

# Comparison of the Coverage of Patents in Beilstein's Handbook of Organic Chemistry and Chemical Abstracts before, during, and after World War II

Helen Schofield

---

## Abstract

Little has been written about the coverage of patents by *Beilstein's Handbook of Organic Chemistry (Beilstein)*, so information professionals are often unaware of the value of the text as a source of patent information in the area of organic chemistry. *Beilstein* covered patents until 1980, and this paper evaluates the coverage and makes comparisons with *Chemical Abstracts*.

The evaluation was made through a series of bibliometric studies, using electronic versions of *Chemical Abstracts* and *Beilstein*. This work focuses on the period from 1930 to 1960, and in particular on U.S., British, German, and French patents.

World War II had a major impact on the production of chemical information as a consequence of its effect on the chemical industry and academic research, and these factors are discussed. The editorial policy adopted by the publications was also a major factor determining the comprehensiveness of coverage of patents, with *Beilstein* concentrating on the quality of information included and *Chemical Abstracts* aiming for comprehensiveness of coverage.

I conclude that *Beilstein* is a valuable source of organic chemistry patents and should not be overlooked when pre-1980 patent information is needed in the area of organic chemistry.

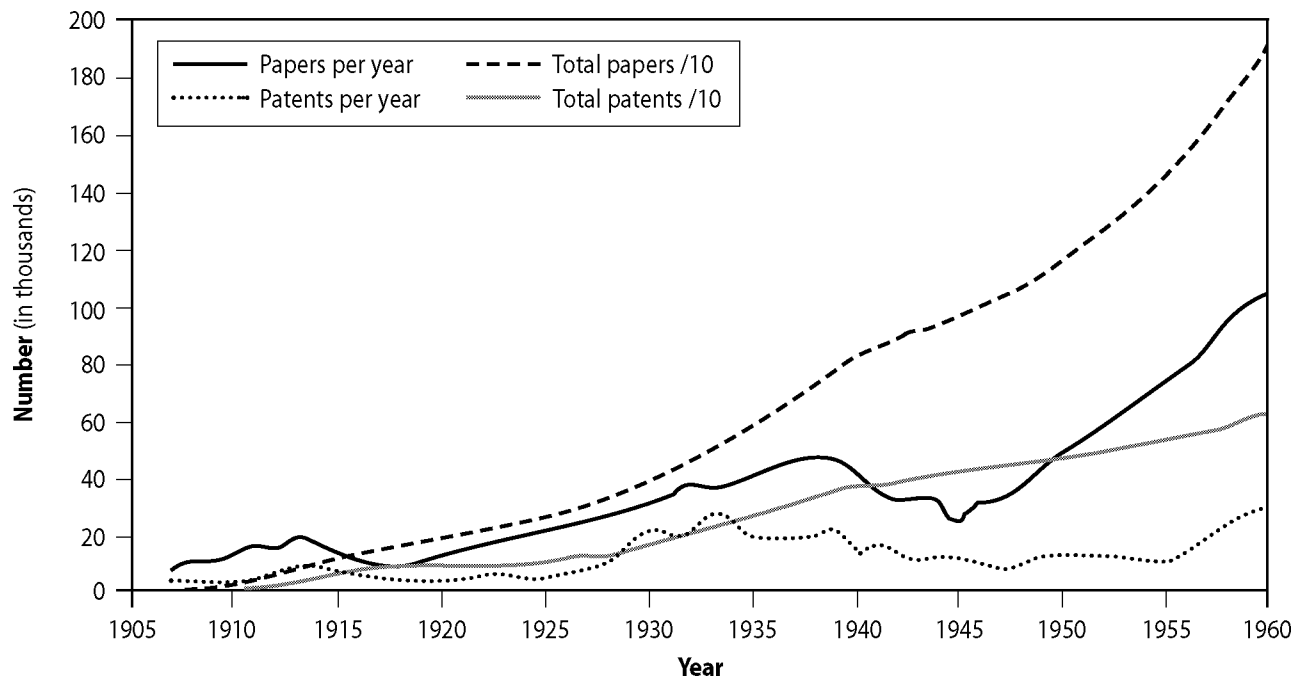
Patent systems were established throughout the world from the seventeenth century onward. According to Julian Smith (1965), the first chemical patent, issued in 1617, was British and concerned an oil for rust

prevention. Since then there has been a steady growth in chemical patents as a percentage of the total number of patents published,<sup>1</sup> and patents form an important part of the chemical literature, being an invaluable source of chemical information. Joseph Fleischer states that "the publication of printed patents has been an important factor in the development of science and industry. Published patents have stimulated technological advances directly and indirectly by revealing information that might otherwise have disappeared" (1951, p. 61).

This paper examines the patent coverage of two major secondary chemical information sources, *Beilstein's Handbuch der organischen Chemie* (Beilstein's handbook of organic chemistry; hereafter referred to as *Beilstein*) and *Chemical Abstracts (CA)*, from 1930 to 1960. One objective is to determine the value of the sources under consideration as information resources for patents published during this period, particularly from the information scientist's perspective. To achieve this aim, bibliometric studies have been undertaken to determine the numbers of patents in the information sources, and external factors that might have influenced the coverage of these publications will be examined to attempt to explain the data obtained from the bibliometric studies.

---

<sup>1</sup> J. Smith (1965, p. 62) states that in the mid-nineteenth century chemical patents accounted for about 5 to 10 percent of the total, and this percentage had risen by 1960 to approximately 25 to 30 percent. Fleischer (1952, p. 239) claims that in the United States chemical patents account for 3.2 percent of the total number of U.S. patents in the period from 1907 to 1909, 10 percent from 1925 to 1929, and 20.2 percent from 1945 to 1948. This growth is mirrored in other countries (Fleischer, 1952). K. L. Pavitt (1982, p. 37) reports that in 1973, 11.3 percent of U.S. patents was for chemicals (except drugs), 1.3 percent was for drugs, and 1.3 percent was for petroleum and related products.



**Figure 1.** Number of papers and patents per year and cumulative numbers of papers and patents in Chemical Abstracts, 1907–1960.

Much has been written about *CA* and its coverage of patents (Crane, 1944; Donnell, 1995; Baker, Horiszny, & Metanomski, 1980), but there is little material about *Beilstein's* patent coverage, although anecdotally *Beilstein* is rated by some as a superior source to *CA* for patent information on synthetic organic chemistry during the period in question.

When making comparisons, significant differences in the philosophies of the two publications should be considered: *CA's* policy has always been to try to be comprehensive in its coverage of the chemical literature, whereas *Beilstein* has aimed to provide evaluated, high-quality information and has never purported to be comprehensive.

