=== REVIEWS ===

Mass Spectrometry as a Crucial Analytical Basis for Omics Sciences

V. G. Zaikin[a] and R. S. Borisov[a, b, c], *

[a]Topchiev Institute of Petrochemical Synthesis, Russian Academy of Sciences, Moscow, 119991 Russia
[b] RUDN University, Moscow, 117198 Russia
[c] Core Facility Center "Arktika," Northern (Arctic) Federal University, Arkhangelsk, 163002 Russia
***e-mail: borisov@ips.ac.ru**

Abstract—This review is devoted to the consideration of mass spectrometric platforms as applied to omics sciences. The most significant attention is paid to omics related to life sciences (genomics, proteomics, metabolomics, lipidomics, glycomics, plantomics, etc.). Mass spectrometric approaches to solving the problems of petroleomics, polymeromics, foodomics, humeomics, and exosomics, related to inorganic sciences, are also discussed. The review comparatively presents the advantages of various principles of separation and mass spectral techniques, complementary derivatization, used to obtain large arrays of various structural and quantitative information in the mentioned omics sciences.

INTRODUCTION

In the last decade of the XX century, virtually new sciences with the names containing the suffix "omics"

Abbreviations: ***Ionization techniques*:** electron ionization (EI); electrospray ionization (ESI); matrix-assisted laser desorption/ionization (MALDI) and its matrixless versions like SALDI; mass spectrometry imaging (MSI); matrix-assisted laser desorption/ionization with mass spectrometric imaging (MALDI-MSI); atmospheric pressure photoionization (APPI); atmospheric pressure photochemical ionization (APPCI); chemical ionization (CI); atmospheric pressure chemical ionization (APCI). ***Methods for separation, decomposition, and analysis of ions:*** Fourier-transform ion cyclotron resonance mass spectrometry (FTICR-MS); orbital ion trap Orbitrap (LTQ, Orbitrap); combination of quadrupole and time-of-flight mass spectrometer mass analyzers (QTOF); ion mobility mass spectrometry (MS-IM); mass spectrometry based on a triple quadrupole analyzer (triple quad); collision induced dissociation (CID = CAD); electron-transfer higher-energy collision dissociation (EThcD); surface induced dissociation (SID); electron-transfer dissociation (ETD); electron capture dissociation (ECD); post-source decay in MALDI (PSD); infrared multiphoton dissociation (IRMPD); ultraviolet photodissociation (UVPD); electron photodetachment dissociation (EPD); selected reaction monitoring (SRM); parallel (multiple) reaction monitoring (MRM). ***Online and offline methods for preseparation of complex mixtures:*** high performance liquid chromatography (HPLC); two-dimensional liquid chromatography (2D-LC); two-dimensional combined chromatography HILIC/HPLC; two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): liquid chromatography (LC), ultra-high pressure liquid chromatography (UHPLC) or ultra-performance liquid chromatography (UPLC); liquid chromatography under critical conditions (LCCC); liquid adsorption chromatography (LAC); liquid exclusive adsorption chromatography (LEAC); capillary electrophoresis (CE); capillary zone electrophoresis (CZE); capillary electrophoresis/mass spectrometry with ESI (CE/ESI-MS); hydrophilic interaction liquid chromatography (HILIC); size-exclusion chromatography (SEC).

emerged [1]. In the literature, such scientific branches are shortly named "omics," meaning they are all engaged in versatile studies of complex mixtures and operate on large amounts of data. Hereafter, we also call these sciences "omics" and their research apparatus "omics technologies." Most of the omics formed to date belong to the life sciences, and proteomics and genomics arose first. Dozens of scientific disciplines like glycomics, metabolomics, lipidomics, transcriptomics, interactomics, degradomics, neuromics, and others have a close relationship with proteomics and genomics. Some of these sciences gave birth to particular scientific fields, engaged in studying narrower arrays of bioorganic compounds, for example, steroidomics, glycoproteomics, phosphoproteomics, and others. In fact, such sciences as pharmacogenomics and pharmacoproteomics, which use omics technologies to develop and investigate the physiological properties of pharmaceuticals, should be considered independent. The use of the suffix "omics" in the names of newly emerging sciences that were not directly related to the life sciences have been criticized for quite a long time; it is curious that the proposal to name peptidomics a science related to the synthesis and research of oligopeptides was also considered ineligible. Nevertheless, the names of petroleomics, polymeric, humeomics, etc., dealing with samples of inanimate nature and using the general research principles of omics technologies, have already confidently taken root in the scientific lexicon.

Working with large amounts of data is not complete without statistical, mathematical, information systems, and databases. Mass spectrometry (MS) is the
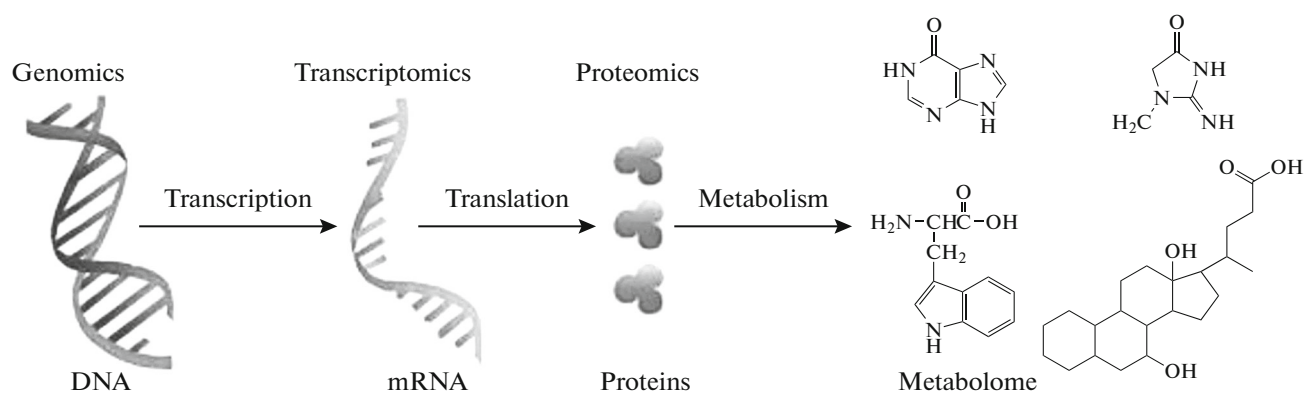
Fig. 1. Relationship between omics sciences included in the methods of systems biology. Reproduced with modifications from [8] with permission. Copyright © 2014 Elsevier Ireland Ltd.

most powerful research and analytical tool in many omics sciences. Highly sensitive, precise, rapid, and selective technique with high dynamic range has proven to be particularly effective and informative when combined with separation techniques, tandem experiments, and informational approaches, providing comprehensive research. We can assume that mass spectrometry laid a foundation for the bunch of omics sciences. Indeed, many of them took shape as sciences after the invention in the 1980s of fundamentally new, highly effective methods of ionization and analysis of ions, among which were electrospray ionization (ESI) [2, 3] and matrix-assisted laser desorption/ionization (MALDI) [4, 5]. The role of these methods was emphasized in several reviews that have been periodically published over the past 20 years and which are cited below in the main part of this review, but here are some more [6, 7]. Within the framework of this review, we did not touch specifically upon bioinformatics systems, various software, and databases, many of which were developed for the interpretation of data and results of omics studies. At the same time, we were aware that without their use, vast arrays of mass spectral data, in particular, obtained using tandem mass spectrometry, cannot be processed manually.

1. MASS SPECTROMETRY IN OMICS LIFE SCIENCES

Biological systems are incredibly complex samples for study because of the enormous number of processes simultaneously occurring in them. Even one living cell is a constantly operating biochemical laboratory, and the interaction of cells in multicellular systems, in its versatility, is an almost inexhaustible object of research. In the early 2000s, the concept of systems biology, a branch of science that studies the interaction of the components of biologic systems and the characteristics of their functions and behavior resulting from this interaction, was introduced into wide scientific use. Its leading methods include interrelated

omics sciences: genomics, transcriptomics, proteomics, and metabolomics (Fig. 1) [8].

1.1. Genomics and Transcriptomics

Winkler [9] proposed the term "genome" in 1920 to designate the totality of genes in all chromosomes in the nucleus of a cell of an organism. Currently, the genome is considered the complete set of DNA in organisms, including all of its genes. The term "genomics" was coined by geneticist Roderick in 1986 during a microsymposium at an international conference in Bethesda (United States) [1]. The essence of this scientific interdisciplinary biological science is the study and comparison of genomes throughout a body, emphasizing their analysis, sequencing, mapping, the study of functional properties, evolutions, and editings.

Transcriptomics is a branch of genomics. Its task is to study a transcriptome, a set of RNAs formed in one or a series of cells at a specific point in time. This set consists of ribosomal (rRNA), transport (tRNA), and messenger (mRNA) types of RNA. Of these, mRNA serves as a transmitter of genetic information encoded in DNA. The sequence of nucleotides in mRNA molecules contains all the information about the amino acid sequence in a protein chain. After translation, these RNAs are rapidly degraded. During translation, tRNAs take part as an intermediate between nucleic acids and proteins: they attach activated amino acid residues and transfer them to the site of synthesis of protein polypeptide chains. rRNAs do not participate in the process of information transfer and constitute the bulk of the ribosomes. Other types of RNA perform their specific functions in biochemical processes. All types of RNA are synthesized on a DNA template, and the sequence of ribonucleotides in them is complementary to the sequence of deoxyribonucleotides in DNA [10].

The elementary monomeric units of DNA and RNA macromolecules are nucleotides, consisting of

nitrogenous bases, monosaccharides, and phosphoric acid residues. Molecules of purine (adenine, guanine) and pyrimidine (thymine, cytosine) structures act as nitrogenous bases for DNA. The same bases are included in the nucleotides of RNA molecules, except that thymine in them is replaced by uracil. In DNA and RNA, these nitrogenous bases bind to monosaccharides deoxyribose and ribose, respectively, forming nucleosides. In nucleotides, the latter is phosphorylated at the primary OH group of a monosaccharide, and this phosphate group phosphorylates the OH group of a monosaccharide of another nucleotide, because of which macromolecules of nucleic acids are formed.

Within the framework of this review, we do not consider the varieties of genomics and transcriptomics; we only note the possible application of mass spectrometry in these fields. ESI-MS and MALD-IMS have the most significant opportunities for the study of DNA and RNA [11]. These methods can be used to determine the molecular weights of such macromolecules, their various complexes, and oligonucleotides, which is essential in studying their transformations.

Tandem mass spectrometry is suitable for determining the primary structure (nucleotide sequence) of native molecules [12]. However, next-generation sequencing (NGS), which generates a massive number of nucleotide sequences, is more efficient and faster.

Of these two methods, MALDI-MS is most often used in the analysis of nucleic acids. However, RNAs are more stable than DNA under MALDI-MS conditions. Therefore, the latter are preliminarily converted into RNA by in vitro transcription. Using MALDI-MS/MS, post-translational modifications in rRNA can be decoded, single nucleotide polymorphism (SNP Genotyping) in the genome and quantitative analysis of gene expression can be done [13]. Another option for determining the type of a specific nucleotide is the use of peptidonucleic acids (PNAs) with missing links: these compounds form pairs with nucleotide chains, in which nucleotides interact with PNA peptides. The nucleotide to be identified turns out to be opposite the missing link in the PNA and remains reactive. The products of the reaction with nucleoside-specific reagents, after reduction, are detected by MALDI (Fig. 2) [14].

The application of MS to genomic research, concerning the in vivo decoding of the types and locations of chemical structural modifications of nucleic acids, makes excellent progress in this field [15]. As the intensity of these transformations is relatively low, the use of a highly sensitive MS in such works is of fundamental significance. Therefore, the use of preliminary chemical modification is of particular interest, which was discussed in sufficient detail in [16].

1.2. Proteomics

Proteomics is one of the most intensively developing omics sciences (Fig. 3); its purpose is a comprehensive, detailed study of the composition, modifications, and functional properties of protein molecules in the cells of living organisms. The essential tasks of proteomics are identifying and quantifying the expression of both individual types of proteins in biological samples and all proteins of the body, depending on their state, participation in a particular biochemical process, or modification under a targeted action. The complete set, diversity, and profile of proteins of a cell, tissue, and organism at each moment are called "proteome." This term was first introduced by M. Wilkins, an Australian scientist, who also coined the term "proteomics" [17, 18]. The proteome is the first link in the information flow resulting from the transcription of the genome through the transcript. It contains millions of protein forms. Among them are dynamically changing post-translationally modified proteins (phosphorylated, sulfonated, glycosylated, lipidated, methylated, etc.), which cause rapid physiological processes and have short lifetimes. Many proteins interact with other cellular components, while others are present in the cell in trace quantities. All these factors pose challenges in identifying, profiling, and quantifying them.

Mass spectrometry has been the primary analytical tool virtually since the inception of proteomics as a science [19, 20]. ESI and MALDI mass spectrometry with various tandem options and high-resolution techniques is most widespread in such studies. This toolkit contributes to research in fundamental proteomics, which aims at obtaining qualitative and quantitative information about proteins, their composition, structure, complete or partial amino acid sequence, isoforms, post-transcriptional and post-translational modifications (PTMs), conformations, absolute or comparative quantitative data; it is called structural proteomics. Functional proteomics, in its turn, studies proteins in their interactions with other proteins and nucleic acids. Such studies are usually based on the results of transcriptomics.

