H.J.E. LOEWENTHAL

Guide for the Perplexed Organic Experimentalist

HEYDEN

Guide for the Perplexed Organic Experimentalist

H. J. E. LOEWENTHAL

Chemistry Department,
Technion—Israel Institute
of Technology

HEYDEN

London . Philadelphia . Rheine

Heyden & Son Ltd., Spectrum House, Hillview Gardens, London NW4 2JQ, UK
Heyden & Son Inc., 247 South 41st Street, Philadelphia, PA 19104, USA
Heyden & Son GmbH, Münsterstrasse 22, 4440 Rheine/Westf., Germany

ISBN 0 85501 169 6

Type set by John Wright & Sons Ltd, Bristol and printed in Great Britain by Henry Ling Ltd, The Dorset Press, Dorchester

Preface

The perplexed organic experimentalist is in my experience the beginning research student and (frequently) the post-doctoral research worker. He is the one who early on discovers that he has to stand on his own two feet, in the task of searching for information on his subject and in all the practical aspects of his work—and that means not only how to run a reaction but also how to choose and acquire the tools and materials of his trade.

All too often it is not by choice that he finds himself in this situation. His supervisor (or, more politely, his 'Senior Collaborator') was himself once a graduate student and post-doctoral researcher. But in the majority of cases he has long since abandoned the laboratory bench and is now busy with administration, with writing and refereeing research grant applications and scientific papers, and with teaching and thinking. In the process he will have forgotten most of the practical knowledge which he had to acquire painfully in his own day and will have become unaware of later developments.

The average graduate student is ill-prepared for searching the literature. Most practical textbooks will have done little to train him to think for himself. Few prepare him for continuing preparative work on a small scale, and fewer still for working with sensitive reagents and under dry and anaerobic conditions. None, so far as I know, do anything to assist him (or his supervisor) to grapple with indifferent suppliers, manufacturers and administrators.

This small book attempts to fill some of the gaps, on the basis that organic chemistry is an experimental science first and foremost, and that in the final analysis—with all of modern instrumentation and computerization—it is the organic chemist's brain and own two hands that are and will remain indispensable. Hence there is little on instrumental spectroscopic and analytical techniques and other special fields, which are quite adequately dealt with elsewhere. The important matter of safety is, with regret, touched upon only in brief; to do proper justice to it would be beyond the scope of this book. Workers in certain areas such as carbohydrates or peptides may find comparatively little of special relevance to them. Others will no doubt take exception to a number of statements in this book. I shall be happy if their misgivings will induce them all to write a similar book of their own.

I am indebted to many friends and colleagues for constructive advice and criticism, and above all to Professor R. A. Raphael, F.R.S. (University of Cambridge), for his encouragement and great help, and to Professor O. Jeger (Swiss Federal Institute of Technology, Zurich) whose hospitality enabled me to get started on this work.

Haifa,
January 1978

H. J. E. LOEWENTHAL

Contents

1
On Searching the Literature

This chapter comes first because that is where it ought to be. No serious programme of research, chemical or otherwise, is ever embarked upon before a thorough investigation of the current state of knowledge of the subject has been undertaken.

It is instructive to examine the way in which the use of the primary and secondary literature is treated in the teaching of organic chemistry nowadays. Inspection of 30 textbooks in English, published or revised since 1960, shows that 25 of them make practically no mention of the subject at all and of these 15 do not even give a bibliography or reading list. All the remaining five relegate it to the end of the book, which will hardly impress its importance upon the student. Only three of the latter make a serious effort to guide the student in the use of that literature. These findings, incidentally, are unrelated to whether the book is called simply 'Organic Chemistry' or whether the title is qualified by terms such as 'Basic', 'Introduction to . . .',

'Modern', 'Comprehensive' or 'Advanced'. Surely this state of affairs ought to raise some searching questions on the education of the organic chemist in the age of the information explosion.

There are books which deal with the subject specifically; among these the ones by Crane *et al.*,[1] by Dyson,[2] by Bottle,[3] by Burman,[4] and the A.C.S. Monograph on the subject[5] can be mentioned; and anyone wanting to study the subject in depth should use these. However, they are designed to cover the literature of chemistry as a whole, and none are up to the latest developments.