World War II had a major impact on the production of chemical information as a consequence of its impact on the chemical industry and academic research. This effect is demonstrated by the number of papers and patents abstracted in *CA*, as shown in Figure 1 (Chemical Abstracts Service, 2002). During this period *CA* may have included more than one patent from a patent family; the policy to include only the first patent family member was adopted in the 1960s (Donnell, 1995).<sup>2</sup>

In addition to the number of papers and patents

included in *CA* each year, Figure 1 also shows the cumulative number of papers and patents in *CA* over the same period. This growth is similar to that observed in other studies. For example, Derek de Solla Price (1975) gathered statistics concerning the number of journal titles published, number of workers in certain industry sectors, and number of abstracts published. He uses these as an indication of the “size of science.” De Solla Price demonstrated that the number of scientific periodicals has increased by a factor of ten during every half century between 1665 (the first date of publication of *Philosophical Transactions of the Royal Society*) and the date of publication of his works, representing a doubling every fifteen years. This trend is mirrored by the publication of abstracts journals and also by the number of scientific papers published. Similar statistics concerning journals were produced by Andrew Osborn (1980, p. 25). A recent study (Schummer, 1997), which details the growth in numbers of new chemical substances reported in the last 200 years, shows a doubling in the number of compounds produced every 12.9 years.

### Why Chemical Abstracts and Beilstein?

*Chemical Abstracts* was first published in 1907. It was born partly out of American chemists' dissatisfaction with

<sup>2</sup> The policy not to include more than one patent from a patent family was dealt with by the publication of the *Chemical Abstracts Patent Concordance*, which began in 1963.

the coverage of American chemical literature by European abstracting journals (Baker et al., 1980). *CA* was sponsored by the American Chemical Society, with its first editor being W. A. Noyes, Sr. The first issue of *CA* contained 11,847 abstracts (Wolman, 1988), taken from 396 journals (Donnell, 1995). The number of journals covered increased to 1,000 by 1922, 2,000 in 1932, and around 9,000 today, along with patents from 27 patent offices. *CA*'s mission was to abstract the complete world's literature of chemistry, at first glance a straightforward objective but in fact one that led to problems with the definitions of three words: *complete*, *abstract*, and *chemistry*. E. J. Crane (1889–1966), editor of *CA* from 1915 to 1958, indicated that publications suitable for inclusion were “studies of new chemical reactions, new information on known reactions, chemical, physical and biological properties of elements or compounds, apparatus of particular interest to the chemist or chemical engineer and procedures that in themselves may not involve chemistry but are essential to an industry that is generally considered chemical” (1958, p. 21). From the outset *CA* covered applied and industrial chemistry as well as pure chemistry.

*CA* is well known as a source of patent information, and the histories of the publication mostly mention its patent coverage. It is the most important secondary source of information for chemists and in a sense is a benchmark against which other chemical information sources are measured. Owing to the recent addition of older *CA* data to electronic versions of the database, one can now undertake research previously not possible.

*Beilstein's Handbook of Organic Chemistry* is not an abstracting service but rather a secondary source that contains evaluated information. Beilstein's objective was to include only compounds and facts that were known to be reliable in terms of current scientific knowledge so that the user is provided with a “concentrate” of the original literature free from errors and trivial or unvalidated information. Despite the huge growth in the literature this was the aim of the publishers until recently. This accuracy makes *Beilstein* different from other services, which usually made no attempt to check the accuracy of the information included. The first edition was published in 1881–1883 and contained approximately fifteen thousand organic compounds. The work comprised two volumes (2,200 pages). Twenty-three journals were covered, and the reference list was 23 pages long. The second edition consisted of three volumes (4,080 pages, 1885–1889), and the third edition was eight volumes (around 11,000 pages, 1892–1899). The third edition

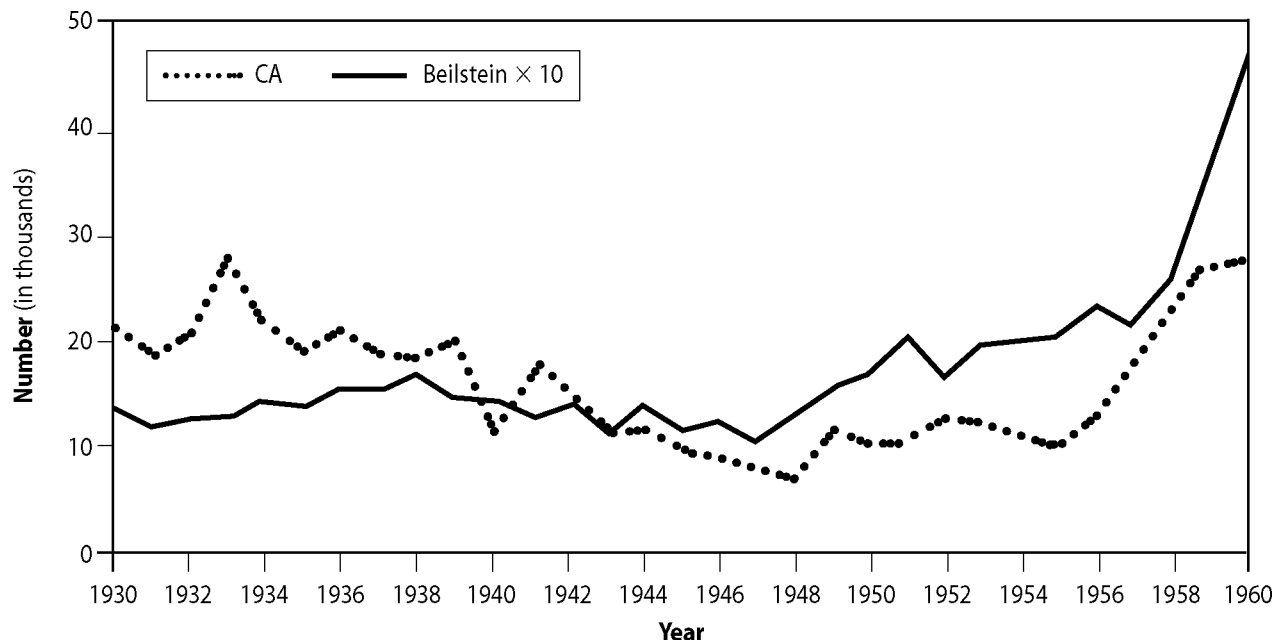
was the last produced by Beilstein himself: up to this time the production of the *Handbook* was almost exclusively his work. Prior to publication of the third edition Beilstein transferred responsibility for publication to the German Chemical Society (Deutsche Chemische Gesellschaft).

Although histories of *Beilstein* exist (see, for example, Huntress, 1938; Luckenbach, 1981a, 1997; Richter, 1938), little has been written about its coverage of patents, and many people are unaware that it covers a significant amount of patent information; indeed most guides to the chemical literature fail to mention that it covers patents at all. This failure is partly because *Beilstein* stopped covering patents in 1980 and at present has nothing to offer the searcher interested in recent patents. However, *Beilstein* covers older synthetic organic chemistry information possibly more thoroughly than any other secondary source. The earliest patents covered in *Beilstein* are from 1869 (GB1948 granted to William Henry Perkin and GB1936 granted to Heinrich Caro, both for the synthesis of 2-hydroxyanthraquinone). *CrossFire Beilstein* (MDL Information Systems GmbH, 2002) enables a study of the coverage of *Beilstein* to be made, as statistics can be generated to establish trends.