There are two types of proteomics "off-target" and "target." The former includes identifying and analyzing the expression of new, previously unexplored proteins, and the latter deals with a comprehensive analysis of a set of already known proteins. In general, carrying out these studies is a challenging task because, according to some assumptions, the human proteome consists of several million different proteins, the amount of which is even greater if we take into account their various proteoforms [21]. We emphasized that various versions of mass spectrometry combined with a separation technique demonstrate exceptionally great possibilities in solving many emerging proteomics problems.
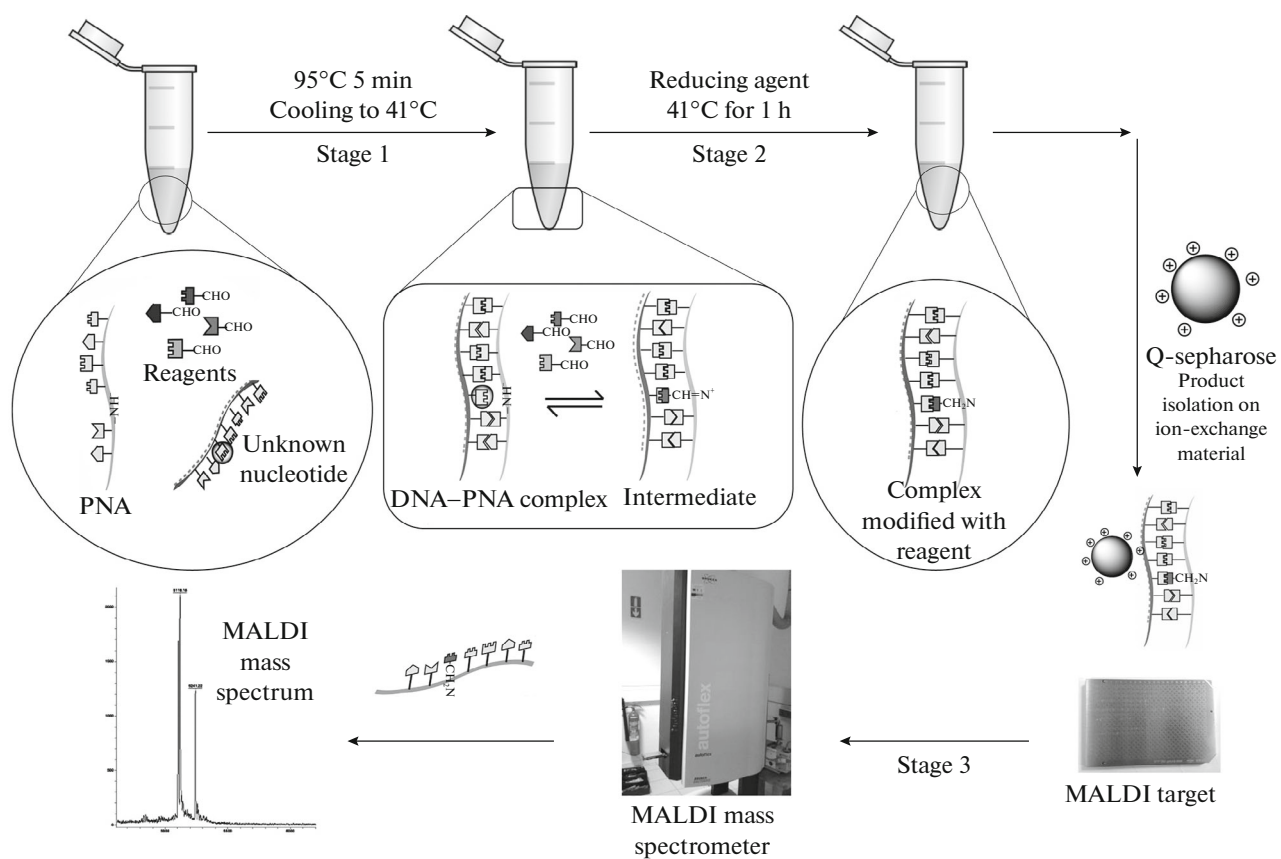
Fig. 2. Schematic diagram of studies to identify a specific nucleoside in the chain using PNA and nucleoside-specific reagents. Reprinted with modifications from [14] with permission. Copyright© 2017 Elsevier B.V.
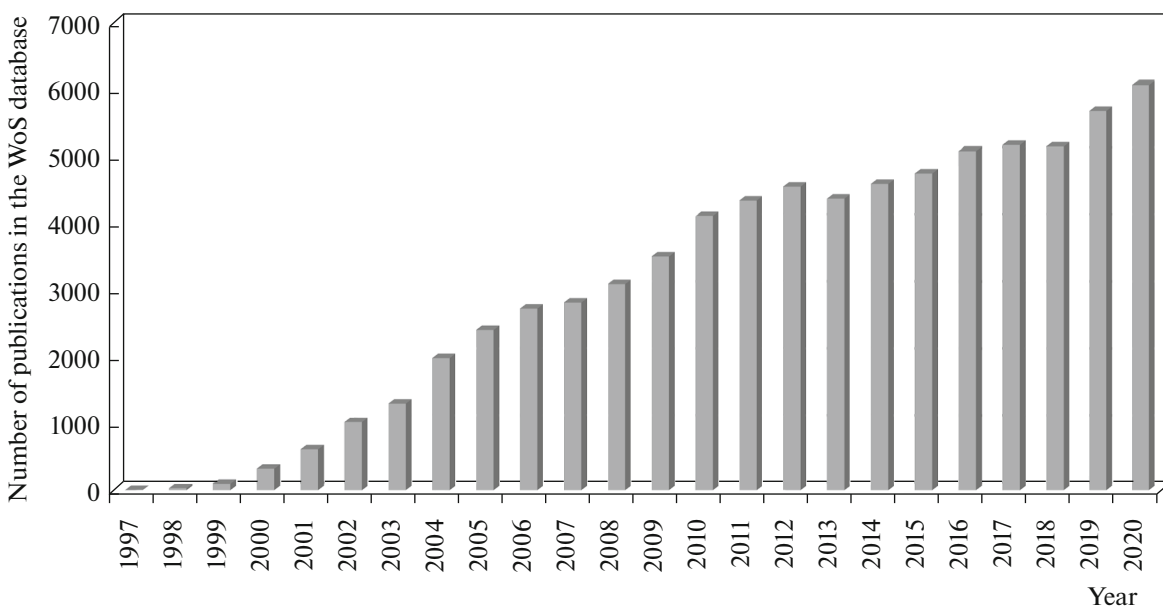

Fig. 3. Number of publications on proteomics in the Web of Science database (WoS, www.webofknowledge.com, Web of Science, a registered trademark of Clarivative) over the years.

The primary task of the mass spectrometric analysis of proteins is their identification and, in the simplest case, the determination of their primary structure. For this, mass spectra of peptides are automatically compared with experimental or predicted (in silico) spectra in databases, or direct de novo sequencing is performed using tandem mass spectrometry, yielding MS/MS spectra [22]. In the latter case, tandem devices like "triple quad," Orbitrap, and others are often used.

Proteomic technologies include the stages of (1) isolation of a set of proteins, (2) direct analysis of them or of peptides obtained by enzymatic hydrolysis using MS and MS/MS methods in combination with separation techniques. Here, versions of polyacrylamide gel electrophoresis 2D-PAGE or SDS PAGE (carried out in the presence of sodium dodecyl sulfate), capillary electrophoresis, reversed-phase HPLC, or multivariate HPLC are used. Then, (3) proteins are identified using databases; for example, the Human RefSeq 75 protein database contains 695 218 MS/MS spectra, among which 225 758 peptide spectra are assigned to 21 164 peptides corresponding to 4493 proteins [23]. Bioinformation systems, for example, SEQUEST, Mascot, are also used in this stage.

Modern mass spectrometers based on ESI, MALDI, or SALDI, ensuring ultrahigh resolution (above 400 000), high accuracy in determining the ion mass (below 1 ppm), and high sensitivity (limit of detection up to the attomole level), in combination with tandem versions, are especially effective for sequencing peptides and proteins. Within the framework of this review, it does not seem possible to consider the incredible abundance of studies in proteomics, carried out over the past 20—25 years with the use of MS. We give just some examples to illustrate a general idea of the irreplaceable possibilities of MS in this field.

Two main versions of MS gained the broadest application in structural proteomics:

(1) The earliest strategy, called "bottom-up," involves the initial separation of complex mixtures of proteins by 2D-PAGE [24], based on isoelectronic points. Using MALDI-MS, separated proteins are detected, profiled, and primarily characterized by comparing the masses of their protonated molecules with those in databases (peptide fingerprint mapping, PFM). The target proteins from the separated proteins are hydrolyzed by proteases (trypsinolysis) directly in the gel, and sets of the resulting peptides are sequenced using MALDI-MS/MS or LC/ESI-MS/MS. Then, the complete or partial amino acid sequence of specific proteins is found using the structures of peptides and the corresponding bioinformatics systems (for example, SEQUEST).

HPLC—MS combined instruments are especially effective in such studies. Both normal-phase and reversed-phase HPLC systems are used, which are compatible with ESI-MS. Various ion traps and scanning mass spectrometers with a relatively narrow range of measured masses but with high resolution and tandem layout can be used to detect and identify peptides.

(2) The so-called "top-down" proteomics is becoming more and more widespread [25, 26]. The development of this approach is associated with the main disadvantage of the bottom-up method, namely, the complexity of "assembling" the initial peptide from its fragments. In the top-down version, a mixture of intact proteins, without preliminary separation or proteolysis, is fed into an ESI mass spectrometer operating in a tandem mode. Each type of protonated protein molecules generated by ESI undergoes induced dissociation (for example, ETD or ECD), and the recorded mass spectra yield virtually complete information about the amino acid sequence and, essentially, enables the detection of various highly labile PTMs, if any, and the location of their positions in each protein. A possibility of obtaining exhaustive structural information is perhaps the essential advantage of this methodological approach. The FTICR-MS and Orbitrap systems are especially effective in such studies, because they ensure high resolution and more reliable structural assignments. A possibility of de novo peptide sequencing by tandem data obtained using the "top-down" strategy [27] should be paid special attention. The disadvantage of the top-down methodology is the limitations on the masses of proteins that can be analyzed by existing mass spectrometric systems, especially those of high resolution. The most effective of them make it possible to study proteins with molecular weights of an order of 30—80 kDa. However, some approaches of this method enable the detection of large polypeptides, for example, thyroglobulin with a molecular weight of more than 660 kDa. The mass spectrum of such molecules in an intact form cannot be recorded; therefore, they are subjected to fragmentation because of a change in the declustering potential [28].

"Shotgun" proteomics probably should be considered the third version of the protein research strategy. In contrast to the bottom-up methodology, which yields a large amount of structural and quantitative information but is labor-intensive and time-consuming, shotgun proteomics excludes the preliminary separation of a mixture of proteins by gel electrophoresis [29]. The protease hydrolyzes the entire test mixture of proteins, and the resulting mixture of peptides is separated by convenient LC options, followed by the online mass spectrometric detection of each component. ESI is the most often used to ionize peptides, while to determine their structure (in particular, sequencing), various opt+ions of tandem mass spectrometry are applied, based, for example, on the use of a triple quadrupole (triple quad), a quadrupole with a time-of-flight analyzer, or a dual time-of-flight analyzer (TOF/TOF). Recently, Orbitrap (LTQ) has been used intensively for these purposes, distinguished by

high mass resolution and yields reliable information about the structure. The use of MALDI tandem mass spectrometry with LC in the offline mode [30] excludes the gel separation stage, but generally has limitations in obtaining information on the structure of intact proteins. However, it seems convenient for a comparative analysis of proteomes, especially in analyzing protein biomarkers in medicine.

An ultrafast method for the determination of proteins for bottom-up proteomics, based on a combination of gradient LC with ultrashort columns and MS, includes the recording of high-resolution single mass spectra (MS1) (Orbitrap) of proteolytic peptides during their elution and the determination of accurate masses and the prediction of the retention times (DirectMS1) [31]. This method demonstrates high sensitivity and helps to identify a large number of proteins. It offers high coverage of their amino acid sequences, eliminates the need to obtain MS/MS spectra, and shortens the time for complete proteomic analysis (identification of thousands of proteins) to several minutes. The LC retention times of the detected peptides are essential parameters for this approach, the dependence of which on the amino acid sequence was demonstrated earlier [32].

In recent years, a combined mass spectrometric technique, including the online separation of proteins and peptides by capillary electrophoresis (CE), based on their electrophoretic mobility, has become more and more widely used in proteomic studies [33]. Capillary zone electrophoresis is most often used in this case. It is most convenient for mass spectrometric identification to use ESI with suitable options to obtain MS/MS spectra. In general, CE−ESI-MS is especially effective for top-down proteomic analyzes; it is also called "native" MS.

One of recent directions in proteomic research is determining the nature and locations in protein molecules of post-translational modifications (PTMs), which arise in proteins after the translation of information by mRNA. About 300 PTM versions are known [34], among which glycosylation and phosphorylation are the most important. Thus, the study of these PTMs actually led to the creation of glycoproteomics and phosphoproteomics. Both tandem bottom-up and top-down mass spectrometric strategies are quite effective for detecting, positioning, and quantitative analysis of PTMs [36, 37]. In some cases, glycosyl and phosphoryl residues can be cleaved off under CID conditions; therefore, preliminary chemical modification is used to analyze glycosyl and phosphopeptides [38]. Activation methods also give valuable information, yielding complementary data sets, for example, CID, ETD, and EThcD (Fig. 4) [39, 40].

