

- (15) Waltersson, K. *Acta Crystallogr.* **1978**, *A34*, 901-905.
 (16) Brown, I. D. *Solid State Ionics* **1988**, *31*, 203-208.
 (17) Garrett, J. D.; Greedan, J. E.; Faggiani, R.; Carbotte, S.; Brown, I. D. *J. Solid State Chem.* **1982**, *42*, 183-190.
 (18) Villiger, H. *DLS Manual*; Inst. Krist. Petr. ETH: Zurich, 1969.
 (19) Brown, I. D. *J. Solid State Chem.* (in press).
 (20) Brown, I. D. *Acta Crystallogr.* **1976**, *A32*, 24-31.
 (21) Donnay, G.; Allman, R. *Am. Mineral.* **1970**, *55*, 1003-1015.
 (22) Brown, I. D. *Chem. Soc. Rev.* **1978**, *7*, 359-376.
 (23) Mitchell, K. A. R.; Schlatter, S. A.; Sodhi, R. N. S. *Can. J. Chem.* **1986**, *64*, 1435-1439.
 (24) Ziolkowski, J. *Thermochim. Acta* **1987**, *110*, 333-336.
 (25) Kahn, A. A. *Acta Crystallogr.* **1976**, *A32*, 11-16.
 (26) Hazen, R. M.; Prewitt, C. T. *Am. Mineral.* **1977**, *62*, 309-315.
 (27) Badger, R. M. *J. Chem. Phys.* **1934**, *2*, 128-131; **1935**, *3*, 710-714.
 (28) *Acta Crystallogr.* **1988**, *A44*, 232.
 (29) Sheldrick, G. SHELX76, Program for Crystal Structure Determination; University of Cambridge.

Problem Solving with the *Beilstein Handbook*[†]

REINER LUCKENBACH* and JOSEF SUNKEL

Beilstein Institute, Varrentrappstrasse 40-42, D-6000 Frankfurt/Main 90, West Germany

Received March 2, 1989

Beilstein is unique among handbooks of organic chemistry in that it provides a collection of critically examined and carefully reproduced data on the known organic compounds. In this respect, *Beilstein* is superior to all other straight bibliographical documentation and series of abstracts. Moreover, the *Beilstein Handbook* is the world's most extensive collection of physical data on organic compounds in printed form. While previous publications mainly concentrated on how to search for specific compounds in the *Handbook*, this paper provides some representative examples of how to find specific information and factual data for the following topics: nomenclature; physical data of "families" of chemical compounds; stereochemical assignments based on physical data; interpolation and extrapolation of physical data; oxidation reactions of a given substance; development of synthesis strategies for known and unknown compounds; "partial structure retrieval" ("similarity searches").

INTRODUCTION

The "Beilstein Information System" is the collective term for the database Beilstein Online and the printed reference work *Beilstein Handbook of Organic Chemistry*. Together they form the most comprehensive source of data and information on organic compounds to be found anywhere.

The purpose of this paper is to provide some practical hints to optimize the use of the Beilstein Information System. In particular, the use of "intuitive guesswork" to derive information from the data available is illustrated with practical examples. Although all the material presented here is taken from the *Beilstein Handbook*, many of the ideas involved are just as relevant to database searches.

1. THE BEILSTEIN HANDBOOK¹⁻¹¹

The *Beilstein Handbook of Organic Chemistry*, commonly referred to as *Beilstein*, is a collection of important published data on the preparation and properties of organic compounds. With well over 350 single volumes comprising more than 270 000 printed pages, the *Beilstein Handbook* today constitutes the world's largest collection of physical data on organic compounds in printed form!

The *Handbook* is published in series, each covering the literature of a certain period (see Table I). All volumes from E V onward are published in English.

Each series consists of 27 (nominal) volumes, almost all of them subdivided into several subvolumes in the more recent series, in which the individual compounds are arranged according to the Beilstein System.^{1,7,8} This system classifies all organic compounds according to their structure, and each volume corresponds to a particular class of compounds (see Tables II and III).

[†] Presented at the 197th National Meeting of the American Chemical Society, Dallas, TX, April 1989.

Table I. Series of the *Beilstein Handbook*

series	abbrev	period of literature covered	color of label on spine
<i>Basic Ser.</i>	H	up to 1910	green
<i>Suppl. Ser. I</i>	E I	1910-1919	dark red
<i>Suppl. Ser. II</i>	E II	1920-1929	white
<i>Suppl. Ser. III</i>	E III	1930-1949	blue
<i>Suppl. Ser. III/IV</i>	E III/IV ^a	1930-1959	blue/black
<i>Suppl. Ser. IV</i>	E IV	1950-1959	black
<i>Suppl. Ser. V</i>	E V	1960-1979	red

^a Volumes 17-27 of *Suppl. Ser. III* and *IV*, covering the heterocyclic compounds, are combined in a joint issue.

Table II. Main Divisions of the *Beilstein Handbook*

main division	vol. no.
(1) acyclic compounds	1-4
(2) isocyclic (carbocyclic) compounds	5-16
(3) heterocyclic compounds	17-27

The various *Beilstein* volumes contain the research data reported within the literature over a specific period, critically sifted and correlated in a reliable and logical form and assessed in the light of current chemical knowledge. The editors are careful to point out errors in the published data, to direct attention to the doubtfulness of some published statements, and to check assertions of speculative nature against subsequent findings. This frequently involves citing very recent publications.

To the organic chemist as well as to every scientist working in the field of organic chemistry, *Beilstein* is an indispensable source of information for which there is no substitute. The alert user will find it can save him from embarking on false trails and stimulate him to further research.

For each compound described in *Beilstein* the following aspects are covered: structural formula; compound name(s); molecular formula; constitution and configuration; natural

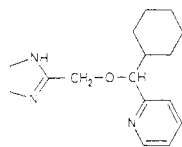
Table III. Heteroclasses of the *Beilstein Handbook*

heteroclass (type and no. of ring heteroatoms)	vol. no.
(1) 1 O atom	17, 18
(2) 2 and more O atoms	19
(3) 1 N atom	20-22
(4) 2 N atoms	23-25
(5) 3 and more N atoms	26
(6) 2 and more different (!) ring heteroatoms, e.g.	27
1 N + 1 O, 1 N + 2 O, ...	
2 N + 1 O, 2 N + 2 O, ...	
other heteroatoms, e.g., B, Si, P	

If there is any reference to your compound within BEILSTEIN, it will be in:

Volume (Band)-No. 23

To find your compound, now consult either the Name and Formula Indexes (Sach- and Formelregister) in the above volume - these are to be found at the end of each part volume (Teil) - or the corresponding Cumulative Indexes (Gesamtregister).