The total numbers of patents in *CA* and *Beilstein* for the period 1930 to 1960 are 490,897 and 52,908, respectively. The breakdown of these by year is shown in Figure 2. (The *CA* data in Figures 1, 2, 5, and 6 are analyzed by *CA* publication year rather than patent publication year, whereas the *Beilstein* data are by patent publication year.) Thus the *Beilstein* data do not reflect any fluctuations in availability of patent documents owing to war or other factors, as there is no indication as to when the patents were added to *Beilstein*.

Clearly *CA* covers many more patents than *Beilstein*, as would be expected in view of the subject coverage of the two products. *CA* covers all branches of chemistry (including, e.g., biochemistry and chemical engineering), whereas *Beilstein* covers only synthetic organic chemistry. Further, patents were only covered in significant numbers by *Beilstein* from around 1925.

One important study of patents in *Beilstein* has been made by Stephen Adams (1999), who examined the representation of patents concerned with ten specific compounds. The compounds chosen reflect a broad range of organic substances and were deemed likely to be extensively patented. Adams's study looked at patents on the compounds under investigation and revealed that a diversity of countries was covered. The study showed that patents with the same assignee and similar



**Figure 2.** Total numbers of patents per year in Beilstein and CA.

publication years but different countries of origin often occur in the *Handbook*. These patents are therefore probably from the same family, and Adams suggests that *Beilstein* is thus a potentially valuable source of patent family information, even though it does not explicitly state that these patents are from the same family. However, with this knowledge, it seems reasonable to assume that *Beilstein's* producers did not have a policy of including only the first member of a patent family. Adams concludes that *Beilstein* covers a substantial volume of unique patent data.

A survey of the literature and inquiries to the Beilstein Institute have not revealed the precise selection criteria for *Beilstein* with regard to patents. However, *Beilstein* was never intended to be exhaustive in its coverage of any type of literature, the emphasis being on quality rather than quantity. So of course fewer patents would be present in *Beilstein* than *CA* for any topic or substance expected to be covered by both sources. Further, if a journal article on a topic was available, it would

probably be included in *Beilstein* in preference to a patent; that is, patents would only be included as a “last resort” if no journal paper was available. These factors must be borne in mind when comparing the two sources.

### Obtaining Patent Information

Table 1 summarizes the means by which the data contained in *Beilstein* and *CA* can be obtained. A bibliometric study of *Beilstein* was undertaken using the *CrossFire* system. Patent data stored in the *Beilstein* database are retrieved by searching for patent country codes and publication years. Figure 3 is an example of a patent record retrieved from *CrossFire Beilstein*. Similar statistics can also be generated for *CA*, which was done by making use of Chemical Abstracts Service (CAS) files on STN (STN International, 2002).

### Trends and Coverage

The electronic information retrieval applications described in Table 1 were used to generate patent statistics

**Table 1.** Format and Means of Delivery of *CA* and *Beilstein* Information

Format	<i>Chemical Abstracts</i>	<i>Beilstein</i>
Print	Printed <i>Chemical Abstracts</i>	<i>Beilstein Handbuch der organischen Chemie</i> (ceased publication in 1998)
Client-server desktop version (designed primarily for end users)	SciFinder (SciFinder Scholar for academic institutions)	CrossFire (via Commander interface and also via CrossFire Web for smaller institutions)
On-line metered service (designed primarily for chemical information scientists)	Through on-line hosts such as STN International and Dialog	Through on-line hosts such as STN International and Dialog

**Figure 3.** Part of a patent record from CrossFire Beilstein. MDL CrossFire Commander: © 1995–2002 MDL Information Systems GmbH. Beilstein Database: © 1988–2002, Beilstein-Institut zur Förderung der Chemischen Wissenschaften licensed to Beilstein Chemiedaten und Software GmbH and MDL Information Systems GmbH.

to estimate the number of patents in *CA* and *Beilstein* for the subject area covered by *Beilstein* (synthetic organic chemistry); to compare the comprehensiveness of the patent coverage of *CA* and *Beilstein* through analysis of the countries of origin of patents over time; and to compare the number of patents in *CA* with the numbers of patents in all subject areas in order to investigate any differences between patent output in chemistry and other subject areas. These comparisons are made for U.S. and British patents. (German patents were not analyzed in this way because data concerning cumulative num-

bers of German patents for all subject areas were not available.)

To estimate the number of patents in *CA* that are in the same subject area as the subject coverage of *Beilstein*, the appropriate subset of *CA* must be analyzed. A convenient method of creating the subset is to use STN to search by *CA* subject “section.”<sup>3</sup> The subject section for organic chemistry for the period under consideration is 10. The numbers of abstracts in *CA* section 10 compared with the numbers in *Beilstein* are shown in Table 2. One should not read too much into these statistics, as

<sup>3</sup> All documents abstracted in *Chemical Abstracts* are allocated to broad subject sections in order to group abstracts on similar subjects; this system was traditionally used to assist the browser of printed *Chemical Abstracts*.

**Table 2.** Comparison of the Number of Patents in *CA* Subject Section 10 (Organic Chemistry) with the Number in *Beilstein*

Year	<i>CA</i> Section 10	<i>Beilstein</i>
1930	1,141	1,349
1935	1,532	1,393
1940	1,283	1,423
1945	1,446	1,134
1950	2,446	1,658
1955	2,949	1,998
1960	5,050	4,752

there are patents in other sections that might have been candidates for inclusion in *Beilstein* and similarly there are patents included in section 10 that, on the basis of their subject coverage, would not be expected to be included. It is not possible to readily ascertain how these numbers balance out.

The coverage by country of patents in *Beilstein* is shown in Figure 4 (for the countries represented highest numerically). As can be seen, German and U.S. patents dominate, with relatively few French and British patents featured. The *CA* data are presented in Figure 5; these show that in *CA* a greater proportion of the patents covered are British or French than in *Beilstein*.<sup>4</sup>

Figure 1 shows a decrease in the number of all papers and patents covered by *CA* during both world wars, with patent numbers appearing to recover more slowly than journals. From 1933 to 1939 the number of patents in *CA* was declining, whereas the number of papers continued to increase. The decline in the number of journal articles did not begin until the start of World War II. Edward P. Donnell (1995) states that for *CA* during both world wars some German and other European publications were difficult to obtain, and missing publications were covered in the years immediately after the wars. Presumably this lack of coverage was not the case for U.S. patents in World War II, as suggested by the data presented in Figure 5, which shows that during the war years the number of U.S. patents abstracted by *CA* did not decrease as sharply as for other countries.