Bottom-up proteomic analysis directly in tissues is another developing field of MS. Dapic et al. [41] analyzed the approaches used in the sample preparation (selection of enzymes for hydrolysis, extraction, purification), assessment of the required amount of starting material with the prospect of analysis at the level of several cells, etc. Derivatization approaches that include LC−ESI-MS and MALDI-MSI analysis, which are increasingly being introduced into clinical practice, were also considered.

Many mass spectrometric approaches have been proposed for the absolute and comparative quantitative determination of proteins in proteomes. The main ones are based on the in vivo or in vitro introduction of isobaric, isotopically labeled residues into the molecules of proteins or proteolytic peptides [38]. Commercially available reagent kits of iTRAC and TMT with various amounts and positions of light and heavy isotopes of O, C, and N became most common for in vitro derivatization. Both reagents introduce the appropriate group to the N-termini of proteins and peptides in parallel samples. Using different options for recording MS/MS spectra, one can compare quantitatively from four to eight samples in the case of iTRAC in one mass spectrometric experiment [42], and their number can be increased to 54 in the case of TMT [43]. However, there are approaches to mass spectrometric quantitative determinations in proteomics without isotope labeling [44].

The entire proteomics toolkit was in demand in connection with the SARS-CoV-19 virus pandemic. Initially, all efforts were directed at studying the mechanisms of infection [45]. However, the rapid increase in the incidence has shifted the focus of work to identifying peptides that could be used in the diagnosis of diseases [46]. In comparison with tests based on polymerase chain reactions, mass spectrometry ensures greater accuracy in identifying sick patients and more rapid analyzes themselves [47].

At the end of this section, we should pay attention to the expansion of protein research by ambient mass spectrometry. The most significant advantages of these methods over conventional ESI-MS and MALDI-MS are the shorter time required to prepare biological or clinical protein samples for the analysis, which can also be studied in the solid state [48].

Interactomics is a branch of functional proteomics. It studies the interaction of proteins with other proteins and other components (mainly nucleic acids) of the cell and the consequences of such interactions. Protein interactomics is distinguished, which studies the interactome, that is, the entire set of proteins interacting with a specific protein, and the organization of complex protein interactions that determine the cell's life. Some physicochemical principles have been proposed, including computer programs, to study and predict such interactions [49, 50]. In this brief review, we only consider those techniques in which the capabilities of MS were manifested to the greatest extent. Currently, there are two mass spectro-
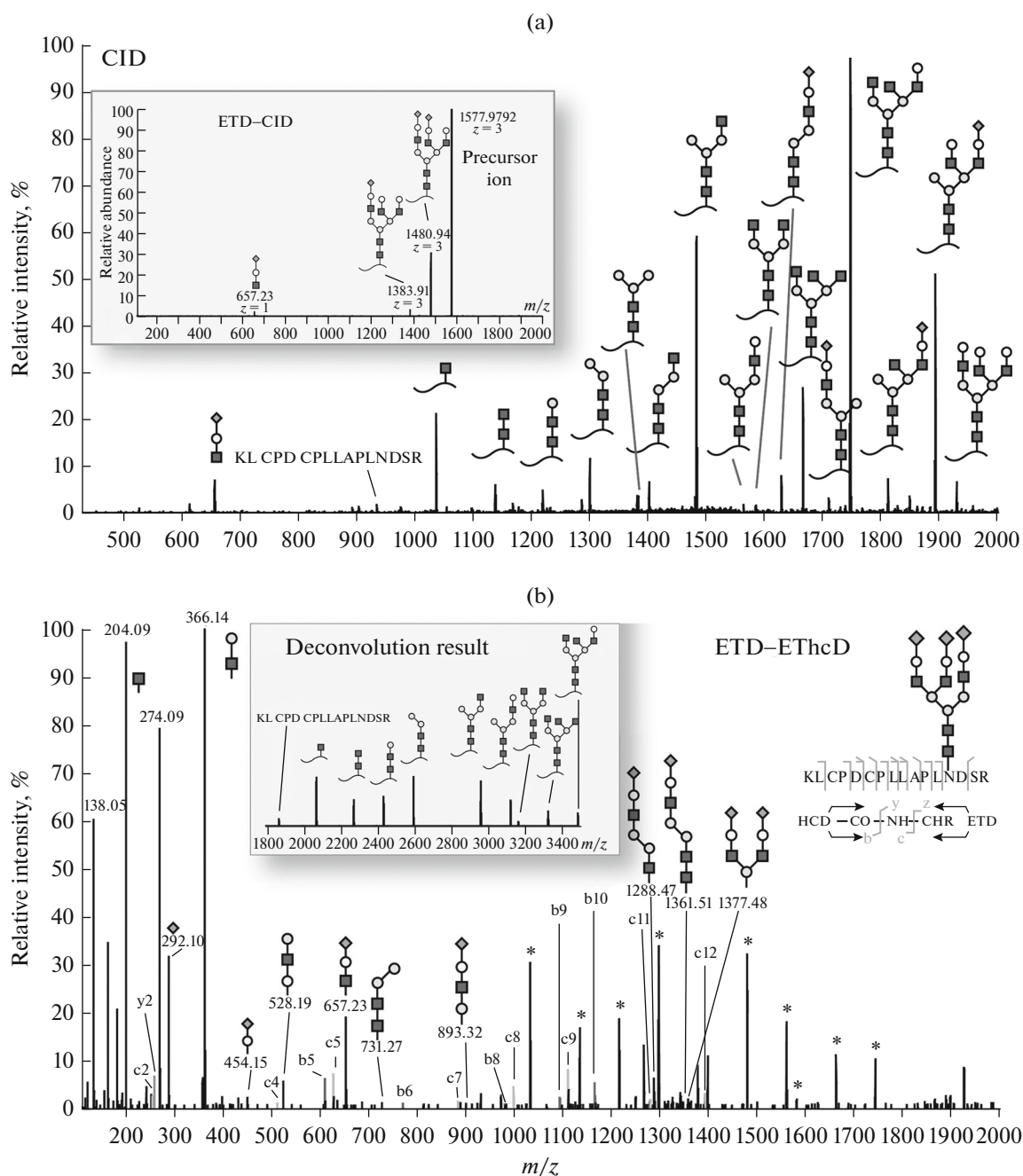
Fig. 4. Mass spectra of the products of the activated decay of the triply charged ion of the N-glycopeptide KLCPD-CPLLAPLNDSR: (a) CID (ETD and CID in the inset) and (b) ETD and EThcD. Reprinted with modifications from [39] with permission (Copyright © 2018 Published by Elsevier B.V.) and from [40] with permission.

metric approaches to the study of protein interactions [51, 52].

One option is the mass spectrometric determination of noncovalent protein interactions, based on the use of the so-called native ESI-MS (to a lesser extent, MALDI-MS). In this case, the mildest experimental conditions were selected, under which the noncovalent complexes are stable. This approach makes it pos-

sible to determine the masses and composition of the complexes and the protein binding constants with high accuracy. This method was used, for example, to study the interaction of the virus with the receptors of carrier cells and confirmed a possibility of using heparin to prevent this binding (Fig. 5) [53].

Another approach to studying the interactions of proteins and their organization (conformation, archi-
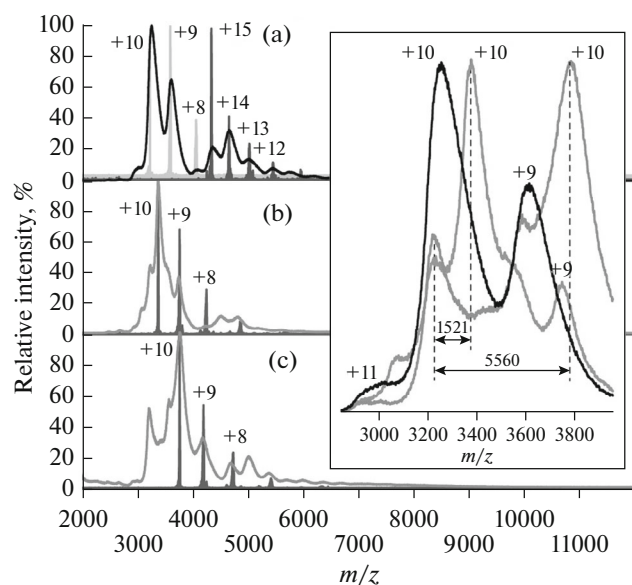
Fig. 5. ESI mass spectra of recombinant forms of the receptor-binding domain of S-protein, obtained (a) in the absence of heparinoids and the presence of (b) fondaparinux and (c) heparin oligomer. Inset: an enlarged region of superimposed spectra corresponding to the +10 charge state. Reprinted with modifications from [53] with permission (Chemistry in Coronavirus Research: A Free to Read Collection from the American Chemical Society). Copyright © 2020 American Chemical Society.

tecture of complexes, networks) in specific systems is based on so-called cross-linking mass spectrometry (XLMS). This principle has been developed for a relatively long time and reviewed in several articles (see, for example, [38, 54−56]). It includes the use of cross-linking spacers of a given length, having functional groups at both ends that readily react with specific groups in the protein side chains. Such cross-linking can only occur if the distance between the reacting groups of the two proteins is not greater than the length of the resulting bridge. The proteins cross-linked in this way undergo proteolysis, and the mass spectrometric sequencing of the resulting peptides (MALDI-MS, ESI-MS) reveals the attachment sites of such cross-linkers in each of the interacting proteins. This approach was used in a series of works devoted to studying protein interactions in various mitochondria compartments (Fig. 6) [57].

Peptidomics should be considered as a branch of proteomics, because its task is an exhaustive qualitative and quantitative analysis and the study of all possible biological properties of low-weight endogenous oligopolypeptides (conventionally up to 10 kDa) [58, 59]. Peptides in the body either have their functional properties (for example, as hormonal or signaling molecules) or are products of the usual enzymatic hydrolysis of proteins from food. The peptide composition and their profiles strongly depend on the way
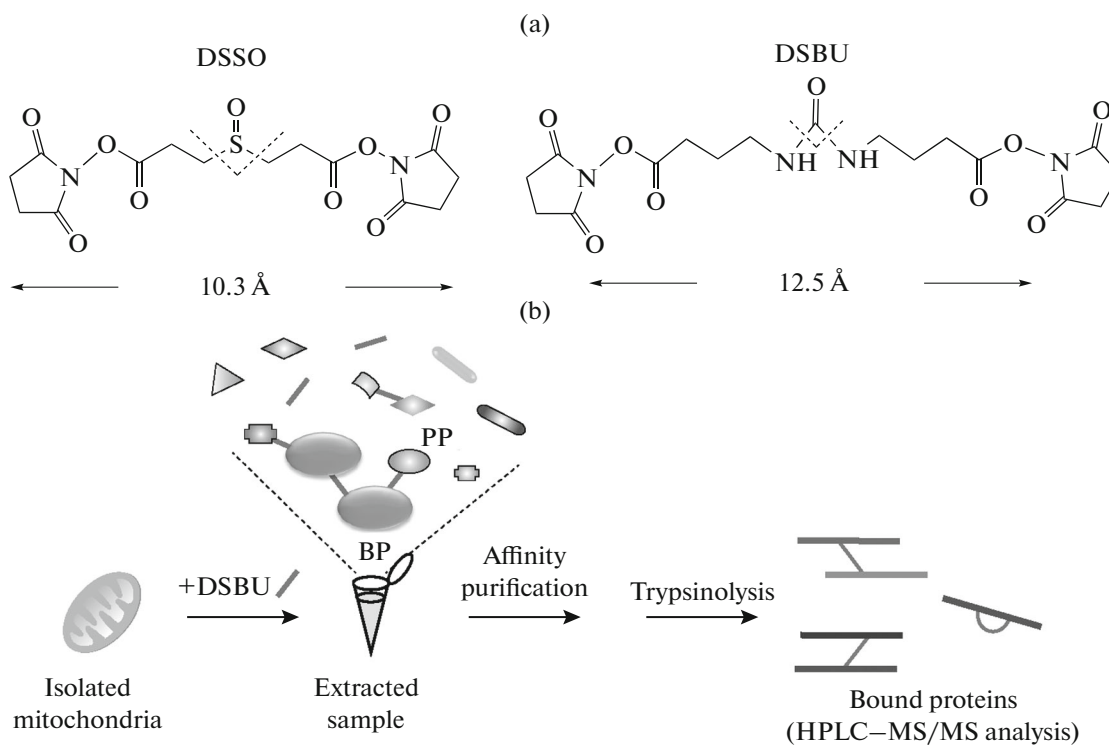

Fig. 6. (a) Chemical structures and lengths of DSSO and DSBU cross-linking agents: (dashed lines) the sites of bond cleavage under CID conditions. (b) Schematic diagram of an experiment using cross-linking reagents interactions: bait protein (BP) is a protein of a known structure, which is used to "catch" promiscuous protein partners in protein−protein interactions; prey protein (PP) is a promiscuous protein that interacts with a bait protein. Reproduced with modifications from [57] with permission. Copyright © 2019, Oxford University Press.

food is prepared from such products. Naturally, peptide profiles can be used as biomarker indicators for detecting pathological processes in the body and various diseases [60]. However, this field can obviously include studies aimed at synthesizing and comprehensively studying the properties of abiogenic oligopeptides.

Various extraction and fractionation protocols were used to study endogenous oligopeptides [58, 59], and HPLC and CE are valid for online separation before mass spectrometric analysis. Naturally, the same mass spectrometric platforms are the most effective in peptidomics as in proteomics, namely, those based on ESI, MALDI, SALDI, and with the involvement of tandem options using CID or ETD. The use of MS-IM seems promising, which makes it possible to determine the structural features of oligopeptides and the position of post-translational modifications (PTMs) in their molecules [61].

