ELSEVIER

Journal of Fluorine Chemistry 100 (1999) 7–11

JOURNAL OF FLUORINE CHEMISTRY

www.elsevier.com/locate/jfluchem

Fluorine chemistry statistics: numbers of organofluorine compounds and publications associated with fluorine chemistry

Helen Schofield*

Chemistry Department and Library, UMIST, PO Box 88, Sackville Street, Manchester M60 1QD, UK

Abstract

Data are presented which demonstrate the considerable growth in the numbers of new organofluorine compounds produced annually, and also the numbers of papers and patents published concerned with fluorine chemistry. An overview of some information sources of relevance to fluorine chemists is also presented.

Keywords: Number of organofluorine compounds; Information retrieval; Databases; Growth in literature

1. Introduction

The overall growth in the chemical literature and the number of new compounds synthesised or isolated annually has been well documented and is recorded in a number of guides to the literature of chemistry [1,2]. More specifically, trends in the growth in the area of organofluorine compounds and papers on the subject during the period 1967–1990 have been reported by Wilkinson [3] and Seebach [4]. Their data are reproduced in Figs. 1–3. Wilkinson asserted that 6.2% of compounds registered in *Chemical Abstracts* (CA) at the time of his review possessed a C–F bond, demonstrating the intense activity in the area of organofluorine chemistry, although the data of Seebach [4] suggest that the percentage of documents discussing such compounds is rather less. The objectives of this study are to update the data of Wilkinson and Seebach, and to provide an overview of sources from which information on fluorine chemistry can be obtained.

2. Growth in the number of compounds since 1980

Nowadays, there are many sophisticated search tools available which enable data on the numbers of compounds and papers to be derived. For this work, the *Chemical Abstracts Service* (CAS) Registry file, as made available by STN International [5], was searched to obtain the number of new compounds registered each year by CAS. This was done by performing a sub-structure search for compounds containing a C–F bond for each year from 1989–1998, and

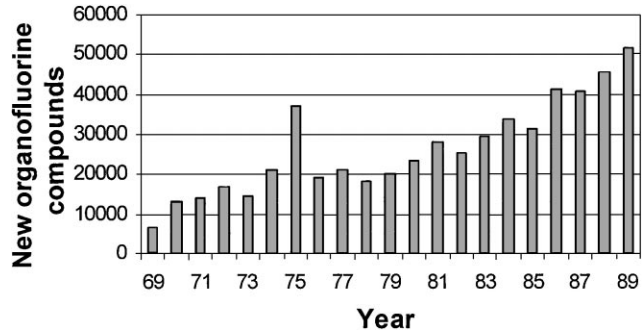

Fig. 1. Number of new organofluorine compounds registered by CAS each year from 1969–1989. Values taken from Wilkinson [3], from data prepared by David Walsh of STN International.

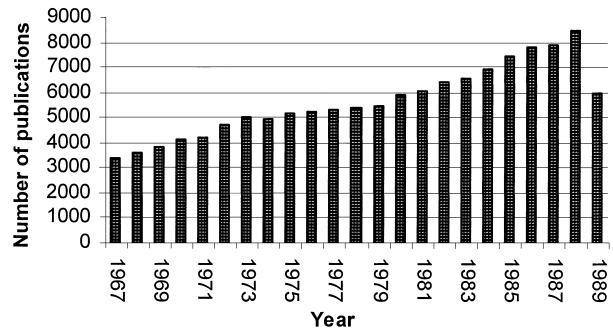

Fig. 2. Number of publications indexed under "fluorine" in *Chemical Abstracts*, from 1967–1989 (data taken from Seebach) [4].

the data are recorded in Fig. 4. This graph shows that, with minor fluctuations, the number of new compounds recorded each year has continued to increase over the time period

E-mail address: helen.schofield@umist.ac.uk (H. Schofield)

0022-1139/99/$ – see front matter
PII: S0022-1139(99)00147-5

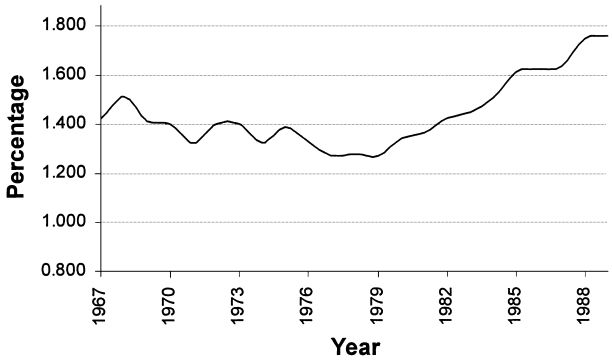

Fig. 3. The number of papers in *Chemical Abstracts* from 1967–1989 indexed under ''fluorine'' as a percentage of the total number of papers. Data taken from Seebach [4].