Figure 2 shows that *CA* covers approximately ten times as many patents as *Beilstein* from 1930 to 1960. The overall trend in patent numbers is approximately

the same in both sources. Figure 4 shows that the number of U.S. patents in *Beilstein* grew steadily, with only minor fluctuations during World War II and its postwar period. German patents are the most numerous around 1930–31, but they decline steadily from 1939 until 1945, continuing to decrease to almost zero. In 1948 the number of German patents starts to rise. The number of French patents is always low, at less than 250 per year, and British patents are also few in number until the mid-1950s, at which time they are represented in numbers similar to German patents. Figure 4 also indicates that for *Beilstein* World War II had a greater effect on German than U.S. patents. Data in Figure 5 generally support the evidence from *Beilstein* concerning the impact of war on chemical patents and the variation between countries.

Figure 6 shows total patent output (considering all subjects together) for the United States and Britain compared with the number of patents in *CA*. The *CA* data tend to follow the overall patent output of the United States owing to the high proportion of U.S. patents in *CA*, with a time delay reflecting the time taken for the patents to be abstracted and published in *CA*.<sup>5</sup>

### Establishing the Comprehensiveness of Chemical Abstracts and *Beilstein*

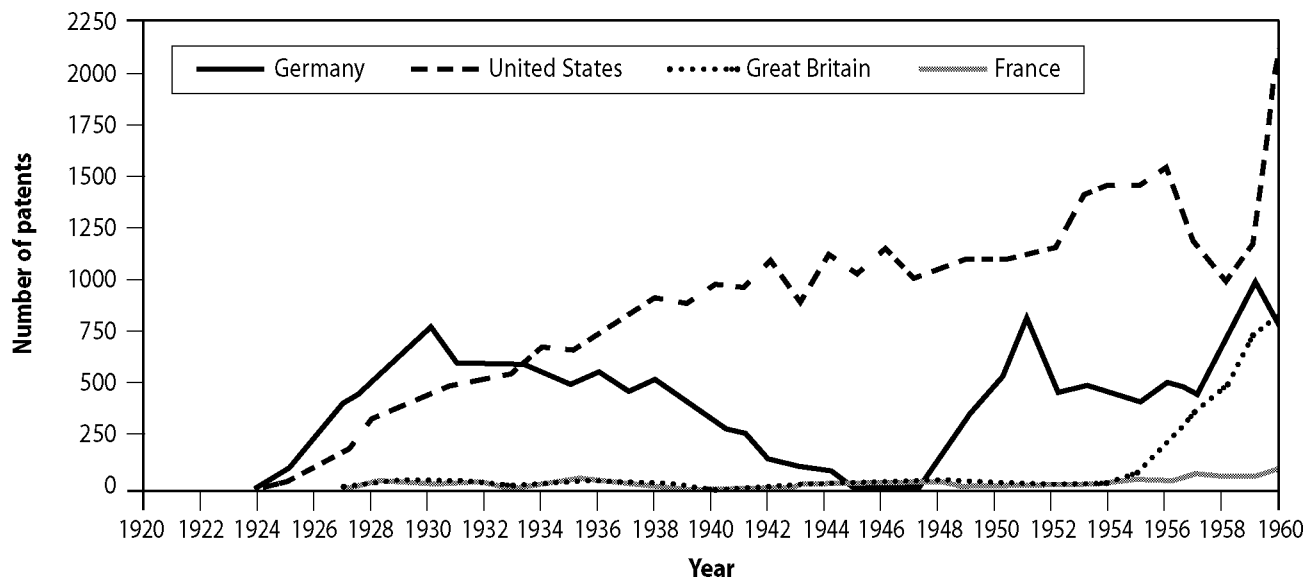
The general principle behind the following case studies was first to identify a list of patents potentially covered by both *Beilstein* and *CA*, that is, from the area of organic chemistry. The list was identified by using patent groups or classification codes, and once it had been generated, the patent numbers were checked against the databases. The studies focus on British and U.S. patents. Examining the coverage of German patents or patents from other countries would also be desirable; this type of study will most likely be possible in the not-too-distant future if MicroPatent or other organizations provide data-mining facilities similar to those available for U.S. patents as described in the second case study below.

#### Case Study 1: British Patents

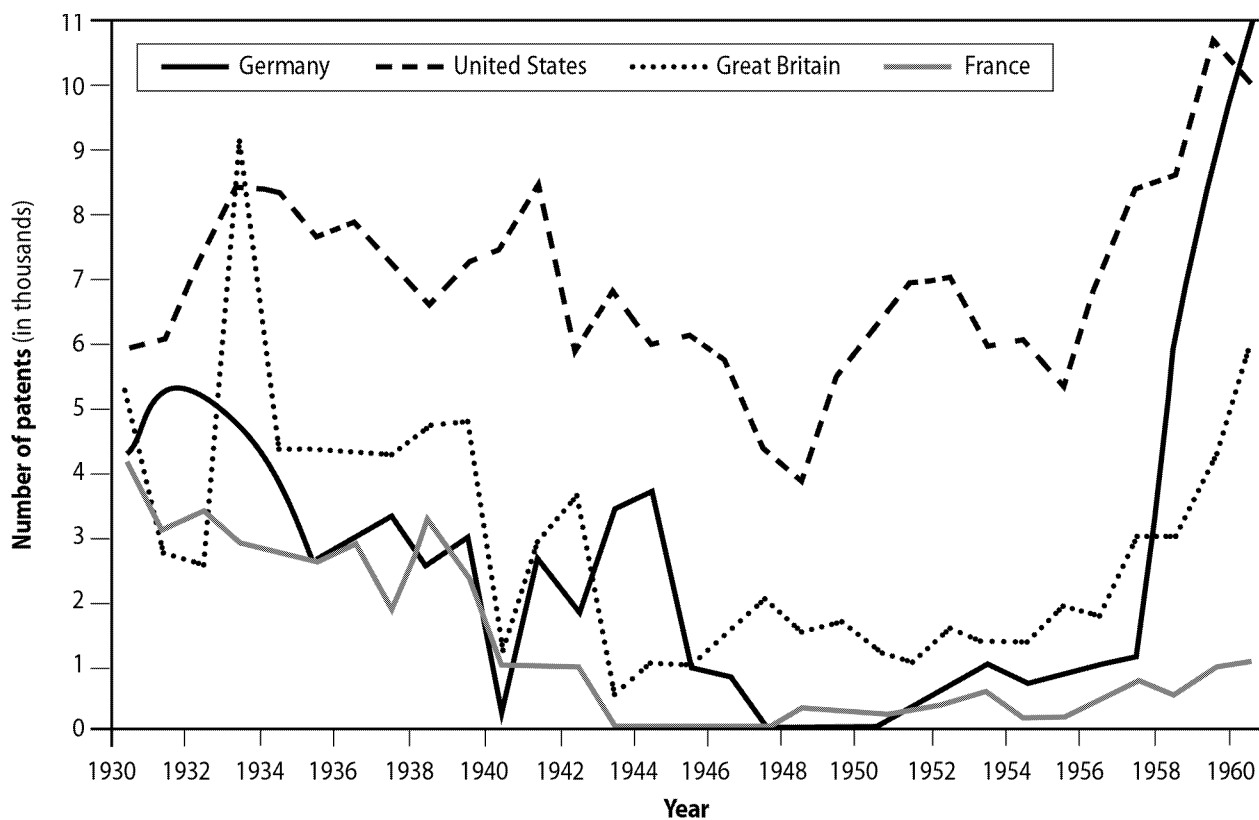
I examined the coverage of British patents by *Beilstein* and *CA*. Although there are a number of Web sites and databases that provide British patent information, often including displays of the full text of patent documents,

<sup>4</sup> *Chemical Abstracts* covered patents from many other countries from 1930 to 1960, but no attempt has been made to establish the comprehensiveness in this paper.