Oligopeptides can be detected directly in tissues and even isolated cells using direct MALDI-MS analysis. In this case, a biological test sample is applied to the target, to which a suitable matrix is admixed, facilitating the ionization of the peptides. For example, neuropeptides in tissues and single neurons can be analyzed [62]. A similar approach can be used for spatially separated visualization of peptides by MALDI-MSI [63].

Metabolomics is an independent scientific field dealing with the qualitative and quantitative determination of low-molecular-weight metabolites in the body caused by enzymatic reactions and interaction with each other, which is the body's metabolism [64]. Therefore, "metabolome" is a complete set of many thousands of low-molecular-weight metabolites in a biological cell, tissue, organ, or organism as a whole. The metabolome is the last link in the transmission of information in the chain genome → transcriptome → proteome → metabolome (Fig. 1). Ultimately, it reflects the course of physiological and pathological cellular processes, which is essential for medical diagnostics. The composition of the metabolome is highly complex because of both the variety of structural types with different physicochemical properties (amino acids, biogenic amines, carbohydrates, lipids, steroids, nucleosides, etc.) and the dramatic difference in the concentration of individual metabolites.

The research process in metabolomics includes the same stages as in proteomics, namely, the selection and preparation of a biological sample (cells, tissue, biological fluid), isolation of metabolites (various methods of liquid– or solid-phase extraction), multicomponent analysis using, most often, a combination of chromatographic techniques and mass spectrometry (GC–MS, GC×GC–MS, LC–MS), and data collection and interpretation using bioinformation systems. Conventional GC–MS is still a reliable tool for analyzing high-volatile metabolites (alcohols, aldehydes, ketones, organic acids, and many others) and derivatized medium-volatile compounds. Among medium-volatile compounds, carbohydrates, amino acids, steroids (see the section on steroidomics), terpenes, and higher fatty acids (see the section on lipidomics) are of the greatest interest; they can be preliminarily derivatized by silylation, alkylation, acylation, and other reactions [38, 65]. Nowadays, LC–MS combinations (HPLC–MS, UPL–MS) are used more frequently, enabling studying a more comprehensive range of metabolites, including polar ones, without preliminary derivatization. However, in conducting quantitative analyzes by isotope dilution, derivatization of target functional groups (NH_2, COOH, OH, SH, carbonyl) greatly facilitates the task of highly reliable detection and profiling of various types of metabolites with ultrahigh sensitivity [66, 67]. Combinations of MS with HILIC and HILIC/HPLC are rather effective in analyzing highly polar metabolites [68], and the CE/ESI-MS method is preferable for the analysis of polar and charged metabolites [69].

In the general case, both "target" (known metabolites) and "off-target" (new metabolites) analysis can be carried out using mass spectrometry [70]. The targeted analysis of metabolites usually involves their identification and quantification. An off-target analysis is limited to detecting unknown metabolites with, where possible, the determination of their structure and profiling.

A detailed analysis of a metabolome containing a vast number of various structural types of metabolites is virtually impossible without automatic identification using standard mass spectral databases. Such databases continue to be developed, and, along with search engines, they usually contain both ordinary and MS/MS spectra of known identified metabolites. Whenever possible, they contain the theoretical MS/MS spectra of metabolites, predicted from metabolic reactions but not yet detected, constructed based on fragmentation patterns (in silico). Among such databases, the Human Metabolome Database (HMDB) is by far the most important and informative; the latest version of HMDB 4.0 was released in 2018 [71]. It contains 7418, 2544, and 26 880 experimental, expected, and predicted GC–MS mass spectra, respectively, and 22 198, 2265, and 279 972 of the MS/MS spectra of the same origin. It also includes a large number of detected, expected, and predicted human metabolites. We should also pay attention to the development of an approach to the automatic deconvolution of GC–MS data and the identification and classification of mixture components, which is essential for metabolic studies. The main goal of these procedures is to enable the scientific community to perform various manipulations with these data within the framework of the molecular network analysis platform (Global Natural Product Social Molecular Networking, GNPS) [72].

Metabolic studies based on various approaches of gas chromatography—mass spectrometry are becoming more widespread in clinical practice [73]. The main focus here is on the early diagnosis of various diseases, including cancer [74], diabetes [75], and heart disease [76].

A special section of metabolomics, pharmacometabolomics, is also of great interest, with the main task to predict the effect of drugs on the profile of metabolites using a mathematical model. Mass spectrometry, indeed, plays a dominant role in this field as a method of analysis [77]. The results of mass spectrometric imaging methods, yielding the distribution of drugs and their metabolites over the body's tissues, are the most interesting [78].

Lipidomics is usually considered as a particular branch of a more general scientific field, metabolomics. It seems that the term "lipidome" appeared in the publications in 2001 to denote a complete set of lipids in a biological system. The term "lipidomics" began to be widely used in 2003; it distinguished the science dealing with the complete identification and quantitative analysis of biogenic lipids in the cell lipidome at the molecular level and the study of their interaction with other metabolites, lipids, and proteins [79]. Currently, lipids are divided into eight main categories: fatty acyls (for example, fatty acids), glycerolipids, glycerophospholipids, sphingolipids, prenols, sterols and other steroids, saccharolipids, and polyketides. Each group consists of subclasses, their molecules containing specific functional groups are characterized with different polarities and vary in size. In total, 43 636 lipids are included in the well-known LIPID-MAPS Structure Database (LMSD), among which more than 21 000 have a known structure and about 22 000 are computer-generated. The study of a particular type of lipids is sometimes referred to as "target" lipidomics.

Highly sensitive MS is the primary experimental method of lipidomics [80, 81]. Early studies that can be attributed to this type of omics were carried out using GC—MS. However, with the invention of ESI, the corresponding mass spectrometric version began to be intensively introduced. As applied to studies by MS, lipidomics is divided into two types: "off-target" (global) and "target." The former is to detect and quantify all lipids (known and unknown) in a lipidome. The latter type of lipidomics aims to detect specific lipid components in the study of complex biological problems. Another particular type of lipidomics is based on mass spectrometric imaging (MS imaging), which includes studying the distribution of lipid components in cells and tissues.

In the last 10—15 years, various LC—MS versions have been most developed, which, in general, provide global lipidome profiling [82]. Lipidome analysis generally includes extraction of lipids from a sample (liquid-phase extraction with mixtures of chloroform or

methyl (*tert*-butyl) ether with methanol, or solid-phase extraction), sample processing (for example, derivatization), and acquisition and processing of mass spectral data. The mass spectrometric termination uses ESI, high resolution, positive or negative ion recording mode, or tandem capabilities. Online LC—MS combinations use normal-phase and reversed-phase LC and HILIC. Several databases have been created for the automatic identification of lipidome components, containing MS/MS spectra for known lipid structures and theoretically generated fragment spectra of candidates, as well as search engines [83].

Along with this methodology, the so-called "shotgun" lipidomics has become widespread, including mass spectrometric analysis of a crude extract of a biological sample, injected directly into an ESI source without preliminary chromatographic separation. This direction was initiated by the first works carried out in the last decade of the XX century [84]. To identify lipid components by shotgun lipidomics, it is convenient to use tandem MS or high-resolution mass spectrometry. MALDI-MS also has significant analytical potential in lipidomics [80].

An applied method for an exhaustive study of lipids in cells and tissues is the combination of LC or HILIC with MS-IM [85, 86]. In this case, lipids of the glycerolipids, sphingolipids, and glycerophospholipids types, in ionic form, are separated and detected based on the ionic mobility of their molecules, reflecting the size, volume, charge, and mass, while their subclasses are separated using HILIC. In many cases, MS-IM helps to determine the classes and subclasses of lipids, the degree of unsaturation of acyl residues, and the position of double bonds, to distinguish the type of their attachment to the glycerol base (sn-regioisomerism), etc., by analysis of the drift time and diagnostic fragment ions.

Tandem mass spectrometry (ESI) with UVPD proved to be very useful for studying the fine structure of acyls in the lipidome. This method reliably determines the position of unsaturation in mono- and diunsaturated fatty acids with nonconjugated and conjugated double bonds, as well as in polyenoic acid molecules [87].

Online or offline derivatization significantly facilitates the determination of the structures of target lipids by various combined chromatographic methods and MS. In particular, various types of chemical modification are effective in analyzing fatty acids and acyls by GC—MS, LC—MS, and MS imaging [88]. Online ozonation has been widely used to determine the double bond position in acyl residues and fatty acids by LC—MS. Aldehyde ions formed upon ozonation and splitting of double bonds are characteristic ions here. In particular, this strategy proved to be effective in the analysis by LC combined with MS-IM [89], making it possible to obtain additional information about the structure and identify the presence of saturated variet-

ies of acid residues. An interesting chemical modification for determining the position of double bonds in acyl residues of phospholipids is the online photochemical reaction of Paternò Büchi with carbonyl compounds (for example, acetone) in a flow-through reactor located between the outlet from the LC column and the inlet to the mass spectrometer. The zone of the microreactor is irradiated with UV light, each eluent undergoes a [2+2] cycloaddition reaction with a ketone, and the resulting four-membered rings under the conditions of a CID mass spectrometer are broken in half, forming characteristic ions [90].

We should also note the online combination of TLC with MALDI-MS, applied in lipidomic studies [91]. Recently, ambient mass spectrometry has been used especially frequently for lipid analysis, which reduces sample preparation to a minimum. The high ionization efficiency of lipid molecules makes it possible to avoid the matrix effect in relation to them. For example, MS with desorption electrospray ionization (DESI) has been successfully used to visualize triglycerides in tissues [92].

Another example of the application of this approach is the use of a method that involves taking a sample with a needle, on which a sample is applied in one way or another (Fig. 7). Then the needle is moistened with a solvent and placed at the entrance cone of the mass spectrometer, after which a voltage drop is created between the cone and the needle, causing desorption of microdroplets with analyte ions [93]. The method is called Touch Spray mass spectrometry.

Less than 10 years ago, the first results of profiling of metabolites and peptides at a one-cell level were reported [94]. It is evident that the unique capabilities of MS will be actively used in lipid studies as well.

Steroidomics is another type of independent omics subclass of lipidomics. A steroidome is a complete set of endosteroid metabolites in a cell or body. The term steroidomics was first proposed in 2004 [95] to denote the scientific industry involved in the detailed characterization and quantitative analysis of steroids (estrogens, corticosteroids, bile acids, oxysterols, various steroid conjugates, metabolites, etc.) in the metabolic profile. Steroidomics should be separated from lipidomics into an independent subgroup because of the fundamental role of steroids in human physiology, the formation of cell structures, signal transmission, and, in particular, their participation in the processes of the onset and development of many diseases, especially, oncological (oncosteroids) [96].

At present, virtually no analytical method can give comprehensive qualitative and quantitative information on the entire composition of the steroid in a single experiment. GC−MS and HPLC−MS, including tandem versions (MS/MS), are most often used to detect and profile steroids in biological fluids and tissues [97, 98]. Currently, both methods can be used for target and off-target steroid assays, including qualitative and
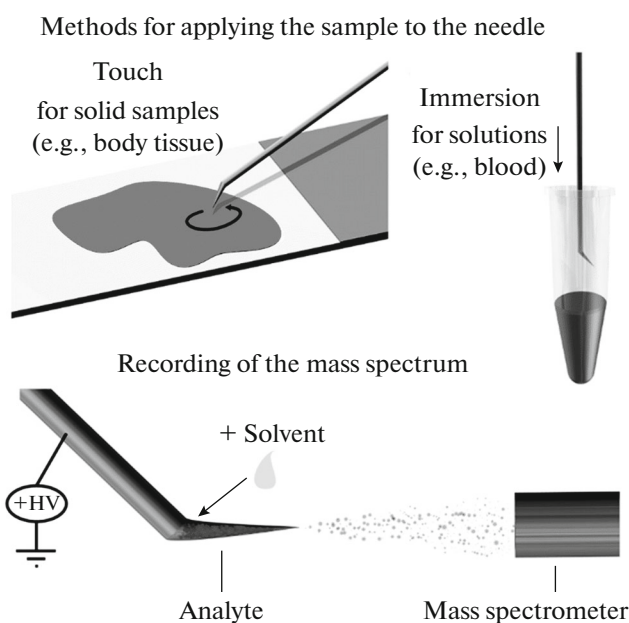

Fig. 7. Principles of touch spray mass spectrometry. Reprinted with modifications from [93] with permission. Copyright © 2014 The Royal Society of Chemistry.

quantitative determinations. The quantitative analysis of steroids is carried out by isotopic dilution using internal standards in the form of ^{2}H-labeled or, better, ^{13}C-labeled steroid structures. When analyzed by GC−MS, steroids and their metabolites are preliminarily released from conjugates (glucuronides, sulfates) by enzymatic hydrolysis.

Because of a significant variety of structural types of the determined steroid systems, each of the above methods is usually limited to determining relatively narrow structurally related classes. The use of derivatization or the involvement of specific online extraction and separation procedures can expand the range of simultaneously analyzed steroids [99]. For example, Chortis et al. [100] applied GC−MS in combination with preliminary enzymatic hydrolysis of primary steroid conjugates followed by the derivatization yielding methyloximes of trimethylsilyl esters [38] to detect adrenocortical carcinoma by analyzing metabolites of the corresponding endogenous steroids. Steroidomics technology was also used to study the metabolic process leading to the formation of bile and other steroid acids in the cerebrospinal fluid.

