

## The Nomenclature of Highly Fluorinated Molecules\*

JOHN A. YOUNG†

Denver Research Institute, University of Denver, Denver, Colorado

Received March 24, 1967

A committee of the Division of Fluorine Chemistry has been attempting to improve accuracy, uniformity, and convenience of nomenclature in this field, and some of their recommendations are briefly presented. Among the problem areas are (1) accommodation of both organic and inorganic nomenclature, (2) limitations on usage of the term "perfluoro" (should it include fluorine in functional groups or attached to atoms other than carbon?), (3) compounds with approximately equal numbers of hydrogen and fluorine atoms, (4) fluorinated analogs of compounds conventionally named as hydrogen addition products of aromatic ring systems, such as dihydronaphthalene, and (5) structures which have no equivalents among nonfluorinated compounds ( $-\text{SF}_5$ ,  $-\text{ONF}_2$ ).

The modern era of fluorine chemistry was born about 1940, when several critical events occurred within a few years: J. H. Simons passed a few volts of electricity through a crude electrochemical cell made of a piece of iron pipe, R. J. Plunkett cut open a cylinder of tetrafluoroethylene which had mysteriously lost pressure, and L. A. Bigelow obtained  $\text{CF}_3\text{COCF}_3$  by direct fluorination. In the best American tradition, the child of these crude origins has developed scientific vigor and economic muscle; Teflon-coated frying pans are on sale at your friendly neighborhood drugstore and Scotchgard-treated clothes at the nearest shopping center.

Prior to World War II most fluorochemical syntheses consisted of a tedious series of consecutive halogenations, dehydrohalogenations, and halogen exchange steps, known familiarly as the "put-and-take" method, and functional compounds containing no hydrogen external to the functional group were rare; fully fluorinated aldehydes and ketones remained virtually unknown as late as 1954. Today, however, much of the vast array of organic compounds can be duplicated in fluorine chemistry, although the synthetic routes employed might seem bizarre compared with conventional organic syntheses. Moreover, the peculiarities of the fluorine atom have engendered many molecular types which have no previous analogs in classical organic or inorganic chemistry. It is this expanding universe of fluorine compounds, both in number and in kind, that necessitates a thorough review of nomenclature in fluorine chemistry at the present time.

Fluorine chemistry does not now face the same problems that vex biochemical, polymer, organic, and inorganic nomenclature. In all these, the most serious difficulties

arise in naming agglomerations of smaller units, often with a repetitive pattern, whereas we are chiefly concerned with naming the smaller units themselves, rather than any polymer structure. Our main goal is merely a concise, convenient, unambiguous, and consistent method for denoting large numbers of fluorine atoms in relatively simple molecules. This is imperfectly done at present, as the illustrations will show. We must also walk a tight-rope between the separate poles of organic and inorganic nomenclature, because fluorine chemistry overlaps both of these fields, and we must try to find an accommodation suitable to both until, in some future Utopia, a unified system applicable to all elements becomes a reality.

In order to document the need for revision, some examples of present nomenclature are discussed below, followed by our recommendations for improvement.

Illustrative examples can be found even in the simplest compounds, the short chain alkanes. Because the location of the fluorine atoms is unambiguous in  $\text{CF}_3\text{CF}_2\text{CF}_2\text{CF}_3$ , locants need not be used. As soon as one hydrogen is present, however, all the fluorine atoms, no matter how many there are, must be numbered. The result was that prior to 1952 a compound as simple as  $\text{HCF}_2\text{CF}_2\text{CF}_2\text{CF}_2\text{CF}_3$  required 41 characters in the name 1,1,1,2,2,3,3,4,4,5,5-undecafluoropentane. Some relief was provided in 1952 when a convention was adopted permitting the use of "perfluoro" in completely fluorinated hydrocarbons and in those in which the number of hydrogen atoms was not greater than four. By this means,  $\text{CF}_3\text{CF}_2\text{CF}_2\text{CF}_2\text{CF}_3$  could be called perfluoropentane rather than dodecafluoropentane and  $\text{HCF}_2\text{CF}_2\text{CF}_2\text{CF}_2\text{CF}_3$  called 1H-perfluoropentane rather than 1,1,1,2,2,3,3,4,4,5,5-undecafluoropentane, the latter change resulting in a saving of over 50% of linear space and encouraging the use of proper nomenclature in oral presentation. Unfortunately, the convention was prohibited for compounds containing functional groups or even other halo-

\* Presented before the Division of Chemical Literature, Symposium on Chemical Nomenclature, 152nd National Meeting of the American Chemical Society, New York, N. Y., Sept. 15, 1966.