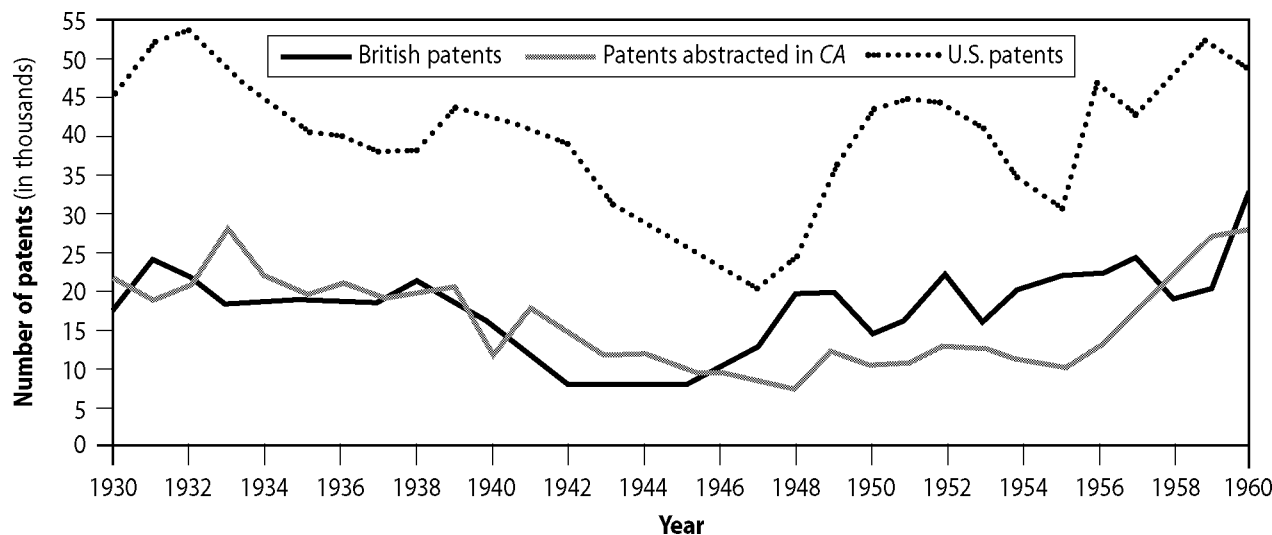
<sup>5</sup> Overall numbers of U.S. and British patents were compiled from data obtained from the U.S. Patent and Trademark Office (2002) and the U.K. Patent Office (2002). *Chemical Abstracts* data are again analyzed by *Chemical Abstracts* publication year rather than patent publication year.



**Figure 4.** Patents per year in Beilstein for the period of 1920 to 1960, broken down by country of origin for Germany, Great Britain, the United States, and France. Data back to 1920 indicate that patents were only covered in significant numbers by Beilstein from around 1925.



**Figure 5.** Patents per year in CA for the period of 1930 to 1960, broken down by country of origin for Germany, Great Britain, the United States, and France.



**Figure 6.** Total number of U.S. (U.S. Patent and Trademark Office, 2002) and British (U.K. Patent Office, 2002) patents per year in all subjects, and patents per year in CA.

it is not possible to automate the comparison of lists of patents retrieved in this way with those included in *Beilstein* or *CA*. In order to examine a broad range of organic chemistry patents, I decided to use printed publications for this part of the study. To generate a list of candidate patents, first an appropriate class group was identified.<sup>6</sup> The class selected was IVb, General Organic Chemistry. Sample sets of patent numbers were scanned to identify patents that had been classified in this subject group.

The patent numbers were checked against *Beilstein* and *CA* by using *CrossFire Beilstein* and *SciFinder*. The samples used and corresponding data are shown in Table 3. Only a very small percentage of the patents in the sample are covered in *Beilstein*. In some cases there appear to be good reasons for this; for example, when the patent abstracts are consulted, the subject matter of the material in the patent is clearly not of the type covered by *Beilstein*. This is especially true for the earliest patents studied, as the scope of the subject group IV is broader than that covered by group IVb. However, examination of the later samples of patents reveals that many would have been expected to have been covered by *Beilstein* if subject matter had been the only selection

criterion applied. These findings are not surprising on the basis of the data reported in Figure 4, which shows that relatively few British patents are included in *Beilstein*, although it might have been expected that a greater proportion of the more recent patents would have been included as the number of British patents increases toward the end of the period under consideration. Although the coverage of British patents by *CA* is more comprehensive, there are still significant gaps (Table 3). These findings are in contrast with the results shown in Table 2, which shows organic chemistry patents covered in similar quantities by *CA* and *Beilstein*, presumably because of the poor coverage of British patents by *Beilstein*.

### Case Study 2: U.S. Patents

For a study of U.S. patents, MicroPatent's database (MicroPatent, 2002)<sup>7</sup> provides an excellent search tool. The database searches by U.S. patent classification, and the results can be output to a Microsoft Excel spreadsheet. For this part of the study a more automated search of the patent coverage of both *Beilstein* and *CA* was possible. Again, an appropriate subject area had to be selected to identify patents that were likely to be included in *Beilstein*.

<sup>6</sup> Countries that have a patent examination system have a patent classification system devised by the patent office, the role of which is to assist in the patent examination process by enabling identification of prior art. Patent classification systems are also an aid to identification of patents of interest to the researcher seeking technical or scientific information. Class numbers are allocated to patents by the patent offices. Patent systems have evolved separately in different countries as have the classification schemes.

<sup>7</sup> MicroPatent's collection of fully searchable, granted U.S. patents goes back to 1836, that is, to patent no. 1. The database is updated six times per year and uses the latest U.S. class numbers; it is therefore not necessary for the searcher to have knowledge of old classification numbers.

**Table 3. Number of British Patents Classified in Group IVb (General Organic Chemistry) and Their Occurrence in *CA* and *Beilstein***

Patent Number Range	Approximate Year	Number of Patents in Group IVb	Number Present in <i>CA</i>	Number Present in <i>Beilstein</i>
500,000–501,000*	1937	87	81	0
600,000–601,000	1945	73	49	3
700,000–701,000	1953	60	25	0
800,000–801,000	1958	81	78	1

\* Group IV had not been subdivided into IVa, IVb, and IVc during this period; so the data refer to Group IV: acetylene; cellulose; chemistry, organic; dyes and dyeing.