E III/IV 23, 2398

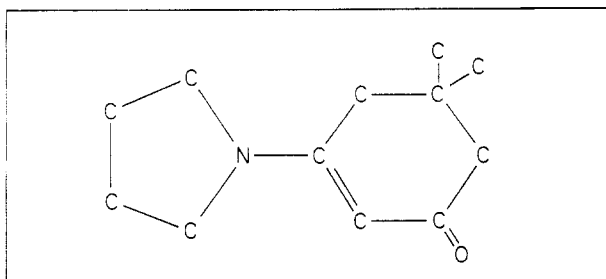
Another Compound (Y/N)?

Figure 1. Example of BEILSTEIN KEY output.

H - Page 4 to 7; Syst. - No. 3037 to 3038

Stammverbindung, 2n+1, C: 4

Another molecule (Y/N)?



E III/IV subvolume(s): 20/1 C 12 H 19 NO

E V subvolume(s): 20/1 Copyright 1987 by Beilstein, Frankfurt.

Figure 2. Example of SANDRA output.

occurrence; isolation from natural products; preparation and purification; structural and energy parameters of the molecule;

physical properties; chemical properties (reactions); characterization and analysis; salts and addition compounds. Altogether more than 350 different kinds of physical data are found in the *Handbook*.

2. HOW TO FIND SPECIFIC COMPOUNDS IN THE BEILSTEIN HANDBOOK

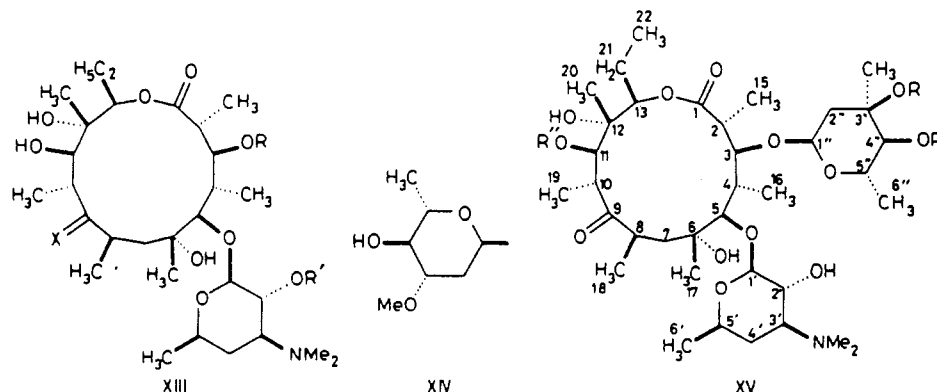
There are at least six different, efficient ways for the user to find the compounds in which he is interested in *Beilstein*: (i) *Beilstein* references in other handbooks and chemical catalogs; (ii) tables of contents of *Beilstein* volumes; (iii) indexes (Compound Name Indexes, Molecular Formula Indexes); (iv) application of the rules of the Beilstein System; (v) BEILSTEIN KEY; (vi) SANDRA. Since method i is self-explanatory and methods ii-iv have been explained in detail elsewhere,^{1-3,7,8,10} the results obtained by using the computer programs BEILSTEIN KEY and SANDRA are presented here.

2.1 Beilstein Key. The BEILSTEIN KEY, which is written in MICROSOFT BASIC and available free of charge from the Institute, helps the *Beilstein* user locate the volume containing a particular organic compound by means of a simple question and answer routine that does not involve any graphics input. BEILSTEIN KEY runs under MS-DOS on any IBM PC (and compatibles) and is a dialogue program that takes the user through a simple step-by-step analysis of the molecule of interest and enables him to identify the relevant volume of the *Handbook*. The program always (!) generates a volume number, even when the compound in question has not yet been reported and as a result is not yet contained in *Beilstein*. For details, see reference 12.

An example of the final output is shown in Figure 1; the structural formula has been added to show the structure of the compound looked for.

2.2 Sandra. The program SANDRA (short for Structure and Reference Analyzer) is a powerful software package that enables the user to draw the structure of the compound of interest, using a fast graphic input system (mechanical mouse). It then analyzes the structure and identifies, in most cases within a few pages, which part of the *Handbook* should deal with the compound of interest. The user can then go directly to the appropriate subvolume to see if the compound is known.

While BEILSTEIN KEY provides its user with the volume number only, SANDRA goes far beyond, as the example shown in Figure 2 indicates. SANDRA output provides the following

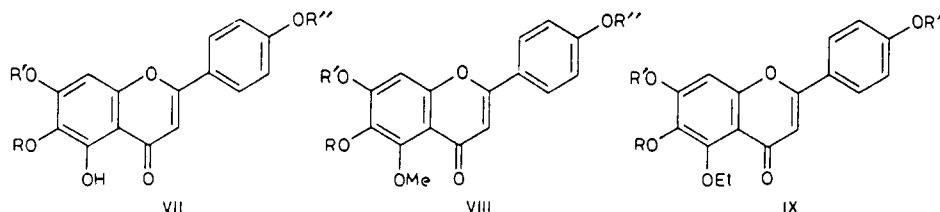


Erythromycin¹⁾, (3*R*)-6*t*-(3-Dimethylamino-β-D-xylo-3,4,6-trideoxy-hexopyranosyloxy)-4*t*-(3*O*,3'-dimethyl-α-L-ribo-2,6-dideoxy-hexopyranosyloxy)-14*t*-ethyl-7*c*,12*t*,13*c*-trihydroxy-3*r*,5*c*,7*t*,9*t*,11*c*,13*t*-hexamethyl-oxacyclotetradecane-2,10-dione, Erythromycin-A C₃₇H₆₇NO₁₃, formula XV (R = CH₃, R' = R'' = H) (E III/IV 7410).

For a review see: Koch, cit. by Florey, Anal. Profiles Drug Subst. 8 [1979] 159.