Recently, a combination of liquid chromatography (HPLC or UHPLC) and tandem mass spectrometry based on ESI, APCI, and APPI has been increasingly used for steroid analysis. The combined LC−MS method is convenient because it enables simultaneous determination of steroid structures of different polarities, often without additional derivatization. However, preliminary derivatization is still used for this type of analysis, using a targeted modification of steroid sub-

structures. Such approaches, including the conversion of carbonyl groups by reaction with hydroxylamine or 2-hydrazinopyridine, were proposed for the highly sensitive determination of endosteroids by HPLC/ESI-MS and UHPLC/ESI-MS [101], respectively. We can also note a preliminary modification of carbonyl groups with the introduction of a fixed charge (Girard's reagent), carried out to detect and determine off-target steroid metabolites by a combined method of LC with high-resolution MS (LTQ-Orbitrap) [102]. In another case, the principles of steroidomics based on UHPLC in combination with high-resolution MS (QTOF) were involved in the process of simultaneous identification and profiling of large arrays of "exogenous" metabolites formed from anabolic steroids as a method of doping control [103]. Note that LC−MS can also be used to determine native steroid conjugates, in particular, sulfates [104].

Glycomics. The term "glycan" refers to any carbohydrate or saccharide that is free or associated with another noncarbohydrate molecule. The complete set of such glycans and glycoconjugates in the body is called "glycome." The science of "glycomics" deals with studying the structure of glycans and their participation in various genetic and physiological processes.

The most common glycoconjugates belong to a series of glycoproteins, where N-glycans (the carbohydrate residue is attached to the side chains of asparagine or arginine) and O-glycans (the sugar residue is attached through the side hydroxyl group of serine, threonine, tyrosine, and some other amino acids) are distinguished. As noted in the section on proteomics, the formation of the corresponding conjugates (glycosylation) is one of the most common post-translational modifications (PTMs) of proteins, which are the subject of research in glycoproteomics (determination of the nature of glycans and the place of their attachment in protein molecules). Glycans in cells can also be attached to lipids (glycolipids) or exist as polysaccharides in a free state.

To determine the structure (sequence, branching) of oligosaccharides attached to proteins, one can use their preliminary enzymatic cleavage from protein molecules with a subsequent study by the tandem MALDI-MS method in combination with derivatization [38]. The monosaccharide composition of such glycans and free oligosaccharides in the cell is found by GC−MS methods after derivatization (silylation, acylation) of the cleaved molecules. MALDI-MS/MS or LC−ESI-MS/MS is effective in determining the complete structure of both carbohydrate and protein parts of intact glycoproteins [105, 106]. D. Harvey published a series of detailed reviews devoted to the application of MALDI-MS to the analysis of carbohydrates and glycoconjugates (see the latter one [107]). MALDI-MS/MS yields a large amount of information on the structure of both saccharides isolated from glycolipids enzymatically and intact conjugates [108].

One of the problems of glycomics is the difficulty of distinguishing carbohydrate stereoisomers with many asymmetric centers. Combinations of mass spectrometry with ion mobility spectroscopy have excellent prospects in this case, because the latter enables the separation of stereoisomeric forms by their different spatial structures (Fig. 8) [109].

We should mention the development of glycoinformation computer platforms, databases, and learning algorithms to interpret glycomic studies automatically and predict PTM protein sites. Many of them are based on the use of various mass spectrometric data [110].

2. OMICS SCIENCES RELATED TO THE APPLICATION OF A COMPLEX OF APPROACHES TO SPECIFIC SAMPLES

2.1. Plantomics

Plantomics should be attributed to life sciences. In general, it combines a variety of suitable omics discussed above for life sciences, such as proteomics, peptidomics, and metabolomics [111]. In plantomics research, such mass spectrometric methods as GC−MS, LC−MS, MALDI-MS, and MALDI-MSI are effectively used, as well as methods of ionization in air, that is, desorption ESI and ESI with laser ablation.

The main purpose for determining the composition of compounds of plant raw materials is to study the physiology of plants [112]. In addition to solving fundamental problems, applied goals related to detecting and identifying biologically active compounds for their further use as medicines are often pursued [113]. As most of these compounds are polar, they are studied using "soft" mass spectrometric methods combined with tandem mass spectrometry. As in other omics studies, the obtained data sets are processed using various computer algorithms [114].

Plantomics, as other omics related to the study of living organisms, both mass spectrometric imaging, which enables the study of the distribution of bioactive compounds in plant tissues [115], and ambient mass spectrometry are widely used. To study plants, a particular version of this method was developed, ionization from the leaf. In this case, a fragment of plant material is moistened with a solvent and placed in a special metal clip near the entrance cone of the mass spectrometer [116]. A voltage drop between the clip and the cone ensures the desorption of microdroplets containing compounds that are part of the plant material (Fig. 9). This approach makes it possible to detect both endogenous [117] and exogenous compounds [118].

The study of lignin structure is an essential field of plantomics, often singled out as a separate science. This biopolymer is the second most widespread after cellulose and is one of the most promising renewable sources of raw materials for processing to obtain fuels
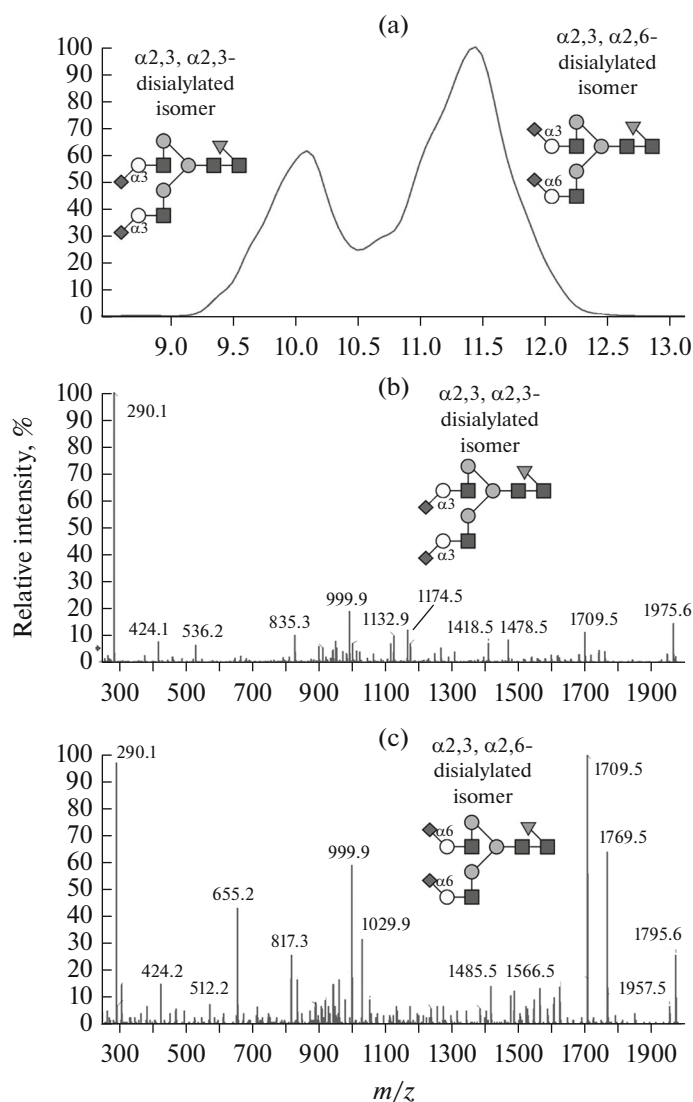
Fig. 8. (a) Separation of doubly deprotonated ions of disialylated glycan isomers and (b, c) mass spectra of individual isomers. Reprinted with modifications from [109] with permission (ACS AuthorChoice Licens). from ACS Publications. Copyright © 2019 American Chemical Society.

and organic synthesis products. Although lignin formally consists of a small set of monomer units, the compounds obtained by polymerization have complex nonlinear structures, which makes them extremely difficult to study. Another problem is the low solubility of lignins, limiting the use of many mass spectrometric methods in omics sciences [119]. The ionization efficiency of different types of lignin molecules and products of lignin degradation also differ significantly; therefore, it is recommended to use several methods for its analysis with the goal of achieving complementary data sets [120].

A separate problem is the use of MALDI mass spectrometry in the study of lignin. Theoretically, this biopolymer is a rather convenient sample for studying by this method, especially because lignin molecules resemble widely used matrix compounds in structure. Another argument favoring MALDI is the low sensitivity of this method to the type of solvent used for sample preparation. However, applied results obtained using MALDI mass spectrometry to study lignin show a relatively low efficiency of desorption/ionization of analytes, while the use of ionic liquids leads to an increase in the ion yield [121]. Nevertheless, the development of approaches for using this method in ligninomics remains a challenge for further work in this field.

2.2. Exposomics

The main task of exposomics is to study the effect of environmental factors on human health. The concept of the environment, in this case, is interpreted
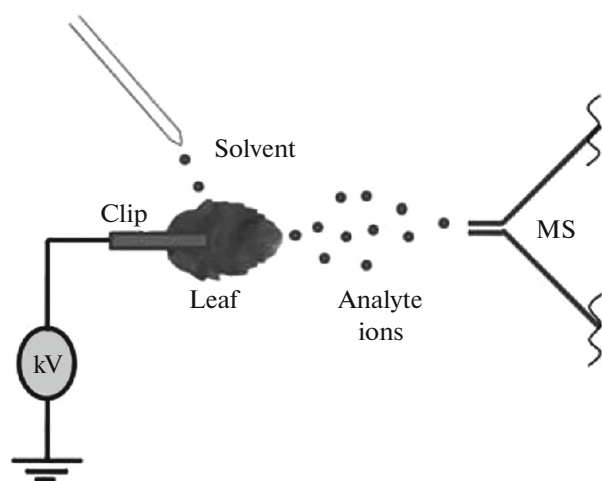
Fig. 9. Ionization from a leaf. Reprinted with modifications from [117] with permission. Copyright © 2012 The Royal Society of Chemistry.

quite broadly, covering all types of effects on a person throughout his life. The sum of such impacts forms an exposome, which includes the impact of endogenous and exogenous processes at the individual level (for example, environmental pollution and infection), as well as general impacts at the global level (for example, climate and socioeconomic factors) [122]. From the point of view of analytical chemistry, exosomics is the study of changes in the metabolome, proteome, transcriptome, and genome of a person as a function of the state of the environment [123]. This approach predetermines the toolkit of this area of research; therefore, it involves all the methods demanded in the listed omics approaches and the methods of environmental analysis and monitoring.

One of the main directions of exposomics is associated with the study of the effects of various ecotoxicants on the human body [122]. An essential task, in this case, is to determine a set of compounds with which the human body interacts. In conventional ecological analysis, the main emphasis is placed on the target determination of compounds included in the list of controlled ones. In exosomics, off-target analysis acquires critical importance, implying the detection and identification of the maximum available amount of organic molecules present in the environment. Only mass spectrometric approaches ensure the achievement of the required limits of detection and the information content of the data obtained [124].

To solve this problem, methods based on gas chromatography-mass spectrometry and approaches with atmospheric pressure ionization are used. Electron ionization GC−MS is best suited for off-target analysis because of a possibility of using mass spectral databases for identification and "manual" interpretation of mass spectra based on the array of already obtained data [125]. To increase the sensitivity and reliability of

the identification of mixture components, two-dimensional gas chromatography and high-resolution mass spectrometry are actively used [126]. The use of a combination of these methods in the original work devoted to determining the sources of the appearance of substituted pyridines in the atmosphere made it possible to identify a set of peat combustion products [127]. Even larger informational content is achieved using complementary ionization methods: the soft method yields an exact empirical formula of an analyte, and electron ionization enables the determination of its structure. This approach, for example, made it possible to identify more than 500 components in snow samples, including aromatic compounds of various structures and phthalates [128].

The use of atmospheric pressure ionization methods combined with tandem experiments and/or high-resolution mass spectrometry is also helpful for the off-target analysis of ecotoxicants [129]. However, the combined use of these methods and GC−MS varieties to obtain a complementary set of data on polar and weakly polar compounds seems much more efficient [130]. For example, using this combination in the off-target analysis of human milk made it possible to detect various halogen derivatives in it using a simple sample preparation procedure (Fig. 10) [131].

Another field of research is the study of changes that occur in a human body under the action of the environment. Metabolic approaches are in especial demand in this case [132], which makes it possible to detect transformation products that have entered the body of xenobiotics [133], to find changes associated with diseases of diverse nature [134], and to identify the causes of personality disorders [135]. However, proteomic tools are also helpful to determine the contribution of the environment to the state of the human body. For example, heatstroke was accompanied by the production of a series of specific peptides [136].

3. OMICS SCIENCES IN INANIMATE NATURE

3.1. Petroleomics

As mentioned above, some sciences dealing with the study of inanimate nature also began to be referred to as omics. These scientific fields deal with vast amounts of data and pursue goals similar to those that are characteristic of life omics. Probably, Marshall's proposal to single out "petroleomics" as a separate scientific area, which is aimed "at an exhaustive description of all chemical compounds of crude oil and their interactions" [137], was perfectly legitimate.