I undertook two studies involving samples of patents. The first study was of U.S. patents concerned with piperidines (U.S. classification numbers 546184 to 546248). A search using MicroPatent retrieved 1,289 patents in this classification range for 1930 to 1960. Of the patents from the MicroPatent master list 860 were present in *CrossFire Beilstein*.

*SciFinder* was used as the source of *CA* information for the comparison. Of the 1,289 patents searched for on *SciFinder*, only 34 were not present. Although these 34 patents represent only a small percentage of the 1,289 patents in this period, examination of the subject material covered suggests that most would have been expected to have been included because their subject coverage was of a type normally included by *CA*.<sup>8</sup> The absence of these patents cannot be explained by another member of the patent family having been included in preference to the U.S. patents as this was not *CA*'s policy at the time.

When the subject matter of the patents that do not appear in *Beilstein* is examined, at least half of them would not be expected to be included because they do not map on to the subject matter covered by *Beilstein*. However, it is not immediately clear why many others were omitted from *Beilstein*. As already mentioned, it has not been possible to determine categorically the selection criteria used concerning inclusion of patents in *Beilstein*.

A second study covering a broader range of patents was carried out, again making use of MicroPatent's services. A set of patents was retrieved covering all aspects of chemistry. This set was narrowed down by using the keyword "synthes\*," where the asterisk is a truncation symbol, and selecting only those patents published in one month in each year for the period of 1930 to 1960.

This search resulted in a set of 729 patents. The number of these contained in *CrossFire Beilstein* was 342, and *CA (SciFinder)* contained 700. Spot checks of the subject content of the patents not included in *Beilstein* again reveals that if coverage was the only criterion for inclusion in *Beilstein*, many of the missing patents would have been included. This finding confirms once more that other factors are taken into consideration when decisions are made about which documents should be included.

### Depth of Coverage

In addition to the breadth of coverage of an information source, it is also important to consider the depth of coverage in terms of the amount of information that can be displayed in an answer retrieved from a search as well as the availability of alternative search approaches. At present *Beilstein*, either on-line through a host such as STN International or through *CrossFire*, wins this contest. *Beilstein*'s records, including old patents, are fully indexed from the scientific content perspective and can be searched by structure, reaction, property, and bibliographic data (Meehan & Schofield, 2001; Heller, 1998). However, the patent metadata, for example, the application date, inventor name, and title, are not consistently included, which means that retrieval from those starting points cannot be relied on. Approaches from these directions are best made using an alternative source such as MicroPatent. At *CA* work is being undertaken to index its pre-1967 literature fully, which will significantly improve the depth of coverage. (See, for example, Ridley [2002] and STN International's Web page [STN International, 2002] for further information on search techniques and displayable fields for *CA*.)

<sup>8</sup> These subjects were new chemical reactions; new information on known reactions; chemical, physical, and biological properties of elements or pure compounds; preparation, purification, and analysis of elements or compounds; apparatus of particular interest to the chemist or chemical engineer; and procedures that in themselves may not involve chemistry but are essential to an industry that is generally considered to be chemical (Crane, 1958, p. 21; see also under "Why *Chemical Abstracts* and *Beilstein*?" in this paper).

## Factors with Potential Impact on Patent Coverage of Secondary Chemical Information Sources

External events potentially affect publication, of primary literature (e.g., patents and journal articles), production of secondary sources, and the secondary sources' coverage of the primary publications. A number of discussions have taken place relating to the impact of war, economic depression, and other national or world events on patenting and its relation to such issues as economics and population growth (see, e.g., Jewkes, Sawers, & Stillerman [1969]). There are indeed many factors that can affect the coverage of an information source, and I will now discuss several of them.

### The Interwar Years

In addition to wars the Depression (1929–1939) clearly affected industrial growth and therefore the production of patents (Figure 6). Job losses and the lack of opportunity for newly qualified scientists and technologists to gain experience in industry and hence contribute to industrial growth were the causes. However, Fred Aftalion (1991, p. 169) states that the Depression had less effect on America's leading industries than the election of Franklin D. Roosevelt in 1933 and his New Deal, whereby the power of unions increased and pricing codes were introduced. Aftalion states that overall in the United States the chemical industry was less affected than others, as companies tended to move to the south of the country where resources (such as cotton, wood, sulfur, salt, and oil) were plentiful and the workforces were not so well organized into unions. The data in Figure 6 support this, with the number of patents in *CA* (i.e., chemical patents) not decreasing to the same extent as the overall numbers from the United States and Britain.

The Depression probably had little effect on the flow of information to the publishers of secondary sources, and so the main factors affecting *CA* and *Beilstein* were more concerned with editorial policies in place at the time.

### Impact of War

Secondary sources can be a good source of data for providing evidence for the impact of outside events on publication. De Solla Price (1975) used such statistics, and his work of most relevance to this study is his analysis of the number of papers abstracted in *Physics Abstracts*. During World War II (1939–1945) the number of pa-

pers in *Physics Abstracts* dropped to a minimum of around one-third of what would have been expected if the pre-war trend had continued; in the decade between 1938 and 1948 around half the number of papers than would otherwise have been expected were published. The data recorded for *CA* in Figure 1 show a similar decrease for chemistry, with the number of patents decreasing even more than that of journal papers. De Solla Price (1975, p. 161) states that war has two effects on science: the stimulation provided by such projects as the development of the atom bomb has a positive effect, but the deployment of personnel and money to war that might otherwise be invested in research tends to counteract the positive effect. The statistics suggest that the more negative effect has a direct impact during the war years but that after the war scientific output, as measured by journal and patent publication, catches up and continues to increase following the same trend as prior to the war period. In other words, the two effects tend to cancel each other out.

E. J. Crane (1944, p. 1,479), the longtime editor of *CA*, had similar, if less positive, observations to make. He states that “war does stimulate research in certain special directions, does overcome some inertial obstacles, does often foster better cooperation within a nation because of common danger, and does lead to some lasting benefits from research, but the results of wartime investigation are, for the most part, of value to human progress only in a negative way. They help put out the fire in the laboratory so that the good work can begin again.” Crane claims that after World War I chemists' productivity did not return to prewar levels for ten years and that during the interwar years productivity seemed to parallel the six years prior to the start of World War I. World Wars I and II had a similar effect, and the decreases are sharper than the increase between the wars.

Part of the picture, as observers have often noted (e.g., Baker et al., 1980), is the classification of much of the war-related research, which was therefore not published until after the war, if at all. A related factor is the capture of military and industrial research. Toward the end of World War II the Americans and British mounted an exercise to obtain German technological information, partly to try to shorten the war and partly to keep the information out of Soviet hands. Many documents were obtained from such companies as I.G. Farben and taken to London and various locations in the United States. Needless to say, in addition to the general devastation of the war this action had a major impact on publications

from German organizations until the rebirth of German industry in the 1950s. Then applications again flowed into the German Patent Office, creating a significant backlog of around 284,000 patent applications that were not cleared until the late 1960s.

