Trends in Analytical Chemistry 69 (2015) 24–33

Contents lists available at ScienceDirect

Trends in Analytical Chemistry

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

General principles of identification by mass spectrometry

Boris L. Milman *

Institute of Experimental Medicine, ul. Akad. Pavlova 12, Saint Petersburg 197376, Russia

ARTICLE INFO

Keywords:
Chemical identification
High-resolution mass spectrometry
Mass spectral library
Mass spectrometry
Metrology
Non-target identification
Prediction of spectra
Qualitative analysis
Spectral interpretation
Target identification

ABSTRACT

The review is devoted to chemical identification using mass spectrometry as the most powerful technique of qualitative analysis. The review begins with consideration of basic principles and means of chemical identification. Following are sections covering techniques and instruments and metrological issues. Procedures for identification outlined next are divided into target identification by methods and unknown/non-target analysis. For the latter, information support, such as mass spectral libraries and chemical databases, programs of formula generation and spectral prediction/interpretation, are reviewed. Finally, identification of samples and some general trends are briefly noted.

© 2014 Elsevier B.V. All rights reserved.

Contents

Abbreviations: B, Magnetic sector mass analyzer; CA, Chemical Abstracts; CI, Chemical ionization; CID, Collision-induced dissociation; EI, Electron ionization; ESI, Electrospray ionization; FN, False negative (result); FP, False positive (result); GC, Gas chromatography; HPLC, High-performance liquid chromatography; HRMS, High-resolution mass spectrometry; IT, Ion trap; LC, Liquid chromatography; LIT, Linear ion trap; LRMS, Low-resolution mass spectrometry; MALDI, Matrix-assisted laser desorption/ionization; MRM, Multiple reaction monitoring; MS^2, Tandem mass spectrometry; MS^3, Three-stage mass spectrometry; MS^n, Multistage/multiple mass spectrometry; NCBI, National Center for Biotechnology Information; Q, quadrupole; QqQ, Triple quadrupole; RI, Retention index; RM, Reference material; RT, Retention time; SIM, Selected/single-ion monitoring; TN, True negative (result); ToF, Time-of-flight; TP, True positive (result); UPLC, Ultra-performance liquid chromatography.

* Tel.: +7 921 766 52 96; Fax: +7 812 692 26 54.
E-mail address: bmilman@mail.rcom.ru; bormilman@yandex.ru.

http://dx.doi.org/10.1016/j.trac.2014.12.009
0165-9936/© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Mass spectrometry (MS) is a key technique for qualitative chemical analysis. It is evident in the fact that the terms of "identification" and "mass spectrometry" occur together in more than a million scientific reports returned in the search results performed by Google Scholar engine [1]. The reason for the potency of MS is that it is superior to other analytical techniques in the combination of features, such as multianalytic property, sensitivity, selectivity, possibility of characterizing compounds by molecular mass or formula, and possibility of combining with chromatography. The latter is important because gas or liquid chromatographs as inlet devices to mass spectrometers separate complex mixtures of chemical compounds for subsequent detection and recognition, with multiple selectivity of the combined techniques exhibited in qualitative determination of different organic compounds in mixtures and matrices.

There are many excellent books {see website [2]} and reviews devoted to MS as a whole. Here, we focus on aspects of qualitative chemical analysis (identification). Many essential aspects of identification were considered in the author's book published in 2011 [3]. In this review, with many references to the book, traditional advances in MS identification of chemical compounds are briefly outlined and appreciably supplemented by new developments in the field.

2. Definitions and general consideration

We begin with the definition of chemical identification. The following definition, analogous to that described in [3], can be proposed: *Chemical identification is assigning an analyte (analytical signal) to one of the set of chemical compounds or to a group/class of compounds.* Compounds may be "known known", "known unknown" or "unknown unknown" [4]. In the first case, there is target analysis, which is a determination of analytes specified before performing analytical procedures. "Known unknown" analytes are identified in non-target or unknown analysis, which is a determination of analytes unknown to a chemist before analyzing the corresponding samples. These may not necessarily be entirely new compounds, which are "unknown unknowns". A process for their identification is often named "structure elucidation".