To date, no universal mass spectrometric method has been developed that could adequately describe the total composition of crude oil. Indeed, "petroleome" is determined by hundreds of thousands of compounds that differ dramatically in structure, elemental composition, molecular weight, volatility, polarity, and ionization efficiency: homologous and isomeric
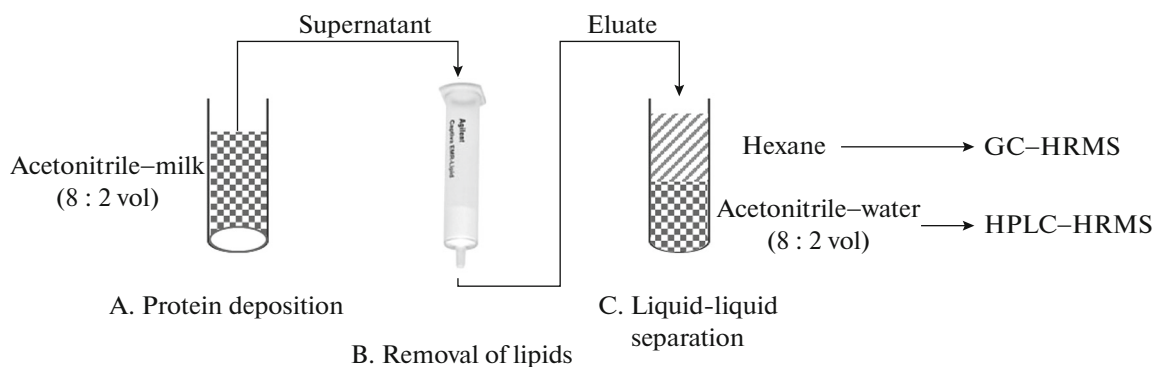
Fig. 10. Sample preparation for off-target analysis of human milk using a complex of complementary mass spectrometric methods. Reprinted with modifications from [131] with permission. Copyright © 2020 Elsevier B.V.

saturated aliphatic, alicyclic, and alkylaromatic hydrocarbons; N-, S-, O-, and metal-containing compounds; naphthenic acids; resins; asphaltenes; etc. So far, it is necessary to use those mass spectrometric varieties suitable for studying relatively narrow structural and functional groups of petroleum compounds. Several reviews have been published on the use of mass spectrometry to solve various oil chemistry and petroleum science problems. Let us mention only comparatively recent ones [138, 139].

Petroleum hydrocarbons were actually the first organic molecules to be studied by mass spectrometry [17]. Later, this method was applied to other volatile petroleum compounds that can be converted to the gas phase. During this period, the mass spectrometric method for determining the structures of the non-group composition of unseparated petroleum fractions has been developed quite intensively [140]. It was based on the mathematical processing of large arrays of total spectra of fractions, classifying this strategy as petroleomics. However, comprehensive and voluminous data can be obtained using the combined GC–MS method, capable of performing a component-wise analysis of oil fractions boiling up to 400–500°C. Until now, EI mass spectrometry (GC–EI-MS) remains the most sensitive, reproducible, and quantitative in studies by this method, although impressive results can be obtained using MS based on CI, APCI, PI, photoionization, or photochemical ionization. GC–EI-MS is also widely used to analyze volatile petroleum hydrocarbons (alkanes, cyclanes, isoprenanes) and relatively low-molecular-weight hetero-organic compounds separated by GC. This type of research includes, for example, popular determinations of sterane and triterpane petroleum biomarkers [141] and adamantanoid structures [142].

The structural and analytical capabilities of gas chromatography–mass spectrometry have significantly expanded with the introduction of two-dimensional GC (GC/GC–MS), which includes a series connection of two columns with stationary phases of different polarity. Corresponding combined systems

enable a more detailed study of mixtures by separating each zone isolated by the first column and detecting isomeric, minor, and other components that cannot be detected by a single GC–MS method.

The GC–MS method in various versions is virtually unsuitable for studying polar and nonvolatile oil components, which are largely present in heavy fractions. At the same time, the widely used ESI-MS, MALDI-MS, and APPI-MS are suitable for the ionization of such compounds. Recently, for the analysis of heavy oil fractions, methods of ionization in air (ambient mass spectrometry), for example, direct analysis in real time (DART) [139], have also begun to be used. All these methods can be applied to the study of both total and preseparated heavy fractions.

In general, such a methodology results in a detailed knowledge of the composition of both crude oil and especially heavy residues, the interest in which is constantly growing because of the expected depletion of oil fields. The most informative in such studies is characterized by ultra-high resolution mass spectrometry based on Fourier-transform ion cyclotron resonance (FT-ICR). That is why petroleomics was singled out as a separate science by A. Marshall, the most known scientist applying this mass spectral technique to oil research [137].

The principal advantage of the FT-ICR–MS method is a possibility of recording and analyzing multiline mass spectra of the total fraction of crude oil, the mass numbers of peaks in which are determined with high accuracy. This makes it possible to detect separately isobaric molecular ions and determine the structural elements of thousands of components of the most complex mixtures (various hydrocarbons, O-, N-, and S-containing structures). The latest achievement was the assignment of more than 240 000 unique gross formulas to analyze heavy oil residues, achieved using special software to optimize the data collection process [143]. As an example, we mention the works based on the application of this method for a detailed study of asphaltenes [144], naphthenic

acids [145], porphyrins [146], sulfur-containing compounds [147], and bitumen [148].

In conclusion of this section, we should pay attention to specific approaches to the representation and visualization of ultrahigh-resolution mass spectra (by now already up to tens of millions), containing tens of thousands of separated peaks of ions of homologous, isobaric, carbo-, and hetero-organic compounds. Along with statistical analysis, Kendrick mass defect diagrams and van Krevelen diagrams [8, 149] are the most useful; they were proposed quite long ago but turned out to be especially necessary for the analysis and comparison of data in petroleomics. These diagrams visualize a series of homologues, compounds with different degrees of unsaturation and aromaticity, with different heteroatoms and ratios, and other structures.

3.2. Polymeromics

Synthetic polymers, like biopolymers, are complex mixtures that contain a large number of structurally similar and differing macromolecules. The characterization of macromolecules and the determination of their structural features require unique technologies that have been classified as "polymeromics" [150]. The field of polymeromics research includes the determination of molecular weight distribution (MWD), polydispersity (PD), nature of end groups, monomeric composition, sequence of units, and branching in the main macrochain.

Online pyrolytic gas chromatography–mass spectrometry should be considered a long-standing omics platform for mass spectrometric studies of synthetic polymers [151]. In this case, the polymer sample is thermally decomposed in a pyrolyzer installed at the inlet of a gas chromatograph connected to a mass spectrometer. The resulting relatively low-molecular-weight products, which are monomeric, *n*-meric, and other pyrolysis products, are separated in a chromatographic column; the components are detected and identified by a mass spectrometer. The corresponding mathematical processing of the recorded chromatograms makes it possible to find the nature of the polymer microstructure, and in the case of copolymers, to quantify the concentration of homo- and heterodiads, triads, and other sequences in the macrochain, enabling the differentiation of statistical, alternating, and block copolymers. This experimental strategy can be attributed to bottom-up polymeromics.

With the creation of mass spectrometry with soft ionization (ESI, MALDI), it became possible to carry out polymer studies according to the top-down type. However, because of the specificity of these methods' ionization principles, nonpolar polymers (for example, hydrocarbon polymers) cannot be directly analyzed. At the same time, many high-molecular-weight molecules containing heteroatoms or functional groups, capable of protonation, are convenient samples for investigation by these methods. One of the essential characteristics obtained in this case is the molecular weight distribution (MWD) and the monomeric composition of macromolecules. The problem, however, is that modern routine mass spectrometers provide measurements of limited masses, somewhere around 5000 Da.

At the same time, tandem mass spectrometry with various methods of induced dissociation (CID, ETD, ECD, PSD) of primary macromolecular ions can reveal the nature of end groups, monomeric composition, and sequence of units in copolymers, branching in macrochains, and other significant structural elements [152−154]. A polymer sample before mass spectrometric analysis can be separated into fractions using LC, 2D-LC, SEC, and others for a more efficient study.

Methods of chemical derivatization are also successfully used in MS studies in polymeromics [155, 156]. A typical example of such a strategy is given in [156], where, using preliminary acetylation, the nature of the end groups was determined and linear and cyclic polyalkylene glycol macromolecules were differentiated (Fig. 11).

3.3. Foodomics

Although food and nutrients, for the most part, come from living organisms, a separate discipline, "foodomics," should be classified as a science of inanimate nature [157]. Its purpose is to study food and nutrients to ensure human health by determining quality, safety, preservation using omics and bioinformation technologies.

Foodomics analyzes complex food items containing nutrients, exogenous dietary supplements in a wide variety of concentrations. Mass spectrometry is the essential analytical platform for such studies, using, particularly, the strategies of genomics, proteomics, lipidomics, and metabolomics discussed above [158].

One of the main tasks in the application of genomic approaches to foodomics is to study the origin of various types of food sources [159]. Nutrigenomics is another important field that has recently become an independent area of knowledge. In nutrigenomics, works aim to determine the effect of consumed food products on the human genome and study a possibility of developing personalized dietary nutrition and the effect of diets on human health [160]. Metabolic and proteomic methods, used to solve these problems, help assess the processes occurring in the human body under the effect of food preferences [161]. For example, the use of off-target metabolome analysis was used to identify biomarkers of red meat consumption that may be responsible for the development of type 2 diabetes [162].

Another goal of foodomics is to control the quality and safety of food. Such tasks include, for example, the study of food allergens. Most of them are of protein origin; therefore, various proteomics methods are used to detect them [163], and the field of knowledge itself has been proposed to call allergenomics [164]. Although this area of research goes beyond foodomics, as it studies the mechanism of allergic reactions and the search for effective protocols for diagnosing allergies, part of this field is devoted to determining the presence of allergens themselves in food. Such studies acquire particular value in connection with the introduction of genetically modified organisms into circulation: the undesirable effects of such a modification can also be expressed in the production of allergens [165].

The use of omics technologies for confirming the origin of food products is also in demand. For example, the study of the composition of polar compounds of natural olive oils using ESI mass spectrometry and statistical analysis identified the markers of their geographical origin [166].

3.4. Humeomics

Humus is formed by chemical and biological degradation of remains of plant and animal origin [167]. Modern concepts consider humus as a supramolecular system formed by the self-assembly of relatively low-molecular-weight heteroatomic compounds. Such a system is hydrogen-bonded and stabilized through the formation of metal-containing complexes and adsorption on the clay surface [168]. As the formation of compounds and supramolecular complexes depends on many factors, studying a humeome is one of the most challenging problems in analytical chemistry.

The solution usually involves a combination of methods, including the preliminary separation of a sample and various analytical approaches [169]. The most informative of these is ultrahigh-resolution mass spectrometry in combination with various ionization methods [170]. In particular, FTICR-MS made it possible to assign more than 6500 unique gross formulas in the study of compounds of humic origin [171]. Comparative studies have shown that the absence of a generally accepted protocol for such studies significantly affects the interlaboratory reproducibility of results [172].

One of the reasons for the interest in humic substances is their use for increasing soil fertility [173] and possible antiviral [174] and other biological activity [175]. The use of mass spectrometry, in this case, makes it possible to determine the relationship between the composition of the samples used and the results of their application.
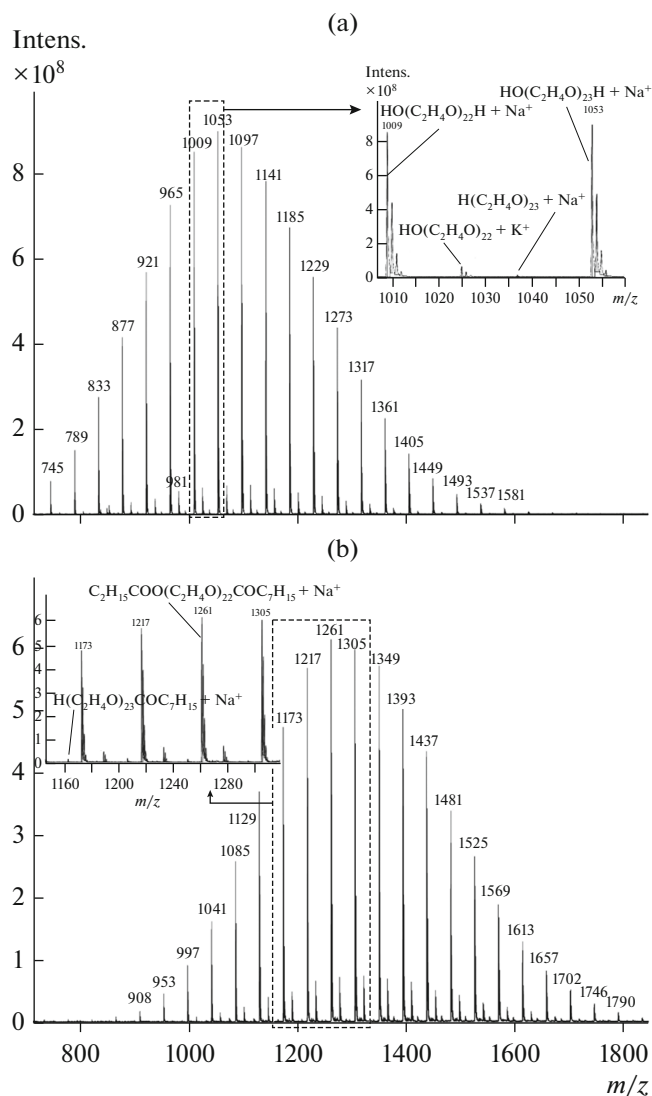

Fig. 11. MALDI mass spectra of (a) polyethylene glycol and (b) products of its derivatization with capryloyl chloride.

CONCLUSIONS

Within the framework of this review, we considered only briefly the fundamental analytical capabilities of mass spectrometric platforms in those omics where they are indispensable or used diversely. Because of the lack of volume, multi-omics technologies based on the application of several approaches in one study, scientific branches from broader omics (for example, proteogenomics, glycoproteomics, phosphoproteomics, lipoglycomics), specific omics in pharmacology (pharmacogenomics, pharmacoproteomics, vaccineomics, nutrigenomics), and others were only mentioned or entirely out of the review. In many of them, mass spectrometry also plays an essential analytical role. With the development of separation, mass spectral, and bioinformatics techniques, omics studies in

the life sciences are increasingly moving to the level of analysis of one or several cells, and their role becomes invaluable in clinical diagnostics.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

AUTHOR CONTRIBUTIONS

All authors contributed to the work equally.

