J. Chem. Inf. Comput. Sci. **1990**, *30*, 33–35 

# 25 Years of *Science Citation Index*—Some Experiences†

RICHARD L. M. SYNGE

19 Meadow Rise Road, Norwich NR2 3QE, England

Received August 9, 1989

*Science Citation Index* (SCI) depends for intellectual content entirely on citations by authors, who are sometimes prodded by editors and referees. Its patchiness is therefore not surprising, but frequently it gives access to relevant and up-to-date documents not easily accessible by other means. Two contrasting "citation families" are described. The first family, dealing with the various ascorbic acid derivatives having C substitution at C-2, actually retrieved very nearly all the relevant documents (other than patent specifications) that were retrieved by a CAS ONLINE substructure search. Organic chemists are clearly careful authors. The second family, dealing with amino acid residues covalently bound in soil organic matter, yielded documents having surprisingly little overlap with those retrieved by using a carefully devised Boolean "profile" on the general subject index of *Chemical Abstracts*. This was only partly because SCI is beset by language-barrier problems to which *Chemical Abstracts* is immune. The SCI management might extend its journal coverage, but otherwise improvement can only come from a more serious attitude to placing references in primary publications. SCI remains a complement to, not a substitute for, other data-bases.

## INTRODUCTION

Eugene Garfield's brainchild is just about celebrating its 25th birthday, and I have watched it closely since its birth. *Science Citation Index* (SCI), like a child, has grown in capability as the years have passed, and the citations and cited documents have piled up. Over those years I have regularly run about 50 papers for citations—a group which has shifted gradually, following shifts in my own interests. On average, those papers have had about two citations per year, which suggests that they have been of about average interest.[1] What is immediately apparent is that the citing papers are a patchy collection—many papers that ought to have been among them are absent, and many of those present discuss aspects of the cited paper not of much interest to *me*. But many valuable documents come to hand, often newly published and not yet abstracted. To go through a 2-monthly issue of SCI takes me less than an hour, and to examine two of the retrieved documents each week is not very onerous.

Of course, some of the papers I run are my own publications, which get far less than the average two citations per year. But, when a citation does happen, it can be very heartwarming and sometimes leads to quick clearing up of a misunderstanding or, better, to personal contact with a fellow enthusiast.

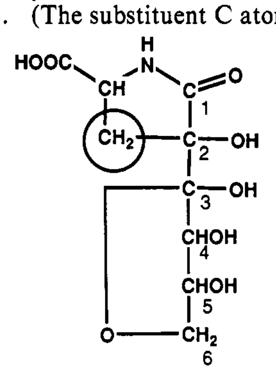
Apart from such regular use of SCI, SCI can on occasion be valuable for quick assessment of an unfamiliar field. The uses and effects of common chemicals is one of the main problems being discussed at this Conference: "Effects of ethanol ingestion on gout and uric acid excretion" is a good example. If one can locate, perhaps using a textbook, a really important document[2] of reasonable antiquity, SCI, with suitable "recycling", can open up a good range of the relevant information. These results may or may not be superior to Boolean searching of the various available data-bases, but they may, as I will go on to show, include documents that the data-bases have missed completely.

Enough of generalities, and I will leave aside the misguided ways in which SCI has been used as a substitute for genuine critical and personal assessment of candidates for jobs or promotions. What I want to do here is to describe the two extreme examples of the various families of "recycled" citations that I have handled over the years.

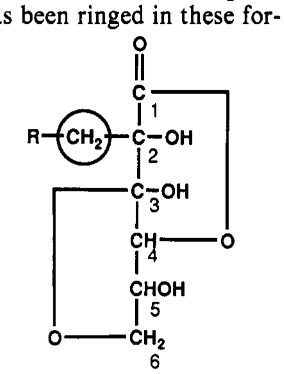
† Paper presented at Joint Conference of Chemical Information Group (Royal Society of Chemistry), Institute of Biology and Institute of Information Scientists on *Multidisciplinary Approaches to Searching for Information about Chemicals* (Oxford, England, March 25–6, 1986).

