes were indexed in CAS's CA File at the index heading “Siloxanes and Silicones” with no polymer registration. Polysiloxanes described by authors only in terms of SRUs were indexed at this heading with no polymer registration. [A description of the indexing enhancements is also available from CAS Customer Service and is offered on the world wide web (WWW) at http://www.cas.org/EO/silox.html.]

Beginning with volume 121 (1994) of CAS Chemical Abstracts, structural representation and machine registration was begun for polymers previously handled only as text terms in the CA File. Backfile conversion is not planned.

This policy change affects the following areas (CAS's exact phraseology (including use of parentheses, vice brackets, for SRUs) is used below; in most cases, when words such as siloxanes or silsesquioxanes are cited, what is meant is polysiloxanes or polysilsesquioxanes):

- siloxanes or silsesquioxanes prepared by hydrolytic polymerization of alkoxysilanes or halosilanes
- siloxanes described only in terms of single-unit SRUs of the type $(-O-SiRR')_n$ (R, R' = monovalent radical)
- siloxanes described only in terms of multiunit SRUs of the type $(-O-SiR_2-)_{x}$, $(-O-SiR'-2-)_{y}$ or $(-R_2SiO-)_{z}(-R'SiO_{1.5})_{y}$ (R, R' = monovalent radical)
- ceramers (hybrid organic–inorganic networks) prepared by hydrolytic polymerization (sol–gel process) of tetraalkoxysilanes containing organic moieties, which may be trialkoxysilyl-terminated organic polymers
- end groups for siloxane SRUs
- copolymers containing siloxane components
- tradenames for siloxanes and silsesquioxanes
- silicone rubbers.

A more detailed discussion of siloxane nomenclature and structure representation (40) includes further information on the changes introduced by CAS (http://www.cas.org/EO/silox.html).

For linear polysiloxanes, an IUPAC document (13) states that if there is a choice for the central atom, the element occurring later in the general element
sequence table (ie, the periodic table) is the central atom. Order of precedence begins in the top, right-hand corner of the periodic table and proceeds (in vertical columns) toward the bottom, left-hand corner; thus, atom seniority is F, Cl, Br, I, At, O, S, Se, Te, Po, N, P, As, Sb, Bi, C, Si, ..., Li, Na, K, Rb, Cs, Fr, He, Ne, Ar, Kr, Xe, Rn.

The linear polysiloxane (diphenylsilylene)oxy is represented by the structure shown as Figure 14. The IUPAC name also differs appreciably from the CAS name. According to the rules for linear organic polymers, this polymer would be oriented and named poly(oxy(diphenylsilylene)) (13).

Lichtenwalner and Sprung (48) present some IUPAC nomenclature for polysiloxanes; most of their examples show silicon on the left. Other papers on polysiloxanes in a 1988 symposium publication (41) show the structures with silicon mostly on the left, although some are drawn with oxygen on the left.

Oligomers and Telomers. Qualitatively, oligomers are “short polymers”, and telomers are “short polymers with specific end groups”.

The terms “telogen” and “telomerization” were defined many years ago to describe an innovative method for the free-radical polymerization of ethylene (49,50). The essence of telomerization, described in a 1942 patent (50), involves reaction of a telogen XY (also called a chain-transfer agent) with an ethylenic molecule to give a telomer of structure X–(A)n–Y. In the product telomer, fragments of the chain-transfer agent are attached to the termini of the molecule (49,50). Figure 15 illustrates the method.

CAS indexes oligomers of definite structures as specific compounds by the regular principles of organic nomenclature. When the precise structure is not known but the number of units is specific, “dimer”, “trimer”, etc, is expressed in the modification at the name of the monomer. The term “oligomeric” may be cited if this aspect is stressed in the original document (2). “Oligomeric,” if applicable, is expressed in the text modification part of the index entry and is not part of the CAS Registry record, except for some records marked with an asterisk. [CAS Registry numbers appended with an asterisk (*) represent substances of unknown or variable composition, complex reaction products, or biological substances that do not meet the criteria for CA indexing. These substances generally have been registered for regulatory agencies and receive a CA Index Name. They have a special code of CTS or GRS in the Class Identifier (CI) field. The information in the Locator (LC) field and/or Source of Registration (SR) field provides an indication of the reason for registration.]