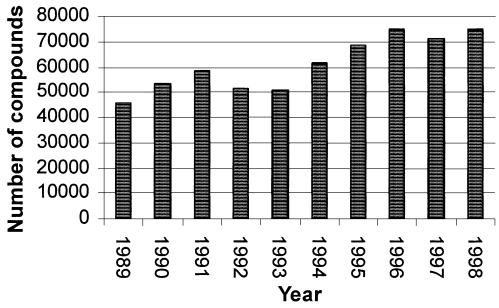

Fig. 4. Number of new compounds containing at least one C–F bond registered by CAS in the period 1989–1998. In total 610 873 C–F containing compounds were registered by CAS during this period.

considered. The data have not been plotted alongside those obtained on behalf of Wilkinson [3], because a check of the data for 1989 revealed a small discrepancy between the results which may be due to a difference in the way in which the searches were carried out. It should also be pointed out that the figures for 1989, 1990 and 1991 might be artificially high because during that period CAS was registering compounds produced in the period 1957–1966 which were not discussed in the literature between then and 1984, distorting the figures by up to 7%. The total number of new compounds containing a C–F bond registered by CAS between 1989 and 1998 is 610 873.

3. Growth in the number of publications since 1980

The growth in publications between 1967 and 1989 is shown in Fig. 2 [2]. The data recorded here are the numbers of publications indexed under ''fluorine'' in *Chemical Abstracts*. Wilkinson [1] also presented a graph which shows a similar trend for the increase in the number of publications concerning organofluorine compounds, although the means by which the data were derived is not given; the numbers were somewhat higher than those recorded by Seebach [4]. Due to the differences in the numbers put forward by Wilkinson and Seebach, the figures have been recalculated across the whole available time period, i.e. 1967–1998, and are recorded in Fig. 5. The data were derived from a search of the CA file produced by CAS and made available through STN International. The data were obtained by searching for the terms ?fluoro?, ?fluorin? or ?fluoride? amongst the title, index terms or words in the abstracts, where ? is a truncation symbol allowing additional characters to precede or follow the search term. The documents retrieved, therefore, are concerned with any aspect of fluorine chemistry or any fluorine-containing compound and discussed in the literature during that particular year, not just new compounds. The documents are broken down into patents and other documents for comparison purposes. In recent years, patents consistently make up about 30% of the total number of publications. This percentage is higher now than during the 1960s and 1970s, probably in part due to an increase in the number of patent offices covered by CAS

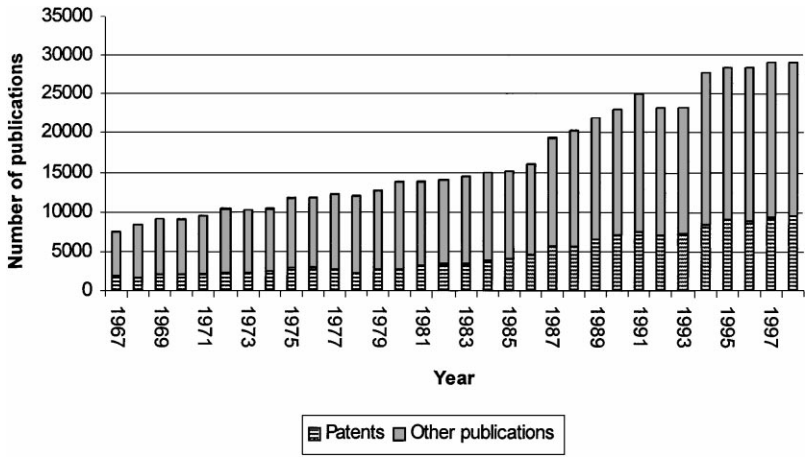

Fig. 5. Growth of publications in *Chemical Abstracts* which contain ?fluoro? or ?fluorin? or ?fluoride? amongst words in the title, index terms or abstracts split into patents and other types of literature. The total number of such publications is 535 875; of these 146 452 are patents and 389 423 other types of publication (e.g. journal articles, books, conference papers).

over the years. The total number of publications retrieved by the above strategy and published in the period 1967–1998 is 535 875 (of which 259 541 were published in the period 1989–1998); of these 146 452 are patents and 389 423 other types of publication (e.g. journal articles, books, conference papers).

4. Fluorine compounds and literature as percentages of total numbers

Fig. 6 reveals the annual variation (1990–1998) in the numbers of new organofluorine compounds as a percentage of the total number of new compounds recorded in the CAS Registry file [6], and also the number of publications containing information about fluorine and its compounds as a percentage of the total number of documents in the CA file, broken down into patents and other documents. The obvious downward trend in the percentage of new organofluorine compounds can possibly in part be accounted for by the large number of compounds registered in recent years which have been produced by combinatorial methods. The percentage of patents has remained roughly constant during the same time period (1990–1998). Fig. 3 reproduces the data recorded by Seebach [4], which, due to the restrictive method of determining numbers of papers by only including those indexed under the term ''fluorine'', indicates a lower percentage to be concerned with fluorine chemistry. The data in Fig. 7 are as recorded in this work through the search technique described above. The present study has used a more all-encompassing approach, but the shapes of the plots during the overlap period are very similar, as would be expected.