The object of this chapter is to provide some hints and guidance to the beginning organic research chemist, and in particular to point out some of the snags and pitfalls he is bound to encounter.

SEARCHING FOR A COMPOUND, ITS SYNTHESIS AND PROPERTIES

The most frequent object of a literature search consists of finding out whether a certain compound is known, how it was prepared and what its properties are (or appear to be). In this, and indeed in any other search, one tries to find everything known in the past up to the point where one's own work should start. In fact, in most cases the search will determine just what that starting point should be.

It should follow that any such search should be made in reverse chronological order, i.e. starting with the latest Index of *Chemical Abstracts*, the most recently issued series of *Beilstein's Handbuch* or the latest review, and then to work backwards in time on the reasonable assumption that whosoever worked last on the point in question will have gathered most if not all the previous information for you. As obvious as this may seem it is mentioned all the same; no one seems to have stressed it particularly before and in fact much of the customary advice (as much as there is of it) implies doing the whole process in the opposite direction. This does not mean that having arrived at the first past author who appears to have reviewed prior endeavours reasonably well you should consider your work

done. As in driving a car, always consider the other fellow worse than you are and then set out to prove it. In most cases, if you dig deep enough you will surely discover some piece of prior and important information which he has missed and which will put you at an immediate advantage.

The beginner is tempted to waste his time placing too much reliance on various handbooks, compendia, 'dictionaries' and 'encyclopedias' found in most libraries and laboratories. For any serious search these are incomplete and liable to be inaccurate, and where literature references are given it is not always clear just to what particular point they do refer.

Chemical Abstracts

This is the most important source of information on chemical compounds, at any rate from the present back to the year 1949 when in the majority of cases *Beilstein's Handbuch* should take over (see below). For this reason that part of any library which houses this work is the most frequented one and that area should be reserved for *Chemical Abstracts* readers only. It should also go without saying that no volume of this work should ever leave that area and that its binding should receive special care.

The most important parts are the *Indexes*: the *Collective* ones in so far as they have appeared and the *Semi-annual* ones if not. Since 1971 the *Subject Indexes* have been sub-divided into a *General Subject* and a *Chemical Substance Index*. Of comparable importance are the *Formula Indexes*; the *Author Index* is of course of secondary importance in this connection.

The main problems encountered by the beginning searcher are those of Nomenclature, of Evaluating the Entry and of Evaluating the Abstract.

Nomenclature

In this *Chemical Abstracts* have gone their own way over the years, undaunted by IUPAC and other rules. The ever-increasing use of computer search and printing does of course provide a generally convincing rationale for this.

A few examples will serve to show what has happened in the course of time:

(a) Up to 1946: α-Toluic acid
1947–71 : Acetic acid, phenyl
Now : Benzeneacetic acid
(b) Up to 1971: Acetic acid, chloroformyl, ethyl ester
Now : Propanoic acid, 3-chloro-3-oxo, ethyl ester
(c) Up to 1971: o-Anisidine
Now : Benzeneamine, 2-methoxy
(d) Up to 1971: Norephedrine
Now : Benzenemethanol, α-(1-aminoethyl)
(e) Up to 1946: Alanine, *N*-acetyl-β-phenyl
1947–71 : Alanine, *N*-acetyl-3-phenyl
Now : Phenylalanine, *N*-acetyl

Compare with the corresponding *N*-phthaloyl derivative:

(f) Up to 1971: 2-Isoindolineacetic acid, α-benzyl-1,3-dioxo
Now : 2*H*-Isoindole-2-acetic acid, 1,3-dihydro-1,3-dioxo-α-(phenylmethyl)

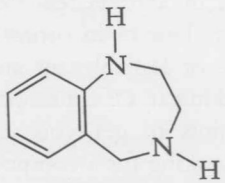
The tendency of the changes introduced recently is ruthlessly to cut out names considered 'trivial' and thus eliminate ambiguity as far as possible. This is indeed true in a case such as example (b). When searched for in previous years in the Formula Index under $C_5H_7ClO_3$ one would have come across something like 'malonaldehydic acid, chloro, ethyl ester', which at least at first glance would have looked like being the right thing. The new name, unreasonable as it seems and indeed is from a functional point of view, leaves no room for doubt. As against that the new 'Phenylalanine' [example (e)], while perhaps 'generally accepted' and hence 'justifiably trivial' is certainly not unambiguous.