## ASCORBIC ACID DERIVATIVES WITH C–C SUBSTITUTION AT POSITION 2

When colleagues and I encountered a compound, ascorbalamic acid (**1**), of this type in leafy vegetables,[3] we fairly quickly related it to what was known[4] about the "ascorbigens" (**2**). (The substituent C atom has been ringed in these formulas.

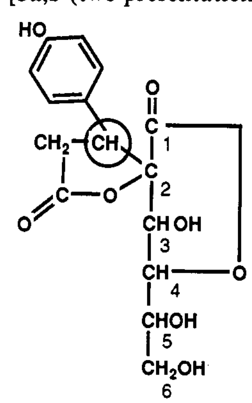



1

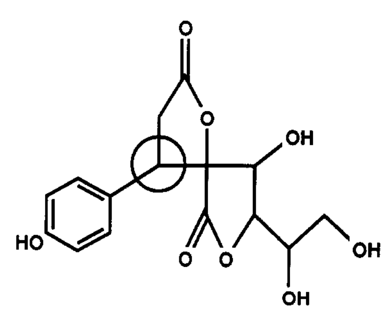



2

Numbering of C atoms is as in ascorbic acid. Stereochemistry is disregarded.) What we did not grasp was its relationship to a group of "*spiro*-dilactones", such as leucodrin [**3a,b** (two presentations)], found in S. African Proteaceae.[5]




3a




3b

We had the good fortune to meet Prof. H. Neukom (Zürich), who used to visit Norwich as a consultant. With his pupil G. Kiss, he had finally elucidated the structure of the ascorbigens after two different incorrect structures had been proposed by previous workers. Neukom told us that Kiss had mentioned

0095-2338/90/1630-0033$02.50/0 

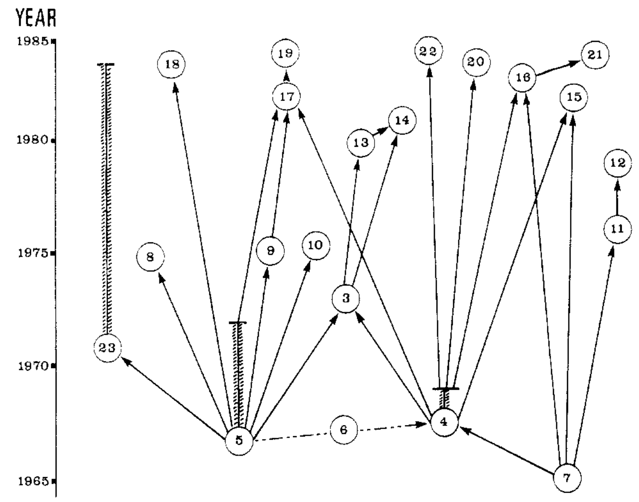
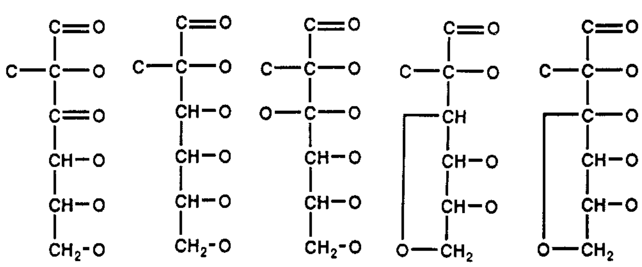



**Figure 1.** Citation family of references to ascorbic acid derivatives with C substitution at C-2. Numbers in circles relate to the reference list. Arrows indicate citation of earlier by later documents. The vertical hatched lines indicate publication period, where several documents are grouped under a single ref. For significance of dashed arrow and ref 6, see text.

the relationship to the S. African compounds in his Ph.D. Thesis,[6] though not in their joint journal publications.[4] I show this as a dotted line in the citation family (Figure 1), which I gradually built up over the years by adding to it each citing document which dealt with a relevant compound (other than the original "indolyl ascorbigen") and then recycling it. Besides the documents referred to in Figure 1, 67 others had to be looked at. Some of those, although irrelevant for present purposes, were in other respects interesting to study.