† Chairman, Nomenclature Committee, Division of Fluorine Chemistry, American Chemical Society.

gens, so that  $\text{HCF}_2\text{CF}_2\text{CF}_2\text{CF}_2\text{CF}_2\text{Cl}$  reverted to 1-chloro-1,1,2,2,3,3,4,4,5,5-decafluoropentane,  $\text{CF}_3\text{CF}_2\text{CF}_2\text{CF}_2\text{CF}_2\text{CF}_2\text{COOH}$  to pentadecafluorooctanoic acid, and  $\text{H}(\text{CF}_2)_7\text{COOH}$  to 2,2,3,3,4,4,5,5,6,6,7,7,8,8-tetradecafluorooctanoic acid. This limitation seems unduly restrictive in present day fluorine chemistry, and our primary objective is the extension of "perfluoro," or some other term conveying the same meaning, to functional compounds.

Although the use of an upper case "H" is officially sanctioned and is quite convenient for denoting a few hydrogen atoms in an otherwise fully fluorinated structure, this use clearly conflicts with its use to designate "indicated" hydrogen atoms in certain ring systems. Conceivably, a compound could be called 2H-perfluoro-2H-azepine, in which case identical terms would have different meanings. Choice of a more suitable term or symbol to describe hydrogen in a "perfluoro" structure would avoid this conflict.

All chemists who deal with highly fluorinated structures use the term "perfluoro" to some extent but they do not always mean the same thing by it. In the days when no more than a handful of N-F compounds were known and the only "perfluoroamines" of any commercial importance were tertiary, in which case no ambiguity can arise, this was not a serious problem. Over the last 10 years improvements in direct fluorination and the availability of dinitrogen tetrafluoride have resulted in a host of known N-F compounds, but "perfluorodimethylamine" can still mean either  $(\text{CF}_3)_2\text{NH}$  or  $(\text{CF}_3)_2\text{NF}$ , depending on the individual preference and geographical location of the user. In organic chemistry this problem is most acute with nitrogen compounds—no one is likely to call  $\text{COF}_2$  "perfluoroformaldehyde"—but it is common to all compounds in which hydrogen is part of a functional group or is attached to a hetero atom; in the inorganic area, where carbon is not king, the problem is a general one. "Perfluoro" may properly include fluorine attached to atoms other than carbon or it may not, but the usage must be consistent.

Aromatic compounds usually have a specific name for the ring with the maximum number of noncumulative double bonds—*i.e.*, benzene and not cyclohexatriene—whereas alicyclic compounds are usually named on the basis of the saturated acyclic compound with the same chain of atoms—*i.e.*, cyclohexane and not hexahydrobenzene. Compounds of intermediate states of unsaturation are: triazine, dihydrotriazine, tetrahydrotriazine, hexahydrotriazine, but pyrrole, pyrroline, pyrrolidine. When the term "hydro" is used to describe a state of intermediate ring unsaturation, use of the term "perfluoro" can create ambiguity. In Figure 1 the first compound is indubitably perfluoronaphthalene. The second, if "hydro" has its usual additive connotation, should be

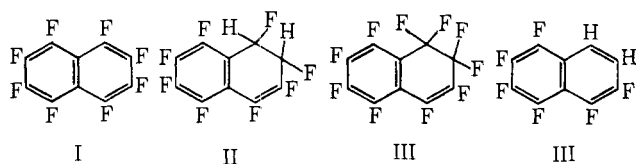


Figure 1.

dihydroperfluoronaphthalene. The third, if "hydro" is taken to describe the state of unsaturation, should be perfluorodihydronaphthalene. Here we have the same two terms whose exact meaning depends on their order of appearance. (The difficulty could be minimized by the use of parentheses, but parentheses are often inadvertently omitted.) This type of ambiguity can arise whenever a ring name contains the term "hydro." In addition, the third example names hydrogen although it is not in fact present, a situation which is fundamentally poor nomenclature. The use of "perfluoro" in cyclic systems obviously raises problems.

