Trends in Analytical Chemistry 69 (2015) 98–104

Contents lists available at ScienceDirect

Trends in Analytical Chemistry

journal homepage: www.elsevier.com/locate/trac

Chromatographic retention indices in identification of chemical compounds

V.I. Babushok *

KTC Consulting, Gaithersburg, MD, USA

ARTICLE INFO

Keywords:
Chemical compound
Data filtering
False-positive identification
Gas chromatography
Hit list
Identification
Kováts index
Linear index
Mass spectrometry
Retention index

ABSTRACT

A brief overview of application of gas chromatographic retention indices (RIs) in the identification of chemical compounds is presented. Recent advances in the development of RI databases and the accumulation of experimental retention measurements in the form of computerized libraries are essential to increasing the efficiency of the identification process with the application of retention data. Currently available libraries of retention data provide the means for the more accurate identification of isomers, which have very similar mass spectra. The use of retention data allows filtering of MS hits, substantially decreasing the number of possible candidates. The application of the calibration function in combination with RI libraries simplifies the identification process. Multimode distributions of replicate measurements of RIs are often the result of isomer misidentification. Practical identification of isomers with close retention times relies strongly on elution-order information.

Contents

1. Introduction

Gas chromatography-mass spectrometry (GC-MS) is widely used for the identification of volatile compounds by matching measured spectra with the spectra in a reference library, which is a common procedure for identification of chemical components in technological and natural mixtures [1–3]. Measurement of retention indices (RIs) of chemical compounds and comparison with available retention-data collections represents the usual approach in confirmation of compound identification. Mass spectral differences in the structures of branched alkyl substituents are typically not significant for their identification, so the combination of retention data with MS provides identification of isomers, which is difficult using mass spectra alone. A great advantage of chromatographic techniques is the ability to distinguish different diastereomers. Diastereomers are characterized by practically identical mass spectra, and the use of GC retention parameters is often the only practical method for the correct identification of such isomers. GC retention data are very attractive for applications due to the simplicity of measurements, application and interpretation. To some extent, retention data are available at no additional cost or effort as a by-product of chromatographic separation for MS measurements.

The aim of the present work is to provide a brief overview of recent progress in the development of retention-data databases and the application of RIs in the identification process, particularly with computerized RI-data libraries. We mostly focus on the traditional GC-MS approach in the identification of chemical compounds using retention data as a filter for removing false-positive MS matches. Note that there are many recent developments in chromatographic techniques (e.g., fast chromatography, multidimensional

* Tel.: +1 (301) 926 1250.
E-mail address: vbabushok@nist.gov.

http://dx.doi.org/10.1016/j.trac.2015.04.001
0165-9936/

chromatographic techniques, and comprehensive GC), where application of RIs is an essential part of the identification process. Retention data provide an additional metric in ranking candidate components in the hit lists generated using MS libraries. Recent advances in the development of computerized retention-data libraries, computer procedures with manipulation of retention data and significant increase in the available volume of retention data provide a basis for an "automatic" non-target use of retention data in the identification process.

2. Retention indices

The fundamental variable tracked in GC is the retention time. However, the retention time is a strong function of experimental conditions, and as a result other parameters have been suggested that are more independent of the experiment. The system of RIs suggested by Kováts [4] is a widely used and recognized system for recording GC data for further use in the identification process. The index system suggested by Kováts in 1958 (and its further development – the linear system of indices for temperature programming conditions [5]), allows the results measured in one laboratory to be used in other laboratories. The RI combines two fundamental GC properties, the relative retention and the specific retention volume [6].

The RI (I) is a generally accepted type of data used for the identification of chemical compounds in GC. Figuratively speaking, Kováts suggested a chemical ruler to characterize different chemical compounds on a specific time scale for identification purposes. According to Kováts, the index of an analyte is its relative time position between the nearest n-alkanes which elute immediately before and after a target analyte. Isothermal Kováts RIs are determined by the relationship [4]:

$$I_x = 100n + 100[\log(t'_x) - \log(t'_n)]/[\log(t'_{n+1}) - \log(t'_n)], \quad (1)$$

where t'_n and t'_{n+1} are adjusted retention times of the reference n-alkane hydrocarbons eluting immediately before and after compound "X", and t_x is the adjusted retention time of compound "X". RI of compound "X" represents one hundred times the number of carbon atoms in the molecule of a hypothetical hydrocarbon, which has the same retention as compound "X" [1].