With CAS ONLINE substructure searches[24] having become publicly available for the post-1966 abstracts, I was curious to find out how many of the published compounds having the substructures of interest



had been retrieved by using SCI in this way. Dr. Peter Rhodes kindly ran searches for me, using the facilities of the Royal Society of Chemistry at Nottingham University. Besides compounds of the group in question, those searches retrieved a few simple hexonic acid derivatives, such as 2-*C*-methylgluconic acid, and also (perhaps of biological interest) some synthetic analogues of that important intermediate in photosynthesis, carboxyribulose diphosphate. Besides these, the searches found all but 1 of the 30 documents (1966 onward) in my citation family (Figure 1). This[15] described synthesis of a 2-*C*-phosphonic analogue, lacking a sought substructure but interesting nonetheless. CAS ONLINE additionally retrieved two relevant patent specifications[25] (a class of document not handled by SCI) and also two papers[26] from a journal not at that time covered by SCI.

The close overlap of these two independent approaches suggests that nearly all journal-published compounds may have been rounded up by SCI before CAS ONLINE was publicly available. I have colleagues who report similar experiences with other organic categories. Organic chemists are well-known to have highly developed territorial sensitivity, but that is surely not an unworthy emotion if it causes authors, referees, and editors to take real care about citing the literature. Moreover, access to that literature has, during the past century, been superbly mediated by *Beilstein*, *Chemisches Zentralblatt*, and *Chemical Abstracts*.

My other example shows a quite different state of affairs.

## AMINO ACID RESIDUES CHEMICALLY BOUND IN THE ORGANIC MATTER OF SOIL

I was first interested in this topic by Prof. R. D. Haworth (Sheffield) around 1965, and he gave me a number of useful references. But those that ultimately turned out useful as starting points in SCI[27–30] could all have been found in the 1971 review by Flaig.[31] I was specially interested in the effects of weathering on the amino acids, and the documents recycled in the citation family (Figure 2) had special relevance to this topic. However, let us imagine that I was trying to write a comprehensive review of the more general topic, starting with Flaig's 1971 review.[31] The various nodes in the trees, with recycling, yielded, in all, 83 post-1966 citing documents, of which 59 seemed relevant. A reasonable next step was to find out how the 13 nodal documents[32–42] published from 1967 to 1984 (the computerized period of *Chemical Abstracts*) were indexed by Chemical Abstracts Service (CAS), so that a search profile could be devised for yielding additional relevant references. This was kindly done for me by Dr. Rhodes, using the CAS ONLINE command ALL. The combination

AMINO ACIDS (biological studies.OR.
occurrence).AND.(FULVIC.OR.HUMIC.OR.
HUMUS.OR.PEAT.OR.SOIL(S))

could retrieve 8 of these 13 nodal documents—more could have been retrieved if ".OR.SEDIMENTS.OR.WATER" had been inserted, but that might have cast the net too widely. [(Added in 1989) It was well that we also ignored HUMIN.[43]]

Use of this search profile on *Chemical Abstracts* (1967–1984 inclusive) produced 226 documents. These were examined as originals when they were close to hand, but mostly through their abstracts. Forty-five documents were judged to be irrelevant, partly because the search terms did not distinguish free from bound amino acids and partly because "soil organisms" were sometimes included under "soil"; 181 relevant documents remained, and the surprise was that their overlap with the 59 relevant documents from SCI amounted to only 13, which included the 8 documents from which the search profile had been constructed. Furthermore, of the 59 relevant documents yielded by SCI, 5 had never been handled by CAS at all, despite their chemical interest.