There are compounds in the fluorine domain which are unstable or unknown in conventional organic or inorganic chemistry. The small size and unequalled electronegativity of fluorine atoms stabilize a wide variety of unusual structures—for example, the O-F and N-F compounds—and are responsible for covalences which are not attainable with hydrogen. The last effect is especially remarkable with sulfur; although neutral alkyl derivatives of  $\text{SH}_4$  and  $\text{SH}_6$  are unknown, compounds such as  $\text{CF}_3\text{SF}_5$  or  $\text{CF}_3\text{SF}_4\text{CF}_3$  are fairly common and  $\text{R}_F\text{SF}_3$  or  $\text{R}_F\text{SF}_2\text{R}_F$  are also known, as well as many inorganic derivatives of  $\text{SF}_6$ . Few of these can be named by conventional means since there is no term for tetravalent or hexavalent sulfur in organic nomenclature, while inorganic nomenclature becomes very awkward for all but the simplest molecules. Take, for instance, the known and relatively simple compound  $\text{CF}_3\text{SF}_4\text{SF}_5\text{CF}_3$ . An inorganic chemist might name this as a derivative of  $\text{S}_2\text{F}_{10}$ ,  $\text{S}, \text{S}'$ -bis(trifluoromethyl)disulfur decafluoride; with inorganic nomenclature now under consideration it might be called 1,1,1,2,2,2,3,3,3,4,4,4-tetradecafluoro-2,3-dithia(VI)-butane. An organic chemist who preferred to exclude hetero atoms under the term "perfluoro" would be forced to call the compound 2,2,2,2,3,3,3,3-octafluoroperfluoro-2,3-dithiabutane, while one preferring to include hetero atoms and a little cavalier about the valence state of sulfur might try perfluorodimethyl disulfide. None of these four names is adequate; the two systematic ones are excessively long for a simple and symmetrical molecule, while the two more nearly trivial names are not sufficiently capable of extension to complex molecules.

The critical need in fluorine nomenclature is a term, syllable, or symbol which would signify that all positions capable of substitution in the parent named chain are occupied by fluorine atoms unless otherwise described. This is exactly what is done with hydrogen in organic nomenclature: the total number of hydrogen atoms is not stated but must be found by satisfying all the conditions of atomic identity, valence state, and structure stipulated in the name, and any atom whose identity is neither specifically stated nor implied in any part of the name is assumed to be a hydrogen atom. We are merely adopting the same convention with regard to fluorine. As the parent "perfluoro" compound, by definition, contains no hydrogen on the named chain, any hydrogen atoms present are not residual but must be named as substituents, and the total number of fluorine atoms is not stated but results from satisfying the stipulated conditions of atomic identity, valence, and structure. Once this transformation has been accomplished, the compound can be named in detail by conventional methods as far as these apply.

The term in present use which most closely approximates this function is "perfluoro." "Perfluoro" is itself not perfect for the purpose since the prefix "per" already has several other distinct and mutually incompatible meanings. "Perfluoro" is presently used with both inclusion and exclusion of hetero atoms, and its nine characters make it longer than desirable; on the other hand it is currently in wide use and has some official sanction. Adoption of an entirely new term might meet some consumer resistance, but would have the great advantage that, like Humpty-Dumpty in Alice, we could make the word mean just what we choose it to mean. It is likely that a new prefix will be coined, a short syllable, suggestive of fluorine by use of the same initial consonant—for instance, "flo". Final decision on this question has not been reached, however, and in this paper the term "perfluoro" will be used to illustrate the proposed methodology since this term is certainly more familiar.

The term "perfluoro" (or whatever new term is adopted) substitutes tacit omnipresence of fluorine for omnipresence of hydrogen; all other atoms are regarded as substituents and must be named. Since this system eliminates almost all locants for fluorine atoms, a great improvement in brevity and oral convenience is achieved;  $\text{CF}_3\text{CF}_2\text{CF}_2\text{CF}(\text{NO})\text{CF}_2\text{Cl}$  is 1-chloro-2-nitrosoperfluoropentane rather than 1-chloro-1,1,2,3,3,4,4,5,5,5-decafluoro-2-nitrosopentane, and  $\text{CF}_3(\text{CF}_2)_7\text{COOH}$  is perfluorooctanoic acid rather than pentadecafluorooctanoic acid. Furthermore, the concept of total fluorination is much easier to grasp than when the number of fluorine atoms must be mentally added up and compared with the unstated number of positions available for substitution. It is to be noted that the term "perfluoro" immediately precedes the group which it modifies and does not apply to hydrogen attached to oxygen.