Factors which may have impact on the statistical data recorded include: (1) variations in the type of fluorine chemistry research over time; (2) the type of organisation where that research is being carried out (e.g. in universities or in industry); (3) trends in applications of existing compounds; and (4) where geographically research is being undertaken. Research into the fluorine compounds of transition metals (including work on nuclear processes) peaked in the 1970s, declining in the 1980s. Research in academic departments throughout the world also peaked during that period, reflecting that activity. As present growth areas include medicinal chemistry, agrochemicals and CFC substitutes and much research in these areas originates from industry, the patent literature could be expected to be increasing. The approximately stable percentage of patents in recent years involving fluorine compounds could be considered an increase when the huge growth due to combinatorial chemistry is taken into account. Since much present-day fluorine chemistry research originates in Japan, and Japan accounts for a high and increasing proportion of patents recorded in *Chemical Abstracts*, this is also likely to contribute to the high level of patenting activity.

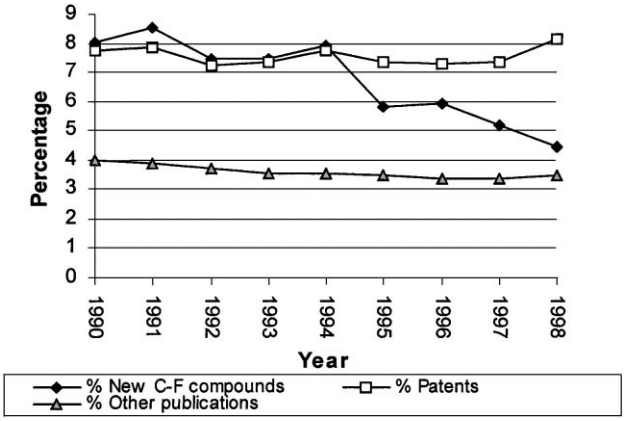

Fig. 6. Number of new C–F containing compounds, number of patents and number of other publications (non-patent documents) concerning fluorine and its compounds expressed as percentages of the total numbers reported each year. The data used are those recorded in Figs. 4 and 5, and for total numbers of compounds, patents and other documents as reported in the CAS Statistical Summary 1907–1998 [6].

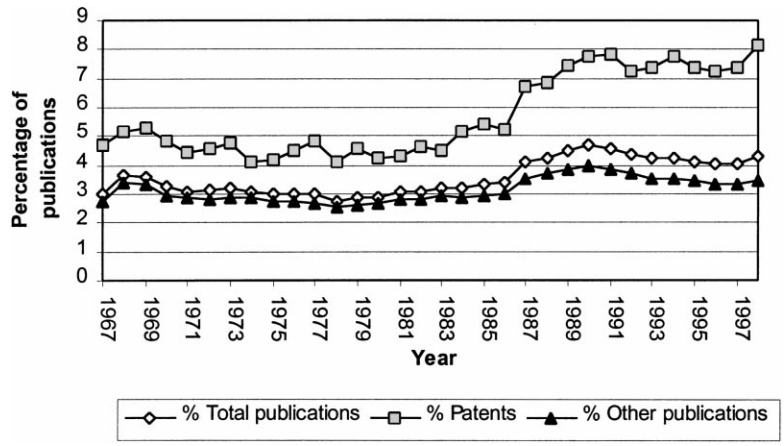

Fig. 7. Records on CAS Online with ?fluoro?, ?fluorin? or ?fluoride? in the title, index terms or abstract (data taken from Fig. 5) expressed as percentages of the total number of documents in the database (as recorded in the CAS Statistical Summary 1907–1998 [6]).

5. New ways to discover information about fluorine compounds

The approach of the millennium is seeing exciting developments in the field of chemical information. New methods of discovering references, manipulating data and identifying structure-activity relationships through use of chemical databases are evolving rapidly. At present there are a number of information retrieval services of relevance to workers in the area of fluorine chemistry. The first database for many enquiries is the *Chemical Abstracts* suite of products, available in a number of different formats, from printed, online through STN International and other database suppliers, *STNEasy* (which provides a web interface to the CAS data), *CA on CD* (which is convenient for providing in-house access to the information but which has less functionality than the on-line products), to *SciFinder*; the last format is rapidly becoming the preferred means of access, being particularly suited for use by chemists themselves rather than information professionals. Many of these products allow structure and sub-structure searching, as well as searching by keywords. At the time of writing it has just been announced that *Chemical Abstracts* will be making their full file back to 1907 available electronically. Until now, only data since 1957 has been available, and only that since 1967 fully searchable.