One feels little confidence in being able to write a comprehensive review on this basis. One could proceed further by feeding the more important documents found through CAS back through SCI. And, to cover the failures of CAS to handle relevant material, such data-bases as those of the Commonwealth Agricultural Bureaux and BIOSIS could be searched, as other contributors to this Conference described.

There is, moreover, an anomaly in the language distributions of the documents, as shown in Table I. One must assume that the language distribution of the CAS documents is fairly representative of the world literature—the high proportion of Russian documents is not surprising, when one considers the longstanding interest in soil science in that country. The fact that the SCI documents are so predominantly in English must arise from some combination of (i) poor coverage of non-English-language literature by SCI and (ii) poor citation of English-language documents by authors not writing in English.

## CONCLUSION

The intellectual content of SCI depends entirely on generous inclusion of references, particularly in primary publications.

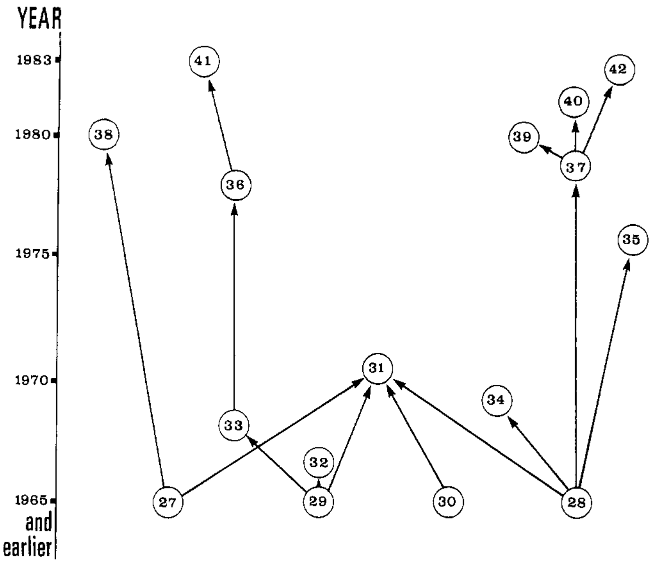



**Figure 2.** Citation family of references to bound amino acids of soil organic matter. Numbers in circles relate to the reference list. Arrows indicate citation of earlier by later documents. Reference 31 was not recycled, because it was too general a review.

**Table I.** Language Distributions of Two Sets of Documents Relating to Amino Acids of Soil Organic Matter (1967–1984)

| | retrieved through | |
|---|---|---|
| | SCI | CAS |
| English | 70 | 85 |
| French | 8 | 8 |
| German | 2 | 7 |
| Russian | 0 | 97 |
| other | 3 | 29 |
| totals | 83 | 226 |

References do not take up much space. The quality of SCI would be much improved if authors, referees, and editors emulated the organic chemists, taking a more serious attitude to references and making a bigger effort to break down language barriers.

SCI is unlikely ever to replace other data-bases, but its value as a complement to them can be increased, and this can only be done by the individual efforts of a lot of people. Such people are all potential users of SCI, so an extra effort at including references in the publications for which they are responsible need not be mere altruism.

## ACKNOWLEDGMENT

I am grateful to Dr. Peter Rhodes and to his employers, The Royal Society of Chemistry, for helping me with the CAS facilities available to them. I also thank Cambridge Scientific Computing Inc. for making available their ChemDraw program for the graphic formulas and Dr. Richard Taylor (School of Chemical Sciences, University of East Anglia) for instructing me in its use.