ADDITIONAL INFORMATION

Vladimir Georgievich Zaikin, ORCID 0000000193608442.

Roman Sergeevich Borisov, ORCID 0000000282037055.

REFERENCES

1. Satya, S.P., *J. Biomol. Tech.*, 2007, vol. 18, no. 5, p. 277.

2. Alexandrov, M.L., Gall', L.N., Krasnov, N.A., et al., *Dokl. Akad. Nauk SSSR*, 1984, vol. 277, no. 2, p. 379.

3. Yamashita, M. and Fenn, J.B., *J. Phys. Chem.*, 1984, vol. 88, no. 20, p. 4451.

4. Karas, M., Bachmann, D., and Hillenkamp, F., *Anal. Chem.*, 1985, vol. 57, no. 14, p. 2935.

5. Karas, M. and Hillenkamp, F., *Anal. Chem.*, 1988, vol. 60, no. 20, p. 2299.

6. Picariello, G., Mamone, G., Addeo, F., et al., *Food Technol. Biotechnol.*, 2012, vol. 50, no. 3, p. 286.

7. Di Girolamo, F., Lante, L., Muraca, M., et al., *Curr. Org. Chem.*, 2013, vol. 17, no. 23, p. 2891.

8. Zhao, Y.-Y. and Lin, R.-C., *Chem.-Biol. Interact.*, 2014, vol. 215, p. 7.

9. Winkler, H.L., *Verbreitung und Ursache der Parthenogenesis im Pflanzen- und Tierreiche* (Distribution and Cause of Parthenogenesis in the Plant and Animal Kingdoms), Jena: Fischer, 1920.

10. McGettigan, P.A., *Curr. Opin. Chem. Biol.*, 2013, vol. 17, no. 1, p. 4.

11. Il'ina, E.N. and Govorun, V.M., *Bioorg. Chem.*, 2009, vol. 35, no. 2, p. 149.

12. Schurch, S., *Mass. Spectrom. Rev.*, 2016, vol. 35, no. 4, p. 483.

13. Ragoussis, J., Elvidge, G.P., Kaur, K., et al., *PLoS Genet.*, 2006, vol. 2, no. 7, 0920.

14. Luque-González, M.A., Tabraue-Chavez, M., Lopez-Longarela, B., et al., *Talanta*, 2018, vol. 176, p. 299.

15. Tretyakova, N., Villalta, P.W., and Kotapati, S., *Chem. Rev.*, 2013, vol. 113, no. 4, p. 2395.

16. Lan, M.-D., Yuan, B.-F., and Feng, Y.-Q., *Chin. Chem. Lett.*, 2019, vol. 30, no. 1, p. 1.

17. Wilkins, M.R., Sánchez, J.-C., Gooley, A.A., et al., *Biotechnol. Genet. Eng. Rev.*, 1996, vol. 13, no. 1, p. 19.

18. Zaikin, V.G. and Tret'yakov, K.V., *Istoriya mass spektrometrii v datakh* (History of Mass Spectrometry in Dates), Moscow: Vseross. Mass.-Spektom. O-vo, 2018.

19. Yates, J.R., *J. Mass Spectrom.*, 1998, vol. 33, no. 1, p. 1.

20. Aebersold, R. and Mann, M., *Nature*, 2003, vol. 422, no. 6928, p. 198.

21. Ponomarenko, E.A., Poverennaya, E.V., Ilgisonis, E.V., et al., *Int. J. Anal. Chem.*, 2016, vol. 2016, p. 1.

22. Medzihradszky, K.F. and Chalkley, R.J., *Mass Spectrom. Rev.*, 2015, vol. 34, no. 1, p. 43.

23. Liu, F., Lössl, P., Rabbitts, B.M., et al., *Mol. Cell Proteomics*, 2018, vol. 17, no. 2, p. 216.

24. Yates, J.R. III., Ruse, C.I., and Nakorchevsky, A., *Annu. Rev. Biomed. Eng.*, 2009, vol. 11, p. 49.

25. Cui, W., Rohrs, H.W., and Gross, M.L., *Analyst*, 2011, vol. 139, no. 19, p. 3854.

26. Pamreddy, A. and Panyala, N.R., *J. Appl. Bioanal.*, 2016, vol. 2, no. 2, p. 52.

27. Vyatkina, K., Wu, S., Dekker, L.J.M., et al., *J. Proteome Res.*, 2015, vol. 14, no. 11, p. 4450.

28. Karabacak, N.M., Li, L., Tiwari, A., et al., *Mol. Cell. Proteomics*, 2009, vol. 8, no. 4, p. 846.

29. Kislinger, Th., Gramolini, A.O., MacLennan, D.H., et al, *J. Am. Soc. Mass Spectrom.*, 2005, vol. 16, no. 8, p. 1207.

30. Zhen, Y., Xu, N., Richardson, B., et al., *J. Am. Soc. Mass Spectrom.*, 2004, vol. 15, no. 6, p. 803.

31. Ivanov, M.V., Bubis, J.A., Gorshkov, V., et al., *Anal. Chem.*, 2020, vol. 92, no. 6, p. 4326.

32. Tarasova, I.A., Masselon, C.D., Gorshkov, A.V., et al., *Analyst*, 2016, vol. 141, no. 16, p. 4816.

33. Gomes, F.P. and Yates, J.R III, *Mass Spectrom. Rev.*, 2019, vol. 38, no. 6, p. 445.

34. Chicooree, N., Unwin, R.D., and Griffiths, J.R., *Mass Spectrom. Rev.*, 2015, vol. 34, no. 6, p. 595.

35. Lanucara, F. and Eyers, C.E., *Mass Spectrom. Rev.*, 2013, vol. 32, no. 1, p. 27.

36. Xiao, H., Sun, F., Suttapitugsakul, S., et al., *Mass Spectrom. Rev.*, 2019, vol. 38, nos. 4−5, p. 356.

37. Palumbo, A.M., Smith, S.A., Kalcic, C.L., et al., *Mass Spectrom. Rev.*, 2011, vol. 30, no. 4, p. 600.

38. Zaikin, V. and Halket, J., *A Handbook of Derivatives for Mass Spectrometry*, Chichester: IM, 2009.

39. Chen, Z., Huang, J., and Li, L., *TrAC, Trends Anal. Chem.*, 2019, vol. 118, p. 880.

40. Alfaro, J.F., Gong, C.-X., Monroe, M.E., et al., *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 19, p. 7280.

41. Dapic, I., Baljeu-Neuman, L., Uwugiaren, N., et al., *Mass Spectrom. Rev.*, 2019, vol. 38, nos. 4−5, p. 403.

42. Pichler, P., Köcher, T., Holzmann, J., et al., *Anal. Chem.*, 2010, vol. 82, no. 15, p. 6549.

43. Everley, R.A., Kunz, R.C., McAllister, F.E., et al., *Anal. Chem.*, 2013, vol. 85, no. 11, p. 5340.

44. Wang, X., Shen, S., and Rasam, S., S, et al, *Mass Spectrom. Rev.*, 2019, vol. 38, no. 6, p. 461.

45. Bouhaddou, M., Memon, D., Meyer, B., et al., *Cell*, 2020, vol. 182, no. 3, p. 685.

46. Ihling, C., Tänzler, D., Hagemann, S., et al., *J. Proteome Res.*, 2020, vol. 19, no. 11, p. 4389.

47. Cardozo, K.H.M., Lebkuchen, A., Okai, G.G., et al., *Nat. Commun.*, 2020, vol. 11, no. 1, 6201.

48. Yao, Z.-P., *Mass Spectrom. Rev.*, 2012, vol. 31, no. 4, p. 437.

49. Chen, F., Gulbakan, B., Weidmann, S., et al., *Mass Spectrom. Rev.*, 2016, vol. 35, no. 1, p. 48.

50. Downard, K.M., *Mass Spectrom. Rev.*, 2016, vol. 35, no. 5, p. 559.

51. Ivanov, A.S., Ershov, P.V., Mezentsev, Yu.A., et al., *Biomed. Chem.*, 2013, vol. 59, no. 2, p. 171.

52. Feng, S., Zhou, L., Huang, C., et al., *Expert Rev. Proteomics*, 2015, vol. 12, no. 1, p. 1.

53. Yang, Y., Du, Y., and Kaltashov, I.A., *Anal. Chem.*, 2020, vol. 92, no. 16, p. 10930.

54. Gonzalez-Lozano, M.A., Koopmans, F., Sullivan, P.F., et al., *Sci. Adv.*, 2020, vol. 6, no. 8, eaax5783.

55. Sinz, A., *Anal. Bioanal. Chem.*, 2017, vol. 409, no. 1, p. 33.

56. Petrotchenko, E.V. and Borchers, C.H., *Mass Spectrom. Rev.*, 2010, vol. 29, no. 6, p. 862.

57. Koshiba, T. and Kosako, H., *J. Biochem.*, 2019, vol. 167, no. 3, p. 225.

58. Dallas, D.C., Guerrero, A., Parker, E.A., et al., *Proteomics*, 2015, vol. 15, nos. 5–6, p. 1026.

59. Govorun, V.M. and Ivanov, V.T., *Russ. J. Bioorg. Chem.*, 2011, vol. 37, no. 2, p. 176.

60. Maes, E., Oeyen, E., Boonen, K., et al., *Mass Spectrom. Rev.*, 2019, vol. 38, no. 3, p. 253.

61. Ibrahim, Y.M., Shvartsburg, A.A., and Smith, R.D., *Anal. Chem.*, 2011, vol. 83, no. 14, p. 5617.

62. Hummon, A.B., Amare, A., and Sweedler, J.V., *Mass Spectrom. Rev.*, 2006, vol. 25, no. 1, p. 77.

63. Minerva, L., Boonen, K., Menschaert, G., et al., *Anal. Chem.*, 2011, vol. 83, no. 20, p. 7682.

64. Nicholson, J.K., Lindon, J.C., and Holmes, E., *Chem. Unzerer Zeit*, 1999, vol. 29, no. 11, p. 1181.

65. Rodrigues, D., Pinto, J., Araújo, A.M., et al., *Metabolomics*, 2018, vol. 14, no. 5, p. 62.

66. Zaikin, V.G. and Borisov, R.S., *Crit. Rev. Anal. Chem.*, 2021, p. 1.

67. Zhao, S., Li, H., Han, W., et al., *Anal. Chem.*, 2019, vol. 91, no. 18, p. 12108.

68. Tang, D.-Q., Zou, L., Yin, X.-X., et al., *Mass Spectrom. Rev.*, 2016, vol. 35, no. 5, p. 574.

69. Ramautar, R., Somsen, G.W., and de Jong, G.J., *Electrophoresis*, 2017, vol. 38, no. 1, p. 190.

70. Han, X. and Gross, R.W., *J. Lipid Res.*, 2003, vol. 44, no. 6, p. 1071.

71. Wishart, D.S., Feunang, Y.D., Marcu, A., et al., *Nucleic Acid Res.*, 2018, vol. 46, no. D1, p. D608.

72. Aksenov, A.A., Laponogov, I., Zhang, Z., et al., *Nat. Biotechnol.*, 2021, vol. 39, p. 169.

73. Mohanty, V., Subbannayya, Y., Najar, M.A., et al., *OMICS*, 2019, vol. 23, no. 2, p. 98.

74. Peng, J., Chen, Y.-T., Chen, C.-L., et al., *Anal. Chem.*, 2014, vol. 86, no. 13, p. 6540.

75. Tsutsui, H., Maeda, T., Min, J.Z., et al., *Clin. Chim. Acta*, 2011, vol. 412, nos. 11–12, p. 861.

76. Yao, H., Shi, P., Zhang, L., et al., *Mol. BioSyst.*, 2010, vol. 6, no. 6, p. 1061.

77. Everett, J.R., Loo, R.L., and Pullen, F.S., *Ann. Clin. Biochem.*, 2013, vol. 50, no. 6, p. 523.

78. Fujimura, Y. and Miura, D., *Metabolites*, 2014, vol. 4, no. 2, p. 319.

79. Han, X. and Gross, R.W., *Mass Spectrom. Rev.*, 2005, vol. 24, no. 3, p. 367.

80. Wang, J. and Han, X., *TrAC, Trends Anal. Chem.*, 2019, vol. 121, 115697.

81. Züllig, T., Trötzmüller, M., and Köfeler, H.C., *Anal. Bioanal. Chem.*, 2020, vol. 412, no. 10, p. 2191.

82. Cajka, T. and Fiehn, O., *TrAC, Trends Anal. Chem.*, 2014, vol. 61, p. 192.

83. Triebl, A., Hartler, J., Trotzmuller, M., et al., *Biochim. Biophys. Acta*, 2017, vol. 1862, no. 8, p. 740.

84. Hsu, F.-F., *Anal. Bioanal. Chem.*, 2018, vol. 410, no. 25, p. 6387.

85. Kyle, J.E., Zhang, X., Weitz, K.K., et al., *Analyst*, 2016, vol. 141, no. 5, p. 1649.

86. Hines, K.M., Herron, J., and Xu, L., *J. Lipid Res.*, 2017, vol. 58, no. 4, p. 809.

87. Fang, M., Rustam, Y., Palmieri, M., et al., *Anal. Bioanal. Chem.*, 2020, vol. 412, no. 10, p. 2339.

88. Xiang, L., Zhu, L., Huang, Y., et al., *Small Methods*, 2020, vol. 4, no. 8, 2000160.

89. Harris, R.A., May, J.C., Stinson, C.A., et al., *Anal. Chem.*, 2018, vol. 90, no. 3, p. 1915.

90. Zhang, W., Shang, B., Ouyang, Z., et al., *Anal. Chem.*, 2020, vol. 92, no. 9, p. 6719.

91. Jurowski, K., Kochan, K., Walczak, J., et al., *Crit. Rev. Anal. Chem.*, 2017, vol. 47, no. 5, p. 418.

92. Unsihuay, D., Qiu, J., Swaroop, S., Nagornov, K.O., et al., *Int. J. Mass Spectrom. Ion Processes*, 2020, vol. 448, 116269.