Hydrogen, like any other substituent, must be named. As was mentioned previously, use of an upper case "H" for this purpose conflicts with use of the same symbol to designate "indicated" hydrogen. The term "hydro" is now used fairly widely, as in  $\alpha,\alpha,\omega$ -trihydroperfluoroalkanol,  $\text{H}(\text{CF}_2)_n\text{CH}_2\text{OH}$ , but this is poor practice as the term "hydro" has always had an additive connotation (dihydronaphthalene, etc.) and has also been used in a quite unrelated context as referring to water (hydrolysis, hydronium, etc.). We have invented the term "hydriyl" specifically to denote substitutive hydrogen in the "perfluoro" system. Thus,  $\text{HCF}_2\text{CF}_2\text{CF}_2\text{CF}_2\text{CF}_2\text{NO}$  becomes 1-hydriyl-5-nitrosoperfluoropentane and  $\text{H}(\text{CF}_2)_4\text{CH}_2\text{OH}$  becomes 1,1,5-trihydriylperfluoro-1-pentanol. The new term is unambiguous and does not interfere with the use of either "H" or "hydro" for any other presently accepted purposes. In Figure 2 the first compound is perfluoro(2H-azepine); since the "2H" is required to describe fully the arrangement of double bonds it is

retained as part of the parent name and is not changed to "2F" or "2-fluoro." The second compound is 2-hydriylperfluoro(2H-azepine), thus avoiding any possibilities of confusion due to different uses of the upper case "H."

The problem of partially saturated rings has been solved by proscribing all ring names which include the term "hydro." In any such compounds, reversion is made to the parent ring name with the maximum number of non-cumulative double bonds, and all compounds derived from that ring are named on this basis, regardless of the state of saturation. In Figure 1, the first compound remains perfluoronaphthalene. The second is unequivocally 1,2-dihydroperfluoronaphthalene, the third is 1,2-difluoroperfluoronaphthalene, and the fourth is 1,2-dihydriylperfluoronaphthalene. The ambiguity of "perfluorodihydro" and "dihydroperfluoro" is thus avoided. There seems no pressing reason to invent a new prefix to describe additive fluorine, as in the third compound; a name ought to describe present structure, not past history, and the structure is quite clear from the name given. Reversion to a parent ring is not necessary when an available name describes the specific state of unsaturation; any ring system accepted in the IUPAC 1957 report, whether trivial, Hantzsch-Widman, or replacement nomenclature, is compatible with the "perfluoro" system.

The thorny problem of naming inorganic fluorine compounds is made more difficult by the present unsettled state of inorganic nomenclature. It is apparent that inorganic nomenclature is becoming more systematic, in the sense of using root terms, prefixes, and suffixes to describe: (1) the chain length and state of saturation; (2) the identities of the component atoms; and (3) the valence states of those atoms. This is essentially what replacement nomenclature, or the oxa-aza system does in organic nomenclature. For instance, *s*-triazine is 1,3,5-triaza-1,3,5-cyclohexatriene. A new inorganic system employing a melange of replacement nomenclature and Stock numbers is now under consideration; as mentioned above,  $\text{CF}_3\text{SF}_4\text{SF}_4\text{CF}_3$  would be named, under this system, 1,1,1,2,2,2,2,3,3,3,3,4,4,4-tetradecafluoro-2,3-dithia(VI)-butane, which could be shortened by the use of "perfluoro" to perfluoro-2,3-dithia(VI)butane.

We would like to see adopted the even more highly systematized recommendations of the subcommittee on phosphorus nomenclature, which are in a sense an extension of replacement nomenclature to include a term describing the valence state of an element in the chain. This term is based on the number of unshared electron pairs remaining on the atom—"ane" for no unshared pairs, "ine" for one unshared pair, "uane" for two pairs, and "uine" for three pairs—and would be applicable to all atoms. According to this system,  $\text{CF}_3\text{SF}_4\text{SF}_4\text{CF}_3$  would be named perfluoro-2,3-disulfabutane ("thia" is restricted to  $-\text{S}-$ ),  $\text{CF}_3\text{ON}=\text{NCF}_2\text{N}=\text{SF}_2$  would be perfluoro-6-oxa-1-sulfina-2,4,5-triaza-1,4-heptadiene, and  $\text{CF}_3\text{SF}_4\text{NFNFSF}_3$  would be perfluoro-1,4-disulfa-2,3-diazapentane. Such an extension of replacement nomenclature would parallel current trends in organic nomenclature, would tend to unify organic and inorganic nomenclature, and would lend itself admirably to the "perfluoro" system for fluorine compounds. Even with present nomenclature, however, the "perfluoro" system can eliminate the necessity of repeated locants for fluorine atoms.