For U.S. chemical patents the impact of World War II was less than in Europe, and that impact was delayed, as shown in Figure 5. Figure 6 indicates that the total patent output may be affected more than total chemical patent output by war but that chemical patent output appears to have taken longer to recover than overall patent publication, possibly owing to the concentration of the chemical industry in Germany prior to the war and Germany's former high output of chemical patents. After the war the German chemical industry had to be rebuilt, and the Americans still had to catch up to Germany's prewar patent productivity.

Another explanation for the drop in *CA* patent numbers occurring later than the drop in overall patent numbers may be the delay, increased by wartime conditions, between printing of the primary publication and its appearance in secondary sources. According to Donnell (1995) and Dale Baker and coworkers (1980), the war disrupted *CA*'s channels for obtaining material, and there was a paper shortage in 1943–44; these factors affected the publication of *CA*, hence potentially artificially reducing the number of patents and papers included in it.

Chemical patents as a percentage of all patents, rather than considered in their absolute numbers, display slightly different trends. Fleischer (1952) used information compiled from data reported in *CA*'s Numerical Patent Index, a cumulative patent index to *Chemical Abstracts* 1907–1936 (Special Libraries Association, 1944), and for non-U.S. patents, in other sources (e.g., *Chemisches Zentralblatt* and the *Journal of the Society of Chemical Industries*). He observed that the percentage of patents that could be considered “chemical” decreased slightly in Britain and Germany during World War II, possibly owing to the need to keep chemical inventions secret, but remained roughly constant in the United States, even though absolute numbers decreased.

As would be expected, after the war the number of patents published began to increase, as reflected in the number of patents reported in the secondary literature (Figures 1, 2, 4–6). Baker and colleagues (1980) also suggest that after the war the number of patent duplicates increased owing to increased patenting in multiple countries (i.e., an increase in the number of patent families). This duplication has probably increased the num-

bers of patents in *CA* out of proportion with actual “unique” patents, and presumably it led to the decision to include only the first patent family member in *CA* after 1962. *Beilstein*, on the other hand, was unlikely to have suffered from these problems in the same way because of its different editorial policy, although as Adams (1999) suggests, there is evidence of more than one member of a patent family being included in *Beilstein* quite often.

### Editorial Policies of Secondary Information Sources

*CA*'s policy is to cover the whole of the chemical literature. However, it is well known that its coverage of the patent literature was not comprehensive in the earlier years of publication but that patent coverage has improved over time as more countries have been included. The first study discussed under “Case Study 2” above indicates that the coverage of U.S. patents by *CA* from 1930 to 1960 is very good, as evidenced by the high proportion of piperidine patents that are included and the similarly high proportion of “random” patents covered as determined through the second study in this case. As seen under “Case Study 1,” the coverage of British patents appears to be less comprehensive. The official policy from 1963 on at *CA* was only to include the first disclosure of a particular invention. Before that date more than one member of a patent family may have been included in *CA*.

As already mentioned, it has not been possible to determine *Beilstein*'s official selection policy regarding patents. However, it appears that the editors focused on U.S. and German patents, with many other countries represented but given patchy coverage. Patent family information is not included per se, but this information can on occasion be obtained through *Beilstein* (Adams, 1999). A note in the first volume (Prager & Jacobson, 1918, p. xxv) of the fourth edition of the *Handbook* implies that the German patents included were extracted from *Chemisches Zentralblatt*, which in turn covers patents from the main patenting countries of the world, though there are differences between the coverage of *CA* and *Chemisches Zentralblatt*. The fact that the patents were extracted from *Chemisches Zentralblatt* is most probably a consequence of the fact that from the start of the preparation of the fourth edition of *Beilstein* (which began in 1918) the material for inclusion was selected and edited by the staff of *Chemisches Zentralblatt*, which was published by the German Chemical Society (Richter, 1938). This arrangement was in place for around twenty

years. For the second supplement of the fourth edition the *Beilstein* editorial staff reverted to consulting the original literature for about forty of the most important journals, referring to *Chemisches Zentralblatt* for the remainder of the journals and presumably other publications. In view of *Chemisches Zentralblatt's* coverage of patents from other countries in addition to Germany, the coverage of *Beilstein* is likely to be closely linked to that of *Chemisches Zentralblatt*. The patent coverage of the two publications would probably have been related into the war period; although publication of *Chemisches Zentralblatt* ceased for a time during the war, *Beilstein* continued to be published, albeit with volumes appearing less frequently. A later note in the *Beilstein Handbook* states that at least some of the patents from the war period were obtained from Bayer (Richter, 1958, p. xii). In addition, *Beilstein* caught up with the literature after the war.

As already stated, *Beilstein* was never intended to be comprehensive in terms of covering all literature, but rather it was to provide a condensed, evaluated reference source. The best assessment of *Beilstein's* coverage of patents can probably be made by considering the general philosophy behind *Beilstein*. Patents were probably considered along with other primary chemical information sources and were analyzed according to the critical screening procedure described by Reiner Luckenbach (1981b; 1996).<sup>9</sup> Namely, papers were analyzed for correctness, depth, completeness, objectives of the publication, accessibility of the publication, origin of information, identity of substance, characterizing data, preparation, purity, methods of measurement, completeness, and accuracy of data (i.e., the philosophy was not to duplicate information but to include the most reliable, easily accessible, and valid source). Such a philosophy might lead to discrimination against patents for good reasons, but clearly patents were sometimes the only source of information on a substance, in which case they were included. For compounds to be included, they needed to have been isolated or definitely identified and of unambiguous constitution, and usually a structure would be needed.

The preference for German and U.S. patents was possibly because *Beilstein* itself is a German publication and the United States the most important patenting country in the world. Since patents were lower-priority documents for inclusion than journal papers, maybe it

was deemed less important to scan all patent countries: the potential return on investment probably was not justified.

### Validity of Comparing Countries

The statistics raise a number of questions about comparisons. How valid is it to compare numbers of patents originating from different countries or even from the same country over a period of time? Does a German patent have the same quantity or quality of information content as a British patent? Does an American patent of 1980 have the same degree of innovation as one published in 1950? Lengths of patents vary between countries, and patent laws also differ; even patent law changes within one country can have a major impact on patenting activity. Much debate has occurred about whether patent output can be regarded as an indicator of a nation's productivity (see, e.g., Gilfillan, 1960; Jewkes et al., 1969; Schmookler, 1960; Thomas, 1936; Wilson, 1937), and concerns have been expressed about the apparent decline in patenting in the United States relative to population growth. A relative decrease in patenting at least in part could be caused by an increase in the number of inventions that are kept as "trade secrets" and never disclosed as patents. It is recognized that know-how and trade secrets are complementary to patents in protecting inventions, as evidenced by the activities of the German dyestuff industry during World War I (Arora, 1997). However, the objective of this paper has not been to compare patenting activity in different countries but rather to examine the comprehensiveness and value of secondary information sources for patent information.