## REFERENCES

(1) *Science Citation Index 1982 Guide*, pp 26–7.
(2) Lieber, C. S.; Jones, D. P.; Losowsky, M. S.; Davidson, C. S. *J. Clin. Invest.* **1962**, *41*, 1863–70.
(3) Couchman, R.; Eagles, J.; Hegarty, M. P.; Laird, W. M.; Self, R.; Synge, R. L. M. *Phytochemistry* **1973**, *12*, 707–18.
(4) Kiss, G.; Neukom, H. *Helv. Chim. Acta* **1966**, *49*, 989–92; *Mitt. Geb. Lebensmittelunter. Hyg.* **1966**, *57*, 443–8; *Experientia* **1968**, *24*, 326.
(5) Perold, G. W.; Pachler, K. G. R. *J. Chem. Soc. C* **1966**, 1918–23. Murray, A. W.; Bradshaw, R. W. *Tetrahedron* **1967**, *23*, 1929–37, 2333–8. Perold, G. W.; Hodgkinson, A. J.; Howard, A. S. *J. Chem. Soc., Perkin Trans. 1* **1972**, 2450–7. Perold, G. W.; Hodgkinson, A. J.; Howard, A. S.; Kruger, P. E. J. *Ibid.* 2457–60.
(6) Kiss, G. *Über die Struktur des Ascorbigens, Diss. No. 5015, Eidg. Tech. Hochschule*; Juris Druck & Verlag: Zürich, 1967.
(7) Jackson, K. G. A.; Jones, J. K. N. *Can J. Chem.* **1965**, *43*, 450–7.
(8) Brimacombe, J. S.; Murray, A. W.; Zahur-ul-Haque. *Carbohydr. Res.* **1975**, *45*, 45–53.
(9) Lowry, J. B.; McAlpine, J. B.; Riggs, N. V. *Aust. J. Chem.* **1975**, *28*, 109–18.
(10) McAlpine, J. B.; Riggs, N. V. *Aust. J. Chem.* **1975**, *28*, 211–5.
(11) Hvoslef, J.; Nordenson, S. *Acta Crystallogr. Sect. B* **1976**, *32*, 1665–9.
(12) Hvoslef, J.; Pedersen, B. *Acta Chem. Scand. Ser. B* **1979**, *33*, 503–11.
(13) Kasai, T.; Okuda, M.; Sakamura, S. *Agric. Biol. Chem.* **1980**, *44*, 2723–4.
(14) Kasai, T.; Okuda, M.; Sakamura, S. *Phytochemistry* **1981**, *20*, 1131–2.
(15) Sekine, M.; Futatsugi, T.; Hata, T. *J. Org. Chem.* **1982**, *47*, 3453–6.
(16) Fodor, G.; Arnold, R.; Mohacsi, T.; Karle, I.; Flippeu-Anderson, J. *Tetrahedron* **1983**, *39*, 2137–45.
(17) Yvin, J. C.; Chevolot-Magueur, A. M.; Chevolot, L.; Lallemand, J. Y.; Potier, P.; Guilhem, J. *J. Am. Chem. Soc.* **1982**, *104*, 4497–8.
(18) Iwagawa, T.; Hase, T. *Phytochemistry* **1984**, *23*, 2299–301.
(19) Seebach, D.; Dust, M.; Naef, R.; Baenziger, M. *Angew. Chem.* **1984**, *96*, 498.
(20) Mukhanov, V. I.; Yartseva, I. V.; Kikot, B. S.; Volodin, Yu. Yu.; Kustova, I. L.; Lesnaya, N. A.; Sofina, Z. P.; Ermakova, N. P.; Preobrazhenskaya, M. N. *Bioorg. Khim.* **1984**, *10*, 433–59.
(21) Fodor, G.; Sussangkarn, K.; Mathelier, H.; Arnold, R.; Karle, I.; George, C. *J. Org. Chem.* **1984**, *49*, 5064–9.
(22) Hanley, A. B.; Belton, P. S.; Fenwick, G. R.; Janes, N. F. *Phytochemistry* **1985**, *24*, 598–600.