Linear indices (non-isothermal indices in accord with the definition of Van den Dool and Kratz [5] from temperature-programming measurements) are defined by the following equation:

$$I_x = 100n + 100(t_x - t_n)/(t_{n+1} - t_n), \quad (2)$$

where t_n, t_{n+1} and t_x are net retention times. Using the Kováts' definition, the measure of a chemical compound is its relative time position between neighboring linear paraffin hydrocarbons. This idea has been extended using other sets of compounds for reference markers. The other widely used set of compounds was suggested by Lee et al. [7] and it uses polyaromatic hydrocarbons as marker compounds. Several RI systems were suggested for specific applications, such as, equivalent chain length, steroid number and tri-n-alkylamine index [1]. RIs determined in one system can be recalculated into the Kováts RIs or into other system, if the corresponding indices of secondary scale standards are known in the Kováts scale or another secondary scale.

3. Retention-index data libraries

A large amount of retention data for different chemical compounds is available in literature sources. Collection of RI data has a relatively long history starting from beginning of the 1960s. To make use of the retention information for identification of various compounds, it is necessary to develop a comprehensive database. Despite a great deal of data published, until recently there were only a relatively small number of retention-data collections available, and many of these were limited to specific sets of compounds (Table 1). At the early stage of retention-data systematization, ASTM efforts on the organization and compilation of GC information were very important (ASTM Committee E-19 on Chromatography) [9]. ASTM

Table 1
Collections of gas chromatographic retention data [8]

Retention data collections	Data scope	Ref.
ASTM collection	General; literature data; retention data on various stationary phases of different polarity	[9]
Sadtler library	General; own measurements; OV-1, SE-54, CW-20M, CW-20M (cross-linked), approximately 2000 compounds	[10]
Pacakova and Feltl collection	General; literature data; data for various polarity stationary phases	[1]
Bogoslovsky, Anvaer and Vigdergaus collection	General; literature data	[11]
CRC Handbook Series on Chromatography	Hydrocarbons, drugs, pesticides, amino acids and amines, terpenoids, literature data	[12]
NIST database	General, 82,868 compounds (2014 release)	[8]
The Golm Metabolome Database	Metabolites, own data, WEB based collection	[13]
FiehnLib	Metabolomics	[14]
Adams collection	Essential oil components; own measurements; DB-5, 1606 compounds.	[15]
Flavornet, Cornell University	Flavor and fragrance compounds (fragrance chemistry); own and literature data; OV-101, DB-5, OV-17, CW-20M; 738 compounds; WEB based collection	[16]
Jennings and Shibamoto collection	Flavor and fragrance compounds	[17]
Kondjoyan and Berdague collection	Flavor and fragrance compounds	[18]
Vinogradov collection	Essential oil components; literature data; nonpolar and polar stationary phases; approximately 2000 compounds, WEB based collection	[19]
HortResearch collection	Pheromones; literature data; WEB based collection	[20]
Massfinder collection	Terpenoids	[21]
ESO 2000, Database of Essential Oils	Essential oil components, 1800 compounds, commercial database	[22]
Flavor Database, Citrus Research and Educational Center, Florida State University	Flavor compounds; DB-1, DB-5, DB Wax; WEB based collection	[23]
Bianchi et al. database	Volatile compounds of food aroma	[24]
Mondello collection	Flavors, fragrances	[25]
Babushok, Linstrom, Zenkevich	Essential oils	[26]
Tarján, Takács	Essential oils	[27]
Gaskin, MacMillan, GC library of gibberellins	Gibberellins and related compounds; own data; OV-1; available on the WEB	[28]
Collection of the DFG Commission for Clinical-Toxicological Analysis	Toxicological compounds; own and literature data; dimethyl-silicone stationary phase; 4500 compounds	[29]
VERIFY Database	Chemical weapon compounds; own data, commercial database	[30]
Maurer, Pfleger, Weber	Drugs, poisons, pesticides, pollutants, metabolites	[31,32]