Even with the data back to 1907, the earliest fluorine literature will not be accessible by computer through *Chemical Abstracts*. For information prior to 1907, however, there is the Beilstein database [7], the data for which are taken from *Beilsteins Handbuch der organischen Chemie*, formerly used widely by the organic chemistry community. The Beilstein database is available online not only through hosts such as STN International, but also through Beilstein Information Systems' *CrossFire* service which has reinstated Beilstein as a major source of chemical information (indeed it is the most comprehensive organic reactions database available). The *CrossFire* service also provides the Gmelin database, which covers inorganic and organometallic chemistry. At the time of writing, there were 407 176 compounds recorded in *Beilstein CrossFire* which include an F atom in their empirical formulae, and 105 312 in Gmelin, although many of these are salts. Beilstein records 34 245 compounds with ''fluoro'' as part of the name. These two databases, as well as providing sophisticated access to substance, reaction and property information and associated references, also open up new ways to retrieve older references. For example, details of Moissan's publications can be easily identified, as can early references about fluorite and other fluorine-containing minerals through Gmelin.Although the sources described above would provide most of the information the fluorine chemist would require, a search of databases available through STN International reveals that a large number of databases contain information on fluorine and its compounds. The numbers of records in these databases since 1990 are recorded in Table 1. Although there is overlap in the coverage of the databases, if a particular application is required, these subject specific services may retrieve more relevant references in some cases.

Increasingly the internet is becoming a means of identifying useful information. A search using an internet browser such as *AltaVista* reveals in excess of 40 000 hits for the search strategy ''fluorine or fluoro or fluoride or fluorocarbon'' However, due to the imprecise search methods available through the internet at present, considerable caution should still be applied when considering this number: the same site can appear many times in the list of hits retrieved, many references can be irrelevant and the authenticity of the source should always be checked before relying on the data obtained.

More traditional sources of information are of course still relevant in this electronic age, however, and many of these can be identified through guides to the chemical literature [2,8,9]. Useful up-to-date bibliographies are provided in, for example, the *Kirk–Othmer Encyclopedia of Chemical Technology* and a spin-off volume specifically on fluorine chemistry [10,11]. Finally, and by no means least, the *Journal of Fluorine Chemistry* must be mentioned as the leading periodical in the area of fluorine chemistry; details of the

Table 1
Number of hits since 1990 for the terms ''fluoro? or fluorin? or fluoride?'' in a range of databases provided by STN International [5]

Database	Number of hits
CAPlus (Chemical Abstracts)	179828
SciSearch (Science Citation Index)	69556
World Patents Index	54242
JICST–E [Japan Science and Technology Corporation (JST), Information Center of Science and Technology (JICST), Japan]	39269
Compendex (Engineering Index)	16258
Inspec (Information Service for Physics, Electronics, and Computing)	14315
PROMT (Predicasts Overview of Markets and Technology)	11095
KKF (The Plastics Rubber Fibres File)	6476
NTIS (National Technical Information Service; Government Reports Announcements)	5370
CAB Abstracts (Commonwealth Agricultural Bureau Abstracts)	5307
American Petroleum Institute	4194
World Surface Coatings Abstracts	4032
Analytical Abstracts	3201

history of the journal can be found elsewhere in this centennial volume.

Acknowledgements

I would like to thank Professor Eric Banks (UMIST) for initiating this project and for his invaluable advice, and Dr Brian Lee (UMIST), Ms Jan Davies (STN International) and Mr David Walsh (formerly of STN International) for their assistance during the course of this work.

References

[1] H. Schulz, U. Georgy, From CA to CAS Online, 2nd ed., Springer, Berlin, 1994.
[2] H. Skolnik, The Literature Matrix of Chemistry, Wiley, New York, 1982.
[3] J.A. Wilkinson, Chem. Rev. 92 (1992) 505.
[4] D. Seebach, Angew. Chem. Int. Ed. Engl. 29 (1990) 1320 (see footnote 90).
[5] http://www.fiz-karlsruhe.de
[6] http://www.cas.org/casdb.html
[7] http://www.beilstein.com
[8] R.T. Bottle, J.F.B. Rowland, Information Sources in Chemistry, 4th ed., Bowker-Saur, London, 1993.
[9] R.E. Maizell, How to Find Chemical Information, 3rd ed., Wiley, New York, 1998.
[10] Kirk–Othmer Encyclopedia of Chemical Technology, 4th ed., vol. 11, Wiley, New York, 1994.
[11] M. Howe-Grant (Ed.), Fluorine Chemistry: A Comprehensive Treatment, Wiley, New York, 1995.