One could go on and on with further examples of often bizarre changes in nomenclature, but the above should suffice to bring home the importance of being on one's guard all the time. Fortunately the latest changes have been accompanied by the issue of a handy and separate Index Guide which copiously lists and cross-references all new names and old. *This must be*

used when in any doubt whatever—all too often a literature search turns out to be woefully incomplete simply because changes in nomenclature have not been taken into account.

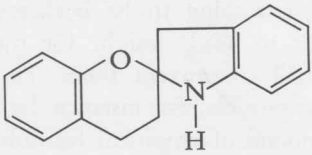
The beginner, not wishing to be bothered with all these complications, tends to head straight for the *Formula Index* which he thinks will circumvent them. He should be disillusioned as early as possible. For instance, let him search there for a specific compound of empirical formula, say, $C_{10}H_{14}O$, and then let us see how he will cope with wading through, and choosing the right one from, more than 500 alternatives. Or, suppose he has to look for a hydroxy-dicarboxylic acid $C_{32}H_{50}O_5$ and compounds related to it. He will probably and eventually find the one he is looking for, but what about its mono- and dimethyl or ethyl or benzyl esters, the corresponding *O*-acetates, -benzoates or -propionates, any one or more of which may be better known, more completely characterized or more easily accessible than the parent acid? In the *Subject* or *Chemical Substance Index* they will all be grouped together, but in the *Formula Index* each one of them has to be looked for separately, and may even not be found at all since it is known that this Index is often incomplete.

Here one ought to mention another aid to compound search which by contrast the beginner seems to avoid as much as he can even though it is quite simple to use: the *Ring Index*. This lists the systematic names of ring systems according to (a) the number of discernible rings, (b) the size of each, and (c) the nature and number of heteroatoms in each ring—arranged according to the alphabetical Hill system. To give examples: (i) the following compound



has two rings, one six- and the other seven-membered, the first has only carbons atoms and the second two nitrogen atoms.

It will appear in the ring index under: two rings; 6,7; $C_6—C_5N_2$, and the corresponding basic system will appear under the name 1,3-Benzodiazepine. (ii) A more complex example:



has four discernible rings, will be classed as: 5,6,6,6 and $C_4N—C_5O—C_6—C_6$, and following up the Ring Index under this will be found the name Spiro [2H-1-benzopyran-2,2′-indoline]. (iii) A steroid having a two-oxygen bridge between positions 5 and 7 will on inspection be seen to incorporate five rings; and since the two-oxygen bridge is common to two of them the compound must be classed as $C_5—C_4O_2—C_4O_2—C_6—C_6$. A possible and ultimately found to be correct candidate is 'cholestane, 5,7-epidioxy'. The old *Ring Index*, published separately with supplements until the early sixties, had the advantage of showing actual structures in each instance. Since 1962 it has formed part of the *Chemical Abstracts* Indexes (starting from the *Collective* one for 1962–66); and for structures no longer being shown it is necessary to refer back and find the appropriate one in the *Subject* or *Chemical Substance Index*.

Evaluating the entry

This and the following section bring home the fact that hundreds of people are involved in preparing *Chemical Abstracts* and that human nature and subjectivity can play a part.

The entry for a compound occasionally carries no topic, but generally one is given in abbreviated form. The art lies in interpreting its wording alone or in context, and thus deciding which entries are more or less relevant and which abstract is worth or not worth looking at. Of course practice makes perfect, but herewith a few hints to get you started assuming, for example, that you are looking for a compound's synthesis:

1. Under the compound are also listed its functional derivatives (esters, salts, complexes, oximes, hydrazones, etc.). The

relevance of these will depend on how easily these derivatives can be converted into the desired compound or whether they can serve as a reasonable substitute. Further evaluation should proceed as below.