¹⁾ Index stem names derived from erythromycin are numbered as in formula XV.

Figure 3. Nomenclature I: erythromycin; numbering convention (*Beilstein* E V 18/10, 398).



5,7-Dihydroxy-2-(4-hydroxy-phenyl)-6-methoxy-chromen-4-one, Dinatin, Hispidulin $C_{16}H_{12}O_6$, formula VII ($R = CH_3$, $R' = R'' = H$).

This structure should also be assigned to the compound formulated by *Rangaswami, Rao* (Proc. Indian Acad. Sci. Sect. A 54 [1961] 51) as 5,6,7-trihydroxy-2-(4-methoxy-phenyl)-chromen-4-one (*Bhardvaj et al.*, Indian J. Chem. 4 [1966] 173, 174) and to *Salvitin* formulated by *Gupta et al.* (Indian J. Chem. 13 [1975] 215) as **5,8-Dihydroxy-2-(4-hydroxy-phenyl)-7-methoxy-chromen-4-one** $C_{16}H_{12}O_6$ (*Horie et al.*, Bull. Chem. Soc. Jpn. 56 [1983] 3773, 3778).

Identity of dinatin and hispidulin (*Bh. et al.* 175).

Isolation from *Ambrosia* species (*Herz, Sumi*, J. Org. Chem. 29 [1964] 3438; *Herz et al.*, Phytochemistry 8 [1969] 877, 879), *Balduina angustifolia* (*Lee et al.*, J. Pharm. Sci. 61 [1972] 626), *Brickellia californica* (*Mues et al.*, Phytochemistry 18 [1979] 1379), *Centaurea arguta* Nees (*Bretón et al.*, An. Quim. 64 [1968] 187, 192), *Clerodendrum indicum* (*Subramanian, Nair*, Phytochemistry 12 [1973] 1195), *Digitalis* species (*Rangaswami, Rao*, Proc. Indian Acad. Sci., Sect. A 54 [1961] 51, 55; *Imre et al.*, Phytochemistry 12 [1973] 2317; 16 [1977] 799; *Kartnig et al.*,

Figure 4. Nomenclature II: trivial names; identity of compounds (*Beilstein E V 18/5, 274*).

Dissociation Constants of Saturated Aliphatic Dicarboxylic Acids (C_2 - C_{10}) in Water					
Compound	BEILSTEIN Citation in Vol. 2; page No. in				
	E IV	E III	E II	E I	H
HOOC-COOH	1820	1541	477	220	506
HOOC-CH ₂ -COOH	1875	1598	516	244	567
HOOC-[CH ₂] ₂ -COOH	1909	1649	544	261	604
HOOC-[CH ₂] ₃ -COOH	1935	1686	564	-	632
HOOC-[CH ₂] ₄ -COOH	1957	1707	573	277	651
HOOC-[CH ₂] ₅ -COOH	-	1741	587	-	670
HOOC-[CH ₂] ₆ -COOH	-	1765	595	-	693
HOOC-[CH ₂] ₇ -COOH	-	1785	602	-	708
HOOC-[CH ₂] ₈ -COOH	2079	-	-	293	718

Figure 5. Physical data: dissociation constants.

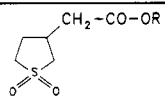
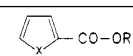
Physical Data on : 	
(taken from E V 18/6, 26)	
R	Physical Data
CH ₃	mp.
C ₂ H ₅	bp ₁₅ ; bp ₅ ; n _D ²⁰
n-C ₄ H ₉	bp ₂ ; n _D ²⁰
i-C ₄ H ₉	bp _{0.6} ; n _D ¹⁸
C ₁₂ H ₂₅	mp.
C ₁₈ H ₃₇	mp.

Figure 6. Physical data on heterocyclic compounds, I.

information (remarks in <> refer to the specific example shown in Figure 2): *Beilstein* series <E III/IV, E V>; *Beilstein* volume and subvolume numbers <20/1>; molecular formula of the compound <C 12 H 19 N O>; back reference ("coordinating reference", H-page) to the *Basic Series* ("Hauptwerk", H) <4 to 7>; *Beilstein* System number (range) <3037 to 3038>; degree of saturation <2n + 1> and number of carbon atoms <C:4> of the relevant "registry compound" <pyrrolidine>.

Physical Data on : 

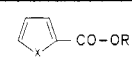
BEILSTEIN		
	x=0	x=S
No. of esters (H-E V)	172	65
No. of esters described only in E V	51 (=30%)	34 (=52%)
no. of physical data citations	49 (R=CH ₃)	37 (R=CH ₃)
...of which are numerical values	24 (=50%)	21 (=57%)

Figure 7. Physical data on heterocyclic compounds, II.

To summarize, *BEILSTEIN KEY* is a simple program designed to give users an insight into the organization of the *Handbook*, whereas the more complex program *SANDRA* is the chosen tool of the practitioner, for whom speed and efficiency are the watchwords. For more details on *SANDRA*, see references 13-17.

3. PROBLEM SOLVING WITH *BEILSTEIN*: HOW TO FIND SPECIFIC INFORMATION IN THE *BEILSTEIN HANDBOOK*

Having summarized how to find particular compounds in *Beilstein*, this section will focus on the types of information to be found in the *Handbook* and will illustrate their compilation and utilization for various purposes.

3.1 Nomenclature. The following chemical names are given: 18-chloroerythromycin and (3*R*)-9*t*-chloromethyl-6*t*-(3-dimethylamino-β-D-xylo-3,4,6-trideoxyhexopyranosyloxy)-4*t*-(3,6-dimethyl-α-L-ribo-2,6-dideoxyhexopyranosyloxy)-14*t*-ethyl-7*c*,12*t*,13*c*-trihydroxy-3*r*,5*c*,7*t*,11*c*,13*t*-pentamethyl-oxacyclotetradecane-2,10-dione. Two questions (at least) might arise: (1) Do both names describe exactly the same compound? (2) Where in the molecule(s) is, for example, position 18 containing the chloro group; i.e., what is the numbering convention of the species?

Both questions can be answered easily with *Beilstein*. First, by use of one of the above procedures, erythromycin is located in volume 18, more specifically, E III/IV 18, 7410, and E V

It has been stated that the mp of isomer II is always higher than that of isomer I and that this property may therefore be used to make stereochemical assignments. Test this statement for a number of substituents.