Figure 2.

At present, two problems must be resolved before official approval of the "perfluoro" system can be sought. The first of these is the term itself: whether to retain "perfluoro" or to invent a new term to express the same intent. For the reasons given earlier the writer favors a new term. The second problem is the inclusion or exclusion, under the meaning of "perfluoro," of hydrogen attached to atoms other than carbon. If "perfluoro" itself were retained and extended unrestrictedly to hetero atoms, names for many organic compounds of oxygen, sulfur, and nitrogen would be completely unacceptable to organic chemists—for instance, perfluoroethanol for  $\text{CF}_3\text{CF}_2\text{OF}$ . In order to retain the traditional functional names which are so universally used in organic chemistry (alcohol, amine, mercaptan) fluorine attached to hetero atoms would have to be excluded or severely restricted. The use of "perfluoro" to cover fluorine attached to *all* atoms is mandatory in the inorganic realm, however, if lengthy repetition of locants is to be avoided. It would, therefore, be necessary to maintain two meanings of "perfluoro," each depending on type of compound or system of nomenclature, and such a situation would be sure to cause confusion and ambiguity. Furthermore, a new system for organic compounds which was not equally applicable to inorganic compounds would have at best only temporary utility,

as the distinction between organic and inorganic chemistry continues to blur with passing time. A new term for "perfluoro" would carry no traditional connotations and might, by virtue of its strangeness, alert the organic chemist to unconventional properties of the functional groups named subsequently. The best course seems to be the employment of a new syllable which would include, in the sense of "perfluoro," fluorine attached to all atoms, not only carbon, with special provision made to protect the integrity of a very few groups, possibly  $-\text{CHO}$ ,  $-\text{COOH}$ , and those in which hydrogen is attached to oxygen.

This account is not meant to be a comprehensive description of the anticipated "perfluoro" system, but merely presents its main principles and intentions. We fully recognize that many familiar names are too deeply entrenched to disappear overnight, even in the event of official adoption of the system, and that some of these names, particularly with very simple molecules, are completely acceptable. The system should be most valuable in furnishing a standardized, consistent nomenclature and in naming highly fluorinated molecules of unusual types; its application to simple molecules and to those which contain relatively large numbers of hydrogen atoms is neither necessary nor desirable.

## Low-Cost Storage and Retrieval of Organic Structures By Permuted Line Notations: Small Collections\*

J. K. HORNER

Stanford Research Institute, Menlo Park, California 94025

Received October 20, 1966

**A low-cost storage and retrieval system is described for a small collection of chemical structures. About 3000 structures from the Stanford Research Institute files were coded according to the Wiswesser Line Notation, a computer program was written, and a permuted line-notation index was generated, all for approximately \$700.**

Earlier reports have shown that permuted chemical line notations in tabulated lists can be used to rapidly locate specific compounds, classes of compounds having similar ring systems, and compounds having the same functional group (1-4). Time and cost data on computer-generated indexes are available for large collections of chemical structures (3) and time data on key punch-generated indexes for small collections are available (2). The object of this communication is to report the experiences of the Life Sciences Research Area of Stanford Research Institute in collecting and storing at a low cost a small collection of chemical structures.

\* Presented at the Wiswesser Line Notational Meeting of the Army Chemical Information and Data Systems (CIDS) Program at US Army Edgewood Arsenal, Edgewood Arsenal, Md., October 6 and 7, 1966.

### BACKGROUND

In 1963 it became evident to scientists at Stanford Research Institute (SRI) that a central file of data on organic structures was needed to replace the antiquated molecular formula files kept in each section. An organic chemist with neither background nor training in data storage and retrieval was assigned the task of assessing the current methods of chemical data storage and of adapting one to fit the needs of SRI.

Definitive information in this field proved scanty until the publication in 1964 of the "Survey of Chemical Notation Systems," (5) a report of the Committee on Modern Methods of Handling Chemical Information. This comprehensive report plus personal contacts with industrial