It should be noted that *chemical compounds* are not the same as *chemical substances*. An (individual) compound has a definite molecular structure. A substance may be composed of one or more different individual compounds. In statistics of known chemical entities recorded by the Chemical Abstracts (CA) database [5], first, low-molecular substances, which are individual compounds and their compositions, and, second, sequences (e.g., DNA, RNA, genes, peptides, and proteins) are counted separately (Fig. 1). Now, the number of molecular entities in each of the two classes is measured in tens of millions (Fig. 1).

The space of biocompounds can be expressed in some other numbers. For example, the human proteome numbers about 10^6 proteins [6]. In 2014, the popular curated databank UniProtKB/Swiss-Prot contained 546,000 entries of known or predicted amino-acid sequences [7]. Far more data, although not totally verified, are contained in NCBI resources [8].

Chemical analysis has also been used permanently to identify or to classify an analyzed sample as the definite kind/type/grade/brand of products, materials, compositions, formulations, and so on. This type of chemical analysis was called qualitative analysis II (identification II) [3]. The definition of identification can therefore be extended: *Chemical identification is also assigning an analyzed sample to one of the classification groups for specimens, materials, products, foodstuffs, pollutions, living organisms, and so on, using the techniques and methods of chemical analysis* [3].

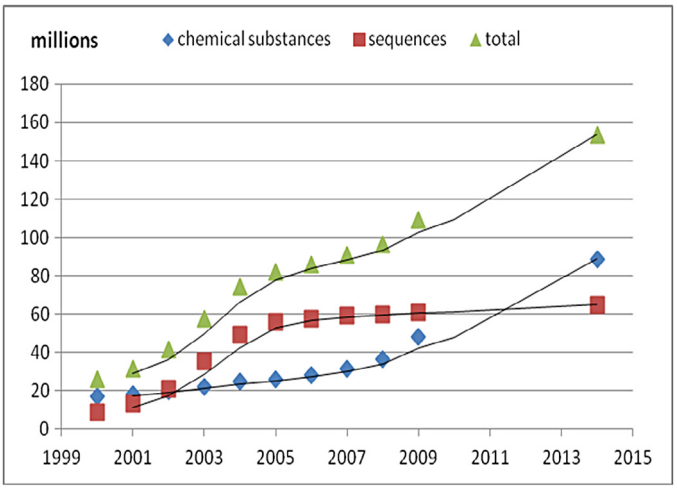

Fig. 1. Estimation of the number of known substances expressed by different registry numbers in the CAS { [5], 3 June 2014, and before [3]}.

A result of chemical identification is derived from matching physical, chemical, and/or biological properties. An analyte is considered to be unambiguously identified as the certain compound/substance X when, first, various properties of an analyte and X are identical and, second, those of an analyte and all other compounds/substances are different. If the properties of an analyte and compounds X, Y, Z and so on cannot be differentiated by some method or technique, there is a case of ambiguous or group identification. If a sample contains a marker compound X′ specific for the kind of the sample, or a group of compounds $X'_1, X'_2, X'_3, \ldots, X'_n$ in a specific ratio, or simply a specific pattern of analytical signals, a sample can be assigned to the particular specimen, material, product, food, pollution, living organism, and so on.

3. Approaches to identification

Five general procedures, especially performed for identification, are given and subject to comment in Table 1. The related idea of chemical identification is shown in Fig. 2. As a whole, the reliability of identification increases in series: $5 \le 4 < 3 < 2 < 1$ (for the numbers, see Table 1). In target determinations by validated confirmatory methods (i.e., in definite identification), only the most reliable means as, first, co-analysis with authentic analytical standards (approach 1) and, second, comparison with reference data obtained in conditions very similar to analytical runs/scans (a version of approach 2) are recommended for use.

In non-target screening, all possible means are required, with the enforced use of means 3, 4 and 5 (Table 1) in cases of very rare and novel analytes, where reference data or materials are not available. Mixed approaches are often implemented (e.g., combination of mass-spectra prediction and matching theoretical ion mass in proteomics) [3]. Identification of not very high reliability typical for screening procedures has been called tentative, preliminary, or putative.

4. Analyte classes, techniques and instruments

For MS, compound properties, such as volatility, polarity of molecule, and level of molecular mass, are of the greatest value. Based on these properties, all compounds are divided into the three following groups.

Download English Version:

https://daneshyari.com/en/article/1248278

Download Persian Version:

https://daneshyari.com/article/1248278

Daneshyari.com