93. Kerian, K.S., Jarmusch, A.K., and Cooks, R.G., *Analyst*, 2014, vol. 139, no. 11, p. 2714.

94. Rubakhin, S.S., Romanova, E.V., Nemes, P., et al., *Nat. Methods*, 2011, vol. 8, no. 4 (suppl.), p. S20.

95. Sjövall, J., *Lipids*, 2004, vol. 39, no. 8, p. 703.

96. Hoang Anh, N.H., Long, N.P., Kim, S.J., et al., *Metabolites*, 2019, vol. 9, no. 10, p. 199.

97. Wudy, S.A., Schuler, G., Sánchez-Guijoa, A., et al., *J. Steroid Biochem. Mol. Biol.*, 2018, vol. 179, p. 88.

98. Shackleton, C., Pozo, O.J., and Marcos, J., *J. Endocr. Soc.*, 2018, vol. 2, no. 8, p. 974.

99. Marcos, J. and Pozo, O.J., *Bioanalysis*, 2015, vol. 7, no. 19, p. 2515.

100. Chortis, V., Bancos, I., Nijman, Th., et al., *J. Clin. Endocrinol. Metab.*, 2020, vol. 105, no. 3, e307.

101. Liu, Q., Chi, Q., Fan, R.-T., et al., *Nat. Prod. Bioprospect.*, 2019, vol. 9, no. 3, p. 201.

102. Bai, Y.-L., Hong, Z.-D., Zhang, T.Y., et al., *J. Anal. Test.*, 2020, vol. 4, p. 1.

103. Boccard, J., Badoud, F., Grata, E., et al., *Forensic Sci. Int.*, 2011, vol. 213, nos. 1−3, p. 85.

104. Sánchez-Guijo, A., Oji, V., Hartmann, M.F., et al., *J. Lipid Res.*, 2015, vol. 56, no. 9, p. 1843.

105. Ruhaak, L.R., Xu, G., Li, Q., et al., *Chem. Rev.*, 2018, vol. 118, no. 17, p. 7886.

106. An, H.J. and Lebrilla, C.B., *Mass. Spectrom. Rev.*, 2011, vol. 30, no. 4, p. 560.

107. Harvey, D.J., *Mass Spectrom. Rev.*, 2018, vol. 37, no. 4, p. 353.

108. Banoub, J.H., Aneed, A.E., Cohen, A.M., et al., *Mass. Spectrom. Rev.*, 2010, vol. 29, no. 4, p. 606.

109. Lane, C.S., McManus, K., Widdowson, P., et al., *Anal. Chem.*, 2019, vol. 91, no. 15, p. 9916.

110. Abrahams, J.L., Taherzadeh, G., Jarvas, G., et al., *Curr. Opin. Struct. Biol.*, 2020, vol. 62, no. 1, p. 56.

111. Gemperline, E., Keller, C., and Li, L., *Anal. Chem.*, 2016, vol. 88, no. 7, p. 3422.

112. Argueso, C.T., Assmann, S.M., Birnbaum, K.D., et al., *Plant Direct*, 2019, vol. 3, no. 4, e00133.

113. Zhestovskaya, E.S., Antokhin, A.M., Taranchenko, V.F., et al., *Mass-Spektrom.*, 2019, vol. 16, no. 2, p. 139.

114. Turova, P., Rodin, I., Shpigun, O., et al., *Phytochem. Anal.*, 2020, vol. 31, no. 6, p. 948.

115. Gemperline, E., Jayaraman, D., Maeda, J., et al., *J. Am. Soc. Mass Spectrom.*, 2015, vol. 26, no. 1, p. 149.

116. Tadjimukhamedov, F.K., Huang, G., Ouyang, Z., et al., *Analyst*, 2012, vol. 137, no. 5, p. 1082.

117. Sarkar, D., Srimany, A., and Pradeep, T., *Analyst*, 2012, vol. 137, no. 19, p. 4559.

118. Malaj, N., Ouyang, Z., Sindona, G., et al., *Anal. Methods*, 2012, vol. 4, no. 7, p. 1913.

119. Kosyakov, D.S., Ul'yanovskii, N.V., Anikeenko, E.A., et al., *Rapid Commun. Mass Spectrom.*, 2016, vol. 30, no. 19, p. 2099.

120. Qi, Y., Fu, P., Li, S., et al., *Sci. Total Environ.*, 2020, vol. 713, 136573.

121. Kosyakov, D.S., Ul'yanovskii, N.V., Sorokina, E.A., et al., *J. Anal. Chem.*, 2014, vol. 69, p. 1344.

122. Xue, J., Lai, Y., Liu, C.-W., et al., *Toxics*, 2019, vol. 7, no. 3, p. 41.

123. Haddad, N., Andrianou, X.D., and Makris, K.C., *Curr. Pollut. Rep.*, 2019, vol. 5, no. 4, p. 378.

124. Lebedev, A.T., *Annu. Rev. Anal. Chem.*, 2013, vol. 6, no. 1, p. 163.

125. Xie, Z., Wang, Z., Magand, O., et al., *Sci. Total Environ.*, 2020, vol. 741, 140200.

126. Röhler, L., Bohlin-Nizzetto, P., Rostkowski, P., et al., *Atmos. Chem. Phys.*, 2021, vol. 21, no. 3, p. 1697.

127. Kosyakov, D.S., Ul'yanovskii, N.V., Latkin, T.B., et al., *Environ. Pollut.*, 2020, vol. 266, 115109.

128. Mazur, D.M., Detenchuk, E.A., Sosnova, A.A., et al., *Sci. Total Environ.*, 2021, vol. 761, 144506.

129. Grandy, J.J., Lashgari, M., Heide, H.V., et al., *Environ. Pollut.*, 2019, vol. 252, p. 825.

130. Andra, S.S., Austin, C., Patel, D., et al., *Environ. Int.*, 2017, vol. 100, p. 32.

131. Pourchet, M., Narduzzi, L., Jean, A., et al., *Talanta*, 2021, vol. 225, 121979.

132. Bloszies, C.S. and Fiehn, O., *Curr. Opin. Toxicol.*, 2018, vol. 8, p. 87.

133. Maitre, L., Robinson, O., Martinez, D., et al., *Environ. Sci. Technol.*, 2018, vol. 52, no. 22, p. 13469.

134. Walker, D.I., Valvi, D., Rothman, N., et al., *Curr. Epidemiol. Rep.*, 2019, vol. 6, no. 2, p. 93.

135. Altmaier, E., Emeny, R.T., Krumsiek, J., et al., *Psychoneuroendocrinology*, 2013, vol. 38, no. 8, p. 1299.

136. Finka, A., Sood, V., Quadroni, M., et al., *Cell Stress Chaperones*, 2015, vol. 20, no. 4, p. 605.

137. Marshall, A.G. and Rodgers, R.P., *Acc. Chem. Res.*, 2004, vol. 37, no. 1, p. 53.

138. Niyonsaba, E., Manheim, J.M., Yerabolu, R., et al., *Anal. Chem.*, 2019, vol. 91, no. 1, p. 156.

139. Borisov, R.S., Kulikova, L.N., and Zaikin, V.G., *Pet. Chem.*, 2019, vol. 59, no. 10, p. 1055.

140. Brodskii, E.S. and Shelepchikov, A.A., *J. Anal. Chem.*, 2013, vol. 68, p. 1111.

141. Gordadze, G.N., Poshibaeva, A.R., Giruts, M.V., et al., *Pet. Chem.*, 2018, vol. 58, no. 3, p. 186.

142. Giruts, M.V., Derbetova, N.B., Erdnieva, O.G., et al., *Pet. Chem.*, 2013, vol. 53, no. 5, p. 285.

143. Lozano, D.C.P., Gavard, R., Arenas-Diaz, J.P., et al., *Chem. Sci.*, 2019, vol. 10, no. 29, p. 6966.

144. Niles, S.F, Chacón-Patiño Martha, M.L., Smith, D.F., et al., *Energy Fuels*, 2020, vol. 34, no. 2, p. 1492.

145. Mapolelo, M.M., Rodgers, R.P., Blakney, G.T., et al., *Int. J. Mass Spectrom. Ion Processes*, 2011, vol. 300, nos. 2−3, p. 149.

146. McKenna, A.M., Purcell, J.M., Rodgers, R.P., et al., *Energy Fuels*, 2009, vol. 23, no. 4, p. 2122.

147. Purcell, J.M., Juyal, P., Kim, D.G., Rodgers, R.P., et al., *Energy Fuels*, 2007, vol. 21, no. 5, p. 2869.

148. Wang, X., Cai, T., Wen, W., et al., *Fuel*, 2018, vol. 230, p. 418.

149. Islam, A., Cho, Y., Ahmed, A., et al., *Mass Spectrom. Lett.*, 2012, vol. 3, no. 3, p. 63.

150. Altuntas, E. and Schubert, U.S., *Anal. Chim. Acta*, 2014, vol. 808, p. 56.

151. Zaikin, V.G., Mardanov, R.G., and Plate, N.A., *Polym. Sci., Ser. A*, 1993, vol. 35, no. 9, p. 1305.

152. Wesdemiotis, C., Solak, N., Polce, M.J., et al., *Mass Spectrom. Rev.*, 2011, vol. 30, no. 4, p. 523.

153. Yol, A.M., Janoski, J., Quirk, R.P., et al., *Anal. Chem.*, 2014, vol. 86, no. 19, p. 9576.

154. Town, J.S., Jones, G.R., Hancox, E., et al., *Macromol. Rapid Commun.*, 2019, vol. 40, no. 13, 1900088.

155. Borisov, R.S. and Zaikin, V.G., *J. Anal. Chem.*, 2010, vol. 65, p. 1423.

156. Borisov, R.S., Polovkov, N.Y., and Zaikin, V.G., *Rapid Commun. Mass Spectrom.*, 2009, vol. 23, no. 20, p. 3309.

157. Cifuentes, A., *J. Chromatogr. A*, 2009, vol. 1216, no. 43, p. 7109.

158. Herrero, M., Simó, C., García-Cañas, V., et al., *Mass Spectrom. Rev.*, 2012, vol. 31, no. 1, p. 49.

159. Morrell, P.L., Buckler, E.S., and Ross-Ibarra, J., *Nat. Rev. Genet.*, 2011, vol. 13, no. 2, p. 85.

160. Kussmann, M., Affolter, M., Nagy, K., et al., *Mass Spectrom. Rev.*, 2007, vol. 26, no. 6, p. 727.

161. Barnett, M., Young, W., Cooney, J., et al., *J. Nutrigenet. Nutrigenomics*, 2015, vol. 7, nos. 4–6, p. 274.

162. Wittenbecher, C., Mühlenbruch, K., Kröger, J., et al., *Am. J. Clin. Nutr.*, 2015, vol. 10, no. 6, p. 1241.

163. Hoffmann-Sommergruber, K., *EuPa Open Proteomics*, 2016, vol. 12, p. 10.

164. Di Girolamo, F., Muraca, M., Mazzina, O., et al., *Curr. Opin. Allergy Clin. Immunol.*, 2015, vol. 15, no. 3, p. 259.

165. García-Cañas, V., Simó, C., León, C., et al., *Mass Spectrom. Rev.*, 2010, vol. 30, no. 3, p. 396.

166. Gil-Solsona, R., Raro, M., Sales, C., et al., *Food Control*, 2016, vol. 70, p. 350.

167. Piccolo, A., in *Humic Substances in Terrestrial Ecosystems*, Amsterdam: Elsevier, 1996, p. 225.

168. Nebbioso, A. and Piccolo, A., *Anal. Bioanal. Chem.*, 2012, vol. 405, no. 1, p. 109.

169. Drosos, M., Nebbioso, A., and Piccolo, A., *Appl. Soil Ecol.*, 2018, vol. 123, p. 513.

170. Zhang, X., Chen, Z., Huo, X., et al., *Sci. Total Environ.*, 2021, vol. 756, p. 144140.

171. Herzsprung, P., Hertkorn, N., von Tumpling, W., et al., *Anal. Bioanal. Chem.*, 2016, vol. 408, no. 10, p. 2461.

172. Zherebker, A., Kim, S., Schmitt-Kopplin, P., et al., *Pure Appl. Chem.*, 2020, vol. 92, no. 9, p. 1447.

173. Kulikova, N.A., Volikov, A.B., Filippova, O.I., et al., *Agronomy*, 2021, vol. 11, no. 1, 150.

174. Zhernov, Y.V., Konstantinov, A.I., Zherebker, A., et al., *Environ. Res.*, 2021, vol. 193, 110312.

175. Lieke, T., Steinberg, C.E.W., Pan, B., et al., *Sci. Rep.*, 2021, vol. 11, no. 1, 174.

Translated by O. Zhukova