X=Cl, Br, NO₂, CH₃, OH, COOH, COOCH₃, COCl, CONH₂, and CN.

Melting points			
X	I (cis)	II (trans)	Beilstein Citations
Cl	18°	102°	E IV 5 51/ E III 5 40
Br	51°	113–114°	E IV 5 72/ 72
NO ₂	89–90°	174°	E IV 5 83/ 83
CH ₃	ca. –90°	–37°	E III 5 102/ 103
OH	113–114°	143°	E IV 6 5209/ 5209
COOH	172–173°	312–313°	E III 9 3818/ 3819
COOCH ₃	14°	71°	E III 9 3819/ 3819
COCl	liquid	67°	E III 9 3819/ 3819
CONH ₂	232°	346°	E III 9 3820/ 3820
CN	65°	140°	E III 9 3820/ 3820

Conclusion: Therefore stereochemical assignments using this method are valid.

Figure 8. Stereochemical assignments based on physical data.

Requested : Boiling Points(bp₇₆₀) of Mono-n-alkylamines
C_nH_{2n+1}NH₂ (n=4–12)

The following values are taken from Beilstein volume no. 4, E III and E IV :

Amines C _n H _{2n+1} NH ₂	
n	bp ₇₆₀ [°C]
4	77.8
5	105.3
6	130.0
7	155.0
8	179.0
9	–
10	220.5
11	–
12	–

Boiling Points of Mono-n-alkylamines C_nH_{2n+1}NH₂

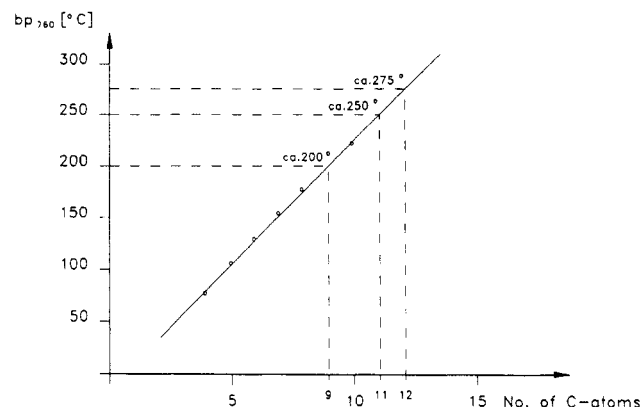


Figure 9. Interpolation and extrapolation of physical data.

Chemisches Verhalten.

but-2-en (Kp₇₆₀: 200–201°; n_D²⁰: 1,4247) erhalten worden (Clauson-Kaas *et al.*, Acta chem. scand. 6 [1952] 531, 532; Clauson-Kaas, Tyle, Acta chem. scand. 6 [1952] 962; Limborg, Clauson-Kaas, Acta chem. scand. 7 [1953] 234). Bildung von Maleinsäure beim Behandeln mit Luft in Gegenwart von Vanadium(V)-oxid bei 290–410° bzw. bei 220–270°: Milas, Walsh, Am. Soc. 57 [1935] 1389, 1392; Giller, Tarvid, Latvijas Akad. Vēstis 1952 Nr. 11, S. 89; C. A. 1955 278. Bildung von 4-Äthoxy-4-hydroxy-cis-croton-säure-lacton beim Schütteln eines Gemisches von Furan und Äthanol mit Sauerstoff in Gegenwart von Eosin unter Belichtung (Glühlampe): Schenck, A. 584 [1953] 156, 171. Bildung eines Peroxids C₈H₈O₄ (Zers. bei 90–100°) bei mehrmonatigem Aufbewahren an der Luft sowie Bildung eines Peroxids C₁₂H₁₂O₇(?) (an der Luft nicht beständig; explosiv) beim Schütteln mit Sauerstoff in Gegenwart von Calciumchlorid: Schenck, B. 77/79 [1944/46] 661, 666, 667. Bildung von Glyoxal und Ameisensäure beim Behandeln mit Ozon: Wibaut, Adv. Chemistry Ser. Nr. 21 [1959] 153, 157. Beim Behandeln mit wss. Wasserstoffperoxid und Schwefelsäure ist 1,4-Dihydroperoxy-but-2-en-1-ol (nicht charakterisiert) (Milas *et al.*, Am. Soc. 76 [1954] 2322, 2325), beim Behandeln mit wss. Wasserstoffperoxid unter Zusatz von Osmium(VIII)-oxid und Methanol sind Maleinaldehyd und meso-2,3-Dihydroxy-succinaldehyd (Clauson-Kaas, Fakstorp, Acta chem. scand. 1 [1947] 216, 218) erhalten worden. Bildung von Maleinaldehyd-bis-phenylhydrazon beim Behandeln mit Peroxybenzoesäure oder Peroxyessigsäure in wasserhaltigem Chloroform und anschließenden Behandeln mit Phenylhydrazin: Clauson-Kaas, Fakstorp, Acta chem. scand. 1 [1947] 415, 419.

Überführung in 2-Chlor-furan durch Behandlung mit Chlor bei 100°: Du Pont de

• Chemical behaviour.

Hg(³P₁) sensitized photoreactions: influence of pressure and additives (Srinivasan, Pure Appl. Chem. 16 [1968] 65). Reaction with singlet O₂ [gas phase; 300–550 K]: kinetics (Ashford, Ogryzlo, Can. J. Chem. 52 [1974] 3544, 3545). Sensitized photooxidation [Rose Bengal; MeOH/PrⁿOH/acetone; –85°] (Koch, Schenck, Chem. Ber. 99 [1966] 1984). Vapour phase oxidation to maleic anhydride [V/Mo/P catalyst; 300–350°]: kinetics (Kreile *et al.*, Kinet. Catal. (Engl. Transl.) 10 [1969] 461; orig. 565). Reaction with Fe(II)/H₂O₂; structure of intermediate radicals [ESR] (Shiga, Isomoto, J. Phys. Chem. 73 [1969] 1139). Reaction with Br₂[H₂O]: kinetics (Williamson, Coller, Aust. J. Chem. 32 [1979] 2423, 2424). Hydrolysis in aq. HClO₄: kinetics (Stamhuis *et al.*, Recl. Trav. Chim. Pays-Bas 83 [1964] 167; Unverferth, Schwetlick, J. Prakt. Chem. 312 [1970] 882). Reaction with OH-radical: ring opening (Lilie, Z. Naturforsch. Teil B 26 [1971] 197, 198). Reaction with C atoms [from 5-diazo-5H-tetrazole]: formation of pent-2-en-4-ynal (Dyer, Shevlin, J. Am. Chem. Soc. 101 [1979] 1303). Addition of cyclopropene (LaRochelle, Trost, J. Chem. Soc. D 1970 1353). Reaction with benzene: photoaddition (Berridge

Figure 10. Oxidation reactions of furan (Beilstein E III/IV, 17, 225–228; E V 17/1, 291–293).