(23) Perold, G. W.; Hundt, H. K. L. *J. Chem. Soc. D* **1970**, 712. Kruger, P. E. J.; Perold, G. W. *J. Chem. Soc. C* **1970**, 2177–33. Perold G. W.; Howard, A. S.; Hundt, H. K. L. *J. Chem. Soc. C* **1971**, 3136–40. Highet, R. J.; Perold, G. W.; Sokoloski, E. A. *J. Org. Chem.* **1976**, *41*, 3860–2. Glennie, C. W.; Perold, G. W. *Phytochemistry* **1980**, *19*, 1463–6. Boeyens, J. C. A.; Naude, M. J.; Perold, G. W. *S. Afr. J. Chem.* **1982**, *35*, 157–60. Perold, G. W. *J. Chromatogr.* **1984**, *291*, 365–7.
(24) Dittmar, P. G.; Farmer, N. A.; Fisanick, V. W.; Haines, R. C.; Mockus, J. *J. Chem. Inf. Comput. Sci.* **1983**, *23*, 93–102.
(25) Koppel, G. A.; Barton, R. L.; Briggs, S. L. *Chem. Abs.* **1985**, *102*, 113888p. Szent-Györgyi, A.; Fodor, G. B.; Veltri, R. W. *Ibid.*, 132016u.
(26) Glennie, C. W. *J. S. Afr. Bot.* **1979**, *45*, 185–91; *Ibid.* **1980**, *46*, 99–107.
(27) Stevenson, F. J. *Proc. Soil Sci. Soc. Am.* **1956**, *20*, 201–4, 204–8.
(28) Carles, J.; Soubies, L.; Gadet, R. *C. R. Acad. Sci. Paris*, **1958**, *247*, 1229–32. Carles, J.; Decau, J. *Sci. Proc. R. Dublin Soc. Ser. A1* **1960**, No. 4, 177–82.
(29) Scharpenseel, H. W., Krausse, R. *Z. Pflanzenernaehr., Dueng., Bodenkd.* **1962**, *96*, 11–34.
(30) Yamashita, T.; Akiya, T. *Nippon Dojo Hiryogaku Zasshi* **1963**, *34*, 255.
(31) Flaig, W. *Soil Sci.* **1971**, *11*, 19–33.
(32) Wang, S. C.; Yang T. K.; Cheng, S. Y. *Soil Sci.* **1967**, *103*, 67–74.
(33) Piper, T. J.; Posner, A. M. N. *Soil Sci.* **1968**, *106*, 188–92.
(34) Decau, J.; Pujol, B. *Ann. Agron.* **1969**, *20*, 35–59; *Ibid.*, 277–303.
(35) Sowden, F. J.; Griffith, S. M.; Schnitzer, M. *Soil Biol. Biochem.* **1976**, *8*, 55–60.
(36) Carter, P. W.; Mitterer, R. M. *Geochim. Cosmochim. Acta* **1978**, *42*, 1231–8.
(37) Goh, K. M.; Edmeades, D. C. *Soil Biol. Biochem.* **1979**, *11*, 127–32.
(38) Casagrande, D. J.; Given, P. H. *Geochim. Cosmochim. Acta*, **1980**, *44*, 1493–507.
(39) Holtzclaw, K.; Schaumberg, G. D.; Le Vesque-Madore, C. S.; Sposito, G.; Heick, J. A.; Johnston, C. T. *Soil Sci. Soc. Am. J.* **1980**, *44*, 736–40.
(40) Preston, C. M.; Mathur, S. P.; Rauthan, B. S. *Soil Sci.* **1981**, *131*, 344–52. Preston, C. M. *Can. J. Soil Sci.* **1982**, *62*, 479–86.
(41) Gonzalez, J. M.; Grimalt, J.; Albaiges, J. *Mar. Chem.* **1983**, *14*, 61–71.
(42) Piispanen, R.; Lähdesmäki, P. *Soil Biol. Biochem.* **1983**, *15*, 381–3.
(43) Rice, J. A.; Maccarthy, P. *Geoderma* **1988**, *43*, 65–73.