Telomers of unknown structure are indexed similarly to copolymers with the term “telomer with...” cited in the modification (see Fig. 15). Telomers with a definite structure are named systematically (2) rather than as telomers per se (see Fig. 16).
Chemical reaction:
\[ \text{CCl}_4 + n \ H_2C=CH_2 \rightarrow \text{Cl}_3C\left[\text{CH}_2=\text{CH}_2\right]_n\text{Cl} \]

Abbreviated CAS Registry record:
RN 25398-31-6
CN Ethene, telomer with tetrachloromethane
MF (C2H4)x:CCl4
Structure: (CH2=CH2)x:CCl4

Fig. 15. CCl4 added as chain-transfer agent in ethylene polymerization.

Fig. 16. Oligomeric poly(ethylene) telomer with CCl4. The term “oligomeric” is part of the CAS Registry record name only in records marked with an asterisk (see text for details).

IUPAC defines an oligomer as “a substance composed of molecules containing a few of one or more species of atoms or groups of atoms (constitutional units) repetitively linked to each other for which the physical properties vary with the addition or removal of one or a few constitutional units from its molecules” (1). IUPAC defines “telomer” and “telomerization” essentially as follows (22):

1. **Telomer**: an oligomer obtained by a telomerization process
2. **Telomerization**: an oligomerization by chain reaction carried out in the presence of a large amount of chain-transfer agent, so that the end groups are essentially fragments of the chain-transfer agent.

No recommendations have been published on how oligomers and telomers are to be named and represented graphically by either structure-based or source-based methods.

IUPAC references to oligomers and telomers are rare, although examples of a telomer (51) and an oligomer have been published. (see Section 4.3, Example 4.3-E1, in Ref. 8). IUPAC uses subscript \( a \), *vice* subscript \( n \), to indicate oligomers.

**Conclusions**

Over a period of about 50 years, representation of polymer structures, both on paper and in databases, has developed from a virtually random system to a highly organized and sophisticated one. Polymers are represented sometimes by structure-based methods, sometimes by source-based methods, and sometimes by both. Both methods survive because each offers advantages and disadvantages. Both methods involve structural representation of polymers by a precisely defined set of rules developed over several decades by CAS, IUPAC, and the Committee on Nomenclature of the Division of Polymer Chemistry of the ACS. Areas still in need of improved representation are copolymers of unterminated SRU types: aften-treated (post-treated) polymers, and dendritic (52), hyperbranched, hyper-cross-linked, star, and star-block polymers. Also needed are hierarchical relationships between intellectually related polymers (53).
Abbreviations

**ACS**: American Chemical Society  
**CA**: Chemical Abstracts  
**CRU**: Constitutional repeating unit  
**IUPAC**: International Union of Pure and Applied Chemistry  
**MDL**: MDL Information Systems, Inc.  
**RN**: Registry number  
**SRU**: Structural repeating unit  
**SSS**: Substructure search  
**STN**: Scientific and Technical Information Network  
**8CI**: CA Eighth Collective Index  
**9CI**: CA Ninth Collective Index  
**12CI**: CA Twelfth Collective Index

BIBLIOGRAPHY

5. CAS 2002 Index Guide, ACS, Washington, D.C., 2002; Appendix II: Section 14A.
15. CAS, Columbus, Ohio, *Searching for Polymer Information in CAS ONLINE*, Mar. 1990 (Publication C1EJ-0390).
47. IUPAC, Pure Appl. Chem. 61, 243–254 (1989); reprinted as Chapt. 8 in Ref. 52.

JOHN L. SCHULTZ
EDWARD S. WILKS
E. I. Du Pont de Nemours

**STYRENE–BUTADIENE COPOLYMERS.** See Volume 4.

**STYRENE POLYMERS.** See Volume 4.

**SULFUR–CONTAINING POLYMERS.** See Volume 4.