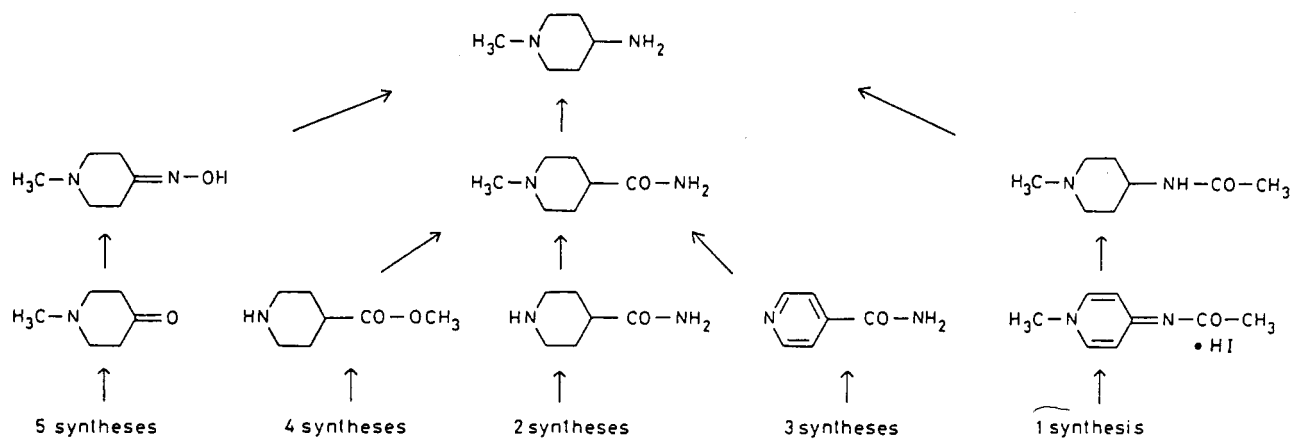


Figure 11. Strategies for the synthesis of 1-methylpiperidin-4-ylamine.

18/10, 398. The latter entry, part of which is shown in Figure 3, shows that both of the above names actually describe the same compound and, at the same time, clearly indicates the carbon atom of the methyl group that, according to nomenclature conventions, constitutes position 18.

A second example of solving nomenclature problems with *Beilstein* is given in Figure 4. This example deals with problems of (different) trivial names (dinatin, hispidulin, salvitin) given to the same compound. Whichever of these trivial names is used [or if the systematic name 5,7-dihydroxy-2-(4-hydroxyphenyl)-6-methoxychromen-4-one is taken], the entry shown in Figure 4 (E V 18/5, 274) will always be found. Lines 3-7 of the entry also show how *Beilstein* provides the user with concise details of how a compound has been reported in the literature with an erroneous assignment of constitution (in this case twice).

3.2 Evaluation of Physical Data. With over 350 different physical properties reported in *Beilstein*, the *Handbook* is an extremely effective tool for the evaluation of physical data. Some representative examples are shown in Figures 5-9.

Figure 5 gives the page numbers within volume 2 (for the *Basic Series*, H, and for *Supplementary Series* E I-IV) where the numerical values of the dissociation constants of saturated aliphatic dicarboxylic acids in water (under different experimental conditions) are given. It should be noted that, due to the *Beilstein* System of compound classification, all of those acids are described in the same *Beilstein* volume (number 2) and, furthermore, that the values of their dissociation constants are to be found within a relatively small range of page numbers (for E II, for example, between 477 and 602).

In Figure 6 the types of physical data reported for various heterocyclic esters are listed. Again it follows from the systematic arrangement of closely related compounds in *Beilstein* that all the values mentioned can be found within a small range, in this case, on a single page of the *Handbook*: E V 18/6, 26.

Figure 7 gives some information on furan-2- and thiophene-2-carboxylic acid esters. While a total of 172 furan esters have been reported in *Beilstein* for the period 1830-1980, the corresponding number for thiophene esters is 65; 30% and 52%, respectively, of these esters have been described in E V only, i.e., between 1960 and 1980. In the lower part of Figure 7, the numbers of citations of physical data for two of the well-documented compounds are listed.

A problem often encountered in examinations of the chemical literature is the report of a pair of stereoisomers whose relative configurations are unknown. Figure 8 provides an example on how, under certain circumstances, stereochemical assignments may be based on physical data. It is shown that stereochemical assignments can be made, on the basis of an obvious physical property relation, in this case by using the