2. A number of entries may refer to the same abstract. This may be significant on its own and making a note of it immediately will save time.
3. No topic: in the end this may be your best bet. Often it means that the paper carries so much information that the abstracter and indexer preferred not giving a topic at all.
4. 'preparation of': in my experience this is prone to refer to a method by which the compound was indeed prepared but which turns out to have little preparative value.
5. 'raman spectrum of', 'nuclear magnetic resonance spectrum of', 'dipole moment of': if this is an unduplicated entry it may refer to the work of someone (spectroscopist or other physical chemist) who obtained the small amount of compound required from a colleague and was possibly too busy arranging his data to make it or even give a literature reference to it.
6. 'bromination of', 'oxidation of', 'reaction of, with phenylmagnesium bromide': an excellent prospect; clearly work done by an organic chemist who had to make the compound himself in larger quantities.
7. 'carcinogenic activity of': possibly useless.
8. 'in grapefruit seed': probably useless.
9. 'bond length in': utterly useless.

Evaluating the abstract

There was a time when the abstract would refer directly to all new compounds made, their melting or boiling points and possibly other physical properties, and outline details of their preparation. Evaluation was then based in the main on how trustworthy one considered the authors and the journal concerned. This is a subject on which I prefer not to make any comment at all here. But those days have gone and the average abstract, often leaning heavily on the author's own summary,

has become not much more than another shunting point *en route* to the original publication. Evaluation is thus reduced to choosing which journal to look at at all, and which to look at first. If it happens to be found in your own library the decision is simple. If not, the question arises whether to take the next train to wherever publications such as *Farmaco, Ed. Sci.*; *Khim. Geterosikli Soedin.* or *Sb. Pr. Pedagog. Fak. Ostrave Rada E.* can be found. This should not even be contemplated unless you know the language; and ordering a Xerox copy having located someone who does is not always a solution. There are few experiences as frustrating as trying to translate an article in Yakugaku Zashhi with the help of a Japanese lecturer of Social Economics.

An important thing to ascertain is whether the article is in the form of a Communication and thus liable to give little experimental detail. If reference is to *Angew. Chem., Chem. Lett., Chimia, J. Chem. Soc., Chem. Commun.* or *Tetrahedron Letters* the situation is clear, but not if to *J. Am. Chem. Soc.* or *J. Org. Chem.* The last two are likely to give a modicum of experimental facts, the first four less so. As far as *J. Am. Chem. Soc.* is concerned you will soon find out that certain authors publish *nothing but* Communications or rarely a full paper. Here the prospect of ever finding fuller detail is not bright at all unless the senior author will condescend to send it to you after you have written to him personally.*

The fact that the reference is to a journal which from its title is not strictly in the organic chemistry field need not deter you. Some of the best and most detailed information on synthesis and preparation can turn up in journals such as *Agric. Biol. Chem., Phytochemistry* and *J. Organomet. Chem.* which cover adjacent fields in which the majority of authors started their career as organic chemists, and it is only natural for them to show that they have not lost their touch and expertise.

* And when you write do not forget to check whether there has not been a change in address from that given in the article, lest you commit the unpardonable offence of not knowing that the man has moved to a more prestigious institution.

And now a few comments on cases where the abstract is of a patent. This is of course a special problem when you are nowhere near a patent library, and when the abstract gives no detail worthy of consideration. By and large I would say that it is definitely advisable to order (by air mail) a copy of the patent, assuming you have made quite sure that the information in question has not subsequently been published in the regular form (and this does happen). I do not think that automatically turning up one's nose at a patent reference is justified, and I can think of a number of cases in my own experience where in the preparation of a rather tricky kind of compound following a patent, and not a paper in a reputable journal, led to success. Moreover, there are many fields and whole groups of compounds which because of incidental commercial importance are not covered anywhere *but* in the patent literature. In the end, however, you should keep in mind always that patents are written by people who first and foremost are lawyers and not scientists, and that their prime object is not dissemination of knowledge for its own sake but the staking of a legal claim—and this usually by not giving away more information than is absolutely necessary for that specific purpose.

Beilstein's Handbuch

This is a unique work, and no self-respecting organic chemist can afford to be unfamiliar with its use. It is of course written in German, but once the English-speaking reader has come to terms with this he cannot but admire the perfectly thorough and logical system by which information is organized, and which has not changed from the early beginning to the present day—this is perhaps the most outstanding characteristic of the work. An excellent starting point for getting aquainted with it is the schematic diagram which appears in the textbook by Hendrickson *et al.*[6] and the chapter by Owen and Rickett[7] which appears in Bottle's book; after this the recent guide byWeissbach[8] should be consulted; this incidentally contains a useful glossary of German words and their translation into English and French.