### Conclusion

*CA* and *Beilstein* are valuable sources of chemical patent information. The major strength of *CA* appears to be the breadth of coverage, in both country and subject matter. However, *Beilstein* provides in-depth coverage of the patents that are included and greater possibilities with regard to search approaches in the period from 1930 to 1960. *Beilstein's* coverage is most reliable for German and U.S. patents, but it also contains small numbers of patents from other countries. The comprehensiveness of *CA's* coverage of patents appears to vary, with U.S. coverage being very good, whereas other countries are not so well represented. The strengths and weaknesses

<sup>9</sup> Although this paper was written in 1981, at the time production of the *Handbook* was so out-of-date that material from the 1950s and 1960s was still being assessed and analyzed. So the procedures used were probably very similar to those in place from 1930 to 1960.

of the secondary sources need to be borne in mind when they are used for information retrieval.

The coverage of patents by secondary sources is affected by a range of factors: factors that affect production of primary literature on which the secondary sources depend, the editorial policy, and the overall aims of the secondary publications.

### Acknowledgments

I would like to thank Frank Cooke, GlaxoSmithKline, for providing access to MicroPatent data and help with the computational work; Sandy Lawson, MDL Information Systems, for information about the criteria used for selecting material for inclusion in *Beilstein*; Brian Lee and Joe Marsh, of the University of Manchester Institute of Science and Technology, for their advice and suggestions; and Paul Peters, representing CAS, for helpful suggestions concerning STN searching of *CA*.

### References

- Adams, S. (1999). The *Beilstein Handbook* as a source of pre-1960 chemical patent information. *World Patent Information*, 21, 75–81.
- Aftalion, F. (1991). *A history of the international chemical industry* (Ch. 4, pp. 102–206). Philadelphia: University of Pennsylvania Press.
- Arora, A. (1997). Patents, licensing, and market structure in the chemical industry. *Research Policy*, 26, 391–403.
- Baker, D. B., Horiszny, J. W., & Metanomski, W. V. (1980). History of abstracting at Chemical Abstracts Service. *Journal of Chemical Information and Computer Sciences*, 20, 193–201.
- Chemical Abstracts Service. (2002). Available: [www.cas.org](http://www.cas.org).
- Crane, E. J. (1944). Growth of the chemical literature: Contributions of certain nations and the effects of war. *Chemical and Engineering News*, 22, 1478–1481, 1496.
- Crane, E. J. (1958). *CA today: The production of Chemical Abstracts*. Washington, DC: American Chemical Society.
- de Solla Price, D. (1975). *Science since Babylon*. New Haven, CT/London: Yale University Press.
- Donnell, E. P. (1995). Growth and change in the world's chemical literature as reflected in *Chemical Abstracts*. *Publishing Research Quarterly*, 10(4), 38–46.
- Fleischer, J. (1951). Exploring the United States chemical patent literature. In *Searching the Chemical Literature, Advances in Chemistry Series, 4* (pp. 61–69). Washington, DC: American Chemical Society.
- Fleischer, J. (1952). The growth of chemical patents. *Chemical and Engineering News*, 30(3), 239–241.
- Gilfillan, S. C. (1960). An attempt to measure the rise of American inventing and the decline of patenting. *Technology and Culture*, 1(3), 201–214.
- Heller, S. R. (1998). *The Beilstein system: Strategies for effective searching*. Washington, DC: American Chemical Society.
- Huntress, E. H. (1938). 1938: The one hundredth anniversary of the birth of Friedrich Konrad Beilstein. *Journal of Chemical Education*, 15, 303–309.
- Jewkes, J., Sawers, D., & Stillerman, R. (1969). *The sources of invention* (2nd ed.). London: Macmillan.
- Luckenbach, R. (1981a). The *Beilstein Handbook of Organic Chemistry: The first hundred years*. *Journal of Chemical Information and Computer Sciences*, 21, 82–83.
- Luckenbach, R. (1981b). The critical screening and assessment of scientific results without loss of information—possible or not? *Angewandte Chemie (International Edition in English)*, 20, 841–849.
- Luckenbach, R. (1996). Past perfect, present perfect, future perfect—quality assessment and quality control mechanisms at Beilstein. *Journal of Chemical Information and Computer Sciences*, 36, 923–929.
- Luckenbach, R. (1997). The *Beilstein Handbook*. In S. R. Heller (Ed.), *The Beilstein system: Strategies for effective searching* (Ch. 2, pp. 13–28). Washington, DC: American Chemical Society.
- MDL Information Systems GmbH. (2002). Available: [www.mdl.com](http://www.mdl.com).
- Meehan, P., & Schofield, H. (2001). *CrossFire: A structural revolution for chemists*. *Online Information Review*, 15, 241–249.
- MicroPatent. (2002). Available: [www.micropatent.com](http://www.micropatent.com).
- Osborn, A. D. (1980). *Serial publications* (3rd ed., Ch. 2, pp. 24–48). Chicago: American Library Association.
- Pavitt, K. L. R. (1982). R&D patenting and innovative activities: A statistical exploration. *Research Policy*, 11, 33–51.
- Prager, B., & Jacobson, P. (1918). *Beilstein Handbuch der organischen Chemie* (Vierte Auflage, Band 1 [4th ed., vol. 1]). Berlin: Springer.
- Richter, F. (1938). How *Beilstein* is made. *Journal of Chemical Education*, 15, 310–316.
- Richter, F. (1958). *Beilstein Handbuch der organischen Chemie* (Vierte Auflage, drittes Ergänzungswerk, Band 1 [4th ed., 3rd suppl., vol. 1]). Göttingen/Heidelberg: Springer Verlag.
- Ridley, D. D. (2002). Information retrieval: SciFinder and SciFinder Scholar. Chichester, NY: John Wiley.
- Schmookler, J. (1960). An economist takes issue. *Technology and Culture*, 1(3), 214–227.
- Schummer, J. (1997). Scientometric studies on chemistry. I: The exponential growth of chemical substances 1800–1995. *Scientometrics*, 39, 107–123.
- Smith, J. (1965). Primary sources—literature on patents. In M. G. Mellon (Ed.), *Chemical publications: Their nature and use* (4th ed.). New York: McGraw-Hill.
- Special Libraries Association. (1944). *Patent Index to*

- Chemical Abstracts, 1907–1936. Ann Arbor, MI: Ann Arbor Science.
- STN International (2002). Available: [www.stn-international.de](http://www.stn-international.de).
- Thomas, E. (1936). Computing progress in chemistry. *Science*, 83, 59–61.
- U.K. Patent Office. (2002). Available: [www.patent.gov.uk](http://www.patent.gov.uk).
- U.S. Patent and Trademark Office. (2002). Available: [www.uspto.gov](http://www.uspto.gov).
- Wilson, R. E. (1937). Looking toward the future of invention. In *Centennial Celebration of the American Patent System; being a complete summary of speeches and data of the banquet and proceedings of the 100th anniversary of the American patent system, Washington, D.C., 23 November 1936* (pp. 20–27). GPO publication no. 112833. Washington, DC: U.S. Government Printing Office.
- Wolman, Y. (1988). *Chemical information: A practical guide to utilization* (2nd ed.). Chichester, NY: Wiley-Interscience.