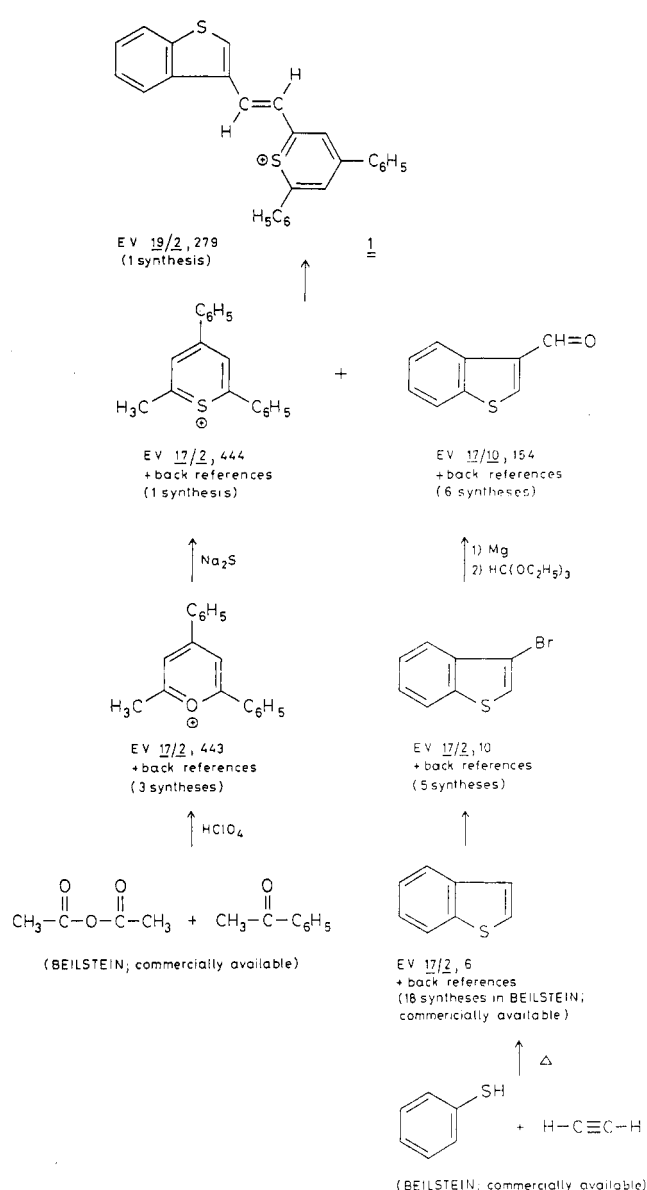
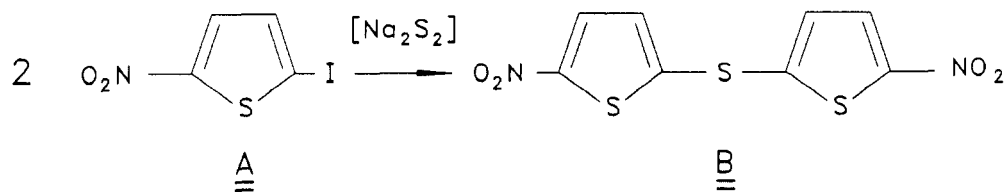


Figure 12. Synthesis of a thiopyranium compound.

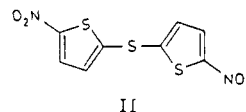
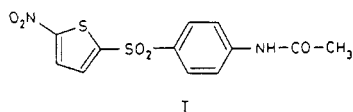
melting point values found in the *Beilstein* references given. The result: in all cases shown the trans isomers have a higher melting point than the corresponding cis isomers. Conclusion: stereochemical assignments using this method are probably valid.

Figure 9 provides an example of how to derive physical property values not directly listed in the *Handbook* (up to

The following synthesis of compound B is described in the literature:



Task : Develop a synthesis for other positional isomers of B !



Bis-[5-nitro-[2]thienyl]-sulfid $\text{C}_8\text{H}_6\text{N}_2\text{O}_4\text{S}_2$, Formel II.

B. Aus 2-Jod-5-nitro-thiophen beim Erhitzen mit Thioharnstoff in Aceton auf 120° (Dann, Möller, B. 80 [1947] 23, 34). Beim Behandeln von 2-Jod-5-nitro-thiophen mit Natriumdisulfid, Natriumpolysulfid oder Natriumthiosulfat in wss. Äthanol (Lew, Noller, Am. Soc. 72 [1950] 5715).

Gelbe Krystalle; F: $107-107,5$ [aus wss. A.] (Hurd, Anderson, Am. Soc. 76 [1954] 1267), $104-106^\circ$ [unkorr.; Block; aus Me.] (Dann, Mö., $104-105^\circ$ [aus Me.] (Lew, No.).

Positional Isomers of 2-Bromo-x-nitro- and 2-Iodo-x-nitro-thiophenes

Beilstein Citations	 Hal=Br Hal=I
E III/IV 17 257, E V 17/1 314	5-NO ₂
E III/IV 17 257, E V 17/1 315	3-NO ₂
E III/IV 17 258	4-NO ₂
E III/IV 17 259, E V 17/1 315	5-NO ₂
E III/IV 17 259	3-NO ₂

Figure 13. Synthesis strategy for bis(x-nitrothienyl) sulfides.

1960, in this case), by interpolation or extrapolation. In this case, the "missing" boiling points (at atmospheric pressure) of the C-9, C-11, and C-12 amines can be estimated to be about 200, 250, and 275°C , respectively.

3.3 Reactions ("Chemical Behavior"). Within the section "Chemical Behavior" ("Chemisches Verhalten") of the *Beilstein* entries, valuable information can be found on the various reactions the compound has undergone. Therefore, problems like "How does furan react with various oxidizing agents?" can easily be solved by consulting this particular section.

As an example, Figure 10 shows two different entries for the compound furan, focusing on part of the subsection "Chemisches Verhalten" (E III/IV) and "Chemical Behavior" (E V), respectively, which provide exhaustive information to

thoroughly answer the aforementioned question.

3.4 Evaluation of the Optimal Synthetic Strategy. The evaluation of optimized synthetic strategies for the preparation of a known or a hitherto unknown compound constitutes a topic of continuing interest and challenge to the experimental chemist. In the context of this paper, some examples are provided to demonstrate how the *Beilstein Handbook* can be integrated in these search procedures.

3.4.1 Syntheses of Known Compounds. As for the synthesis of a known compound, Figure 11 shows part of the "synthesis tree" for the compound 1-methylpiperidin-4-ylamine (first row). Looking up this compound in *Beilstein* reveals (up to 1960) three different synthetic routes, starting from the three different precursors shown in the second row. The five com-

Thiophen-2-carbaldehyd C_5H_4OS , Formel IX (H 285; E I 148; E II 313; dort als 2-Formyl-thiophen und als α -Thiophenaldehyd bezeichnet).

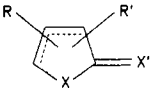
B. Beim Behandeln von Thiophen mit Phosphorylchlorid und *N*-Methyl-formanilid ohne Lösungsmittel (King, Nord, J. org. Chem. 13 [1948] 635, 638; Weston, Michaels, Am. Soc. 72 [1950] 1422; Org. Synth. Coll. Vol. IV [1963] 915; s. a. Buu-Hoi et al., Soc. 1950 2130, 2132) oder in Toluol (Emerson, Patrick, J. org. Chem. 14 [1949] 790, 792) sowie mit Dimethylformamid und Phosphorylchlorid (Campagne, Archer, Am. Soc. 75 [1953] 989) und anschließenden Hydrolysieren. Beim Behandeln von Orthoameisensäure-triäthylester mit [2]Thienylmagnesiumjodid in Toluol und anschließend mit wss. Salzsäure (Cagniant, Bl. 1949 847, 849). Beim Behandeln von [2]Thienylmethanol mit Aceton, Natriumdichromat und wss. Schwefelsäure (Em., Pa., l. c. S. 793), mit aktiviertem Mangan(IV)-oxid in Äther (Sugasawa, Mizuhami, Pharm. Bl. 3 [1955] 393) oder mit Kaliumperoxodisulfat in Wasser (Horii et al., J. pharm. Soc. Japan 76 [1956] 1101; C. A. 1957 3553). Beim Erwärmen von 2-Chlormethyl-thiophen mit Hexamethylen-tetramin in Chloroform und Behandeln des Reaktionsprodukts mit heissem Wasserdampf (Dunn et al., Am. Soc. 68 [1946] 2118; Wiberg, Org. Synth. Coll. Vol. III [1955] 811). Beim Erhitzen von 2-Chlormethyl-thiophen mit Hexamethylen-tetramin und wss. Essigsäure und anschließend mit wss. Salzsäure (Angyal et al., Soc. 1953 1742, 1747).

Thiophene-2-carbaldehyde C_5H_4OS , formula VIII (H 285; E I 148; E II 313; E III/IV 4477).

Prep. From thiophene and DMF [Ph_3PBr_2] (Bestmann et al., Justus Liebigs Ann. Chem. 718 [1968] 24, 31). From thiophene, by treatment with either $Cl_2CH-OMe$, $Cl_2CH-OBu^n$ [$SnCl_4$] (Rieche et al., Chem. Ber. 93 [1960] 88, 91, 93), or $Cl_2CH-SMe$ [$SnCl_4$] (Gross, Matthey, Chem. Ber. 97 [1964] 2606, 2608, 2612). From thiophene, by treatment with (i) chloromethylene-malononitrile [$AlCl_3$], and (ii) aq. KOH (Ertel, Friedrich, Chem. Ber. 110 [1977] 86, 91, 95). From 2-dichloromethyl-thiophene [$CaCO_3/H_2O$] (Yagupol'skii et al., J. Org. Chem. USSR (Engl. Transl.) 8 [1972] 846, 848; orig. 838). From thiophen-2-yl-methanol [Ag_2CO_3] (Fétizon et al., J. Heterocycl. Chem. 13 [1976] 525). From thiophene-2-carboxylic acid, by treatment with (i) 2-mercapto-phenol/ $POCl_3$, and (ii) H_2O (Costa et al., J. Heterocycl. Chem. 11 [1974] 943, 947, 948). From thiophene-2-carbonyl chloride and $NaAl(OBu^n)_3H$ [-60°] (Zakharin et al., J. Org. Chem. USSR (Engl. Transl.) 2 [1966] 2153, 2155; orig. 2197). From thiophene-2-carbox-nitrile, by treatment with 2-mercapto-phenol [BF_3/Et_2O], followed by hydrolysis (Degani, Fochi, Synthesis 1976 757). From thiophene-2-carbonitrile, by treatment with $SnCl_2/HCl$ (Nishimura, Imoto, Nippon Kagaku Zasshi 82 [1961] 1685). Preparation of thiophene-2- $[^{18}O]$ carbaldehyde (Andrieu et al., Bull. Soc. Chim. Fr. 1971 1314, 1317).

Figure 14. Methods of introducing a formyl group into the 2-position of thiophene (Beilstein E III/IV 17, 4477; E V 17/9, 349).

Requested : All compounds, of the following type, described in the period 1960–1980 :



R, R' = H, alkyl, halogen

X, X' = O, S, Se

1 endocyclic double bond

Monooxo compounds $C_nH_{2n-4}O_2$

Result : Described in E V 17/9, 111–288 :

152 Compounds

Figure 15. “Partial structure retrieval” (“similarity searches”) in the Beilstein Handbook.

pounds listed in row 3 are the starting materials for those in row 2, and for the five compounds in the third row, a total of 15 synthetic pathways can be found in the section “preparation” of those five compounds in Beilstein up to 1960.

This synthesis tree now allows the practical chemist to choose his own “personally optimized” way to the target molecule, depending on the availability of starting materials, preferred reaction conditions, and so on.

Figure 12 contains another example of developing a strategy for the synthesis of a given compound. While the target compound is found in E V, volume 19, most of the relevant precursors happen to be described in volume 17 (various Beilstein series). With the exception of the last step in the preparation of the thiopyranylum compound shown, there are usually various different synthetic routes to the precursors.

3.4.2 Syntheses of Unknown Compounds. If a compound of interest has not yet been reported in the literature, Beilstein

offers the possibility to develop a synthesis strategy for it, based on the known synthesis of related analogues or homologues. An example for the development of a synthesis strategy for hitherto unknown positional isomers of a given compound is outlined in Figure 13.

As shown in the upper part of this figure, a synthesis of compound B is described in the original literature. To evaluate, by means of the Beilstein Handbook, the possibility of analogously synthesizing different positional isomers of compound B, one should first refer to the synthesis of B itself, as reported in Beilstein. The second part of Figure 13 shows part of the corresponding Beilstein entry (E III/IV 17, 1232) showing that compound B has been prepared from 2-iodo-5-nitrothiophene.

The next step is to search for all known positional isomers of the starting compound, 2-iodo-5-nitrothiophene, in the Beilstein Handbook. The lower part of Figure 13 is a compilation of the Beilstein entries of all the known positional isomers of the iodo compounds, together with the corresponding bromo compounds, which might work in the reaction yielding positional isomers of compound B. All the iodo and bromo isomers may now be subjected to the reaction with one of the reagents mentioned in the Beilstein entry shown in the middle part of Figure 13.

In a second example, the task is to find different synthetic methods of introducing a formyl group into the 2-position of thiophene or substituted thiophenes. The most convenient way to solve this problem by using the Beilstein Handbook is to look for the section “preparation” of either thiophene-2-carbaldehyde itself or of some of its derivatives. Figure 14 shows the information given in this section for thiophene-2-

carbaldehyde in the *Beilstein Series* E III/IV and E V. The expressions "Beim Behandeln" ("on treatment") and "From" precede a variety of methods of introducing a formyl group into the 2-position of thiophene either directly or by varying a substituent already present in the 2-position of the molecule.

3.5 "Partial Structure Retrieval" ("Similarity Searches") in the *Beilstein Handbook*. Finally, Figure 15 demonstrates the overall result of a "partial structure retrieval" or "similarity search" carried out by browsing in the *Beilstein Handbook*. The quotation marks have been added to the two expressions because these searches, for which one example is shown here, do not merit, in the literal sense, the expression "partial structure retrieval", as this expression is, of course, reserved for substructure searches done by a computer.

Nevertheless, taking carefully into consideration the limitations imposed by the rules of the *Beilstein System*,^{1,7,8} it is possible to perform such searches as described in Figure 15 with a remarkable result: In this case, the extremely useful "browsing" method reveals a total of 152 compounds that fulfil the requirements, found within about 170 pages in subvolume 17/9 in E V.

It should be pointed out that browsing in the *Beilstein Handbook* is particularly easy and is very likely to yield a satisfactory result because of the close proximity of the entries for structurally related compounds, which is due to the ordering of the entries according to the rules of the *Beilstein System*.^{1,7,8} A more detailed explanation of how to browse efficiently in *Beilstein* will be provided in a future publication.

4. THE *HANDBOOK* AS AN ESSENTIAL AND INTEGRAL PART OF THE BEILSTEIN INFORMATION SYSTEM

The main purpose of this paper was to demonstrate how effective use of the *Beilstein Handbook of Organic Chemistry* provides a wealth of information over and above the numerical, factual, and bibliographic data relating to specific compounds. In particular, we wished to show that, apart from the obvious uses of *Beilstein*, it is quite possible to derive information or estimate numerical data on compounds that are either unknown or inadequately reported in the literature. This fact may not have been sufficiently recognized by all *Beilstein* users so far. Consequently, we have provided some examples to demonstrate how, in addition to finding *specific compounds*, one is able to also find *specific information and factual data*

in the collection. These additional search possibilities will also ensure in future that the *Beilstein Handbook* will remain an integral and essential part of the *Beilstein Information System* and that the *Handbook* may eventually even increase its reputation throughout the world.

REFERENCES AND NOTES

- (1) Beilstein Institute. "How to Use *Beilstein*"; Frankfurt/Main, 1980 (available free of charge).
- (2) Luckenbach, R. *Der Beilstein. CHEMTECH* 1979, 9, 612.
- (3) Luckenbach, R. "Kennen Sie *Beilstein*?" *Chem. Unserer Zeit* 1981, 15, 46.
- (4) Luckenbach, R.; Sunkel, J. *Das wissenschaftliche Handbuch: 100 Jahre Beilstein. Naturwissenschaften* 1981, 68, 53.
- (5) Luckenbach, R. *The Beilstein Handbook of Organic Chemistry: The First Hundred Years. J. Chem. Inf. Comput. Sci.* 1981, 21, 82.
- (6) Luckenbach, R.; Ecker, R.; Sunkel, J. The Critical Screening and Assessment of Scientific Results without Loss of Information - Possible or Not? *Angew. Chem., Int. Ed. Engl.* 1981, 20, 841; *Angew. Chem.* 1981, 93, 876.
- (7) Sunkel, J.; Hoffmann, E.; Luckenbach, R. Straightforward Procedure for Locating Chemical Compounds in the *Beilstein Handbook. J. Chem. Educ.* 1981, 58, 982.
- (8) Hoffmann, E.; Sunkel, J.; Luckenbach, R. *Beilsteins Handbuch der Organischen Chemie: Aufbau, Inhalt, Benutzungsweise und Bedeutung. CLB, Chem. Labor Betr.* 1982, 33, 205, 403, 451.
- (9) Luckenbach, R. *Beilsteins Handbuch der Organischen Chemie: 5. Ergänzungswerk. Naturwissenschaften* 1984, 71, 419.
- (10) Luckenbach, R. *Le Beilstein - aujourd'hui et demain. L'actualité chimique* 1985, No. 8, 39.
- (11) Luckenbach, R.; Jochum, C. Chemical Handbooks Online: Status Quo and Perspectives. *CODATA Bull.* 1986, No. 64, 28; and in *Computer Handling and Dissemination of Data*; Glaeser, P. S., Ed.; Elsevier Science Publishers, 1987; p 211.
- (12) Beilstein Institute. *BEILSTEIN KEY* (flow chart); Frankfurt/Main, 1987 (available free of charge).
- (13) Lawson, A. J. In *Graphics for Chemical Structures*; Warr, W., Ed.; ACS Symposium Series 341; American Chemical Society: Washington, DC, 1987; Chapter 8.
- (14) Lawson, A. J. In *Software-Entwicklung in der Chemie 2*; Gasteiger, J., Ed.; Springer Verlag: New York, 1988; Chapter 1.
- (15) Lawson, A. J. Structure Searches with SANDRA. *Beilstein Brief* 1987, No. 1.
- (16) The SANDRA program, version 2.0 (ISBN 3-540-14047-6) is available from Springer Verlag.
- (17) SANDRA runs on IBM PC/XT/AT or IBM compatible personal computers with at least 256K memory. The program has been tested on IBM PC/XT/AT, PS/2, Compaq Portable, Deskpro, Olivetti M 24 (in the U.S.: Olivetti AT&T 6300), Tandon PCA 40, Toshiba T 1100/T 3100, and HP Vectra. SANDRA version 2.0 supports IBM graphics (640 to 200 pixels), Hercules (720 to 348 pixels), EGA (640 to 350 pixels), Olivetti M 24 (640 to 400 pixels), and Toshiba (640 to 400 pixels). Standard graphics as well as all mentioned high-resolution graphics cards have been tested. A Microsoft or compatible mouse must be used as graphics input device.