F O U N D A T I O N S O F C O M P U T I N G A N D D E C I S I O N S C I E N C E S
Vol. 50 (2025) No. 2
ISSN 0867-6356
DOI: 10.2478/fcds-2025-0007 e-ISSN 2300-3405

sciendo

Algorithm for Determining the Sum Formula of Metabolites from Mass Spectrometry Spectra

Marcin Borowski *, Karolina Baumgard

Abstract. This work addresses the challenge of determining chemical sum formulas from mass spectrometry data through a three-fold approach: establishing a formal problem framework, analyzing its computational complexity, and developing an approximate algorithm that synergizes reference database integration with functional group property analysis to enable accurate compound identification and structural characterization. By bridging theoretical foundations in computational complexity with practical solutions for chemical structure elucidation, the proposed methodology advances analytical capabilities in mass spectrometry data interpretation.

Keywords: scip, tabu search, mass spectrometry, spectra identification

1. Introduction

Recent developments in mass spectrometry have significantly increased the volume of data generated during the analysis of metabolites and lipids. Metabolomics and lipidomics, which focus on the characterization of small-molecule compounds in biological systems, now face substantial challenges in processing and interpreting these complex datasets. Computational approaches, including optimization and machine learning techniques, have been successfully applied in related fields, such as protein structure prediction [3], demonstrating their potential for handling high-dimensional biological data. In this study, we present an advanced algorithm designed to facilitate the analysis of mass spectra associated with metabolomic and lipidomic investigations.

A key approach to characterizing an unknown compound involves determining its sum formula, that is, its elemental composition and the number of atoms of each element present. Given a measured mass, a set of candidate sum formulas can be generated – a problem analogous to sequence identification in proteomics [6]. The

*Institute of Computing Science, Poznan University of Technology, Poznań, marcin.borowski@put.poznan.pl

size of this set depends on the elements considered and the precision of the mass measurement. Although increased mass accuracy can reduce the number of possible formulas, it does not necessarily lead to a unique solution. This limitation underscores the importance of leveraging additional information encoded in the mass spectrum, similar to how structural constraints refine predictions in protein analysis [12]. Consequently, the determination of the sum formula represents a critical initial step in the *de novo* identification of metabolites.

2. Related work

Accurate determination of chemical sum formulas from mass spectrometry data is a longstanding challenge in analytical chemistry and metabolomics. Numerous computational strategies have been developed to address the ambiguity inherent in formula assignment, especially as the molecular mass increases and the number of plausible formulas grows rapidly.

Early approaches relied primarily on high-precision mass measurements and elemental constraints to generate candidate formulas. However, even with sub-ppm mass accuracy, the number of chemically plausible formulas can be substantial, particularly for larger molecules [17, 11]. To address this, additional information such as isotopic patterns and heuristic chemical rules are routinely incorporated into formula prediction workflows.

A prominent example is the Rdisop package [17], which determines sum formulas using both exact mass and isotope patterns, leveraging algorithms developed by Böcker et al. (2006, 2008, 2009). Rdisop integrates with Bioconductor tools and applies rules such as the nitrogen rule and double bond equivalents (DBE) to filter candidates. Similarly, MZmine 2 offers a universal formula prediction tool that combines heuristic rules, MS/MS fragmentation analysis, and isotope pattern matching to improve accuracy. In evaluations with Orbitrap MS data, this approach correctly ranked the true formula first for 79% of the compounds tested, outperforming earlier methods in many cases [19].

Recent advances have introduced machine learning and deep learning to further refine formula prediction. For example, FIDDLE is a deep learning framework that predicts chemical formulas directly from tandem mass spectra, dramatically reducing the candidate formula space and improving classification accuracy compared to traditional computational methods and other tools such as SIRIUS and BUDDY [9]. These methods benefit from the ability to learn complex MS/MS feature relationships and to leverage large spectral libraries.

Exact algorithms, such as integer linear programming (ILP), have also been applied to related problems such as de novo peptide identification [7]. In the context of sum formula identification, constraint-based solvers (e.g., SCIP) and metaheuristic approaches (e.g., tabu search, genetic algorithms) have been explored for their ability to efficiently search large solution spaces while adhering to chemical plausibility constraints.

Despite these advances, limitations remain. The high mass precision alone is in-

sufficient for the determination of unique formulas, especially for compounds of high mass [17, 11]. Heuristic and machine learning approaches can improve ranking and reduce candidate lists, but often require extensive training data and careful validation. The integration of multiple orthogonal data types (e.g. isotope patterns, MS/MS fragmentation, retention time) is increasingly recognized as essential for robust formula assignment.

Our work builds on these foundations by formally framing the sum formula identification problem as a constrained optimization task and developing both exact (SCIP) and heuristic (tabu search) algorithms. We benchmark these methods against real metabolomic data, demonstrating their strengths and limitations relative to existing approaches.

Building on these prior works, we now formalize the sum formula identification problem and present our computational approach.

3. Mass spectrometry

Mass spectrometry is an analytical technique used in fields such as chemistry, biochemistry, medicine, or pharmacy [8]. It involves the separation of ions according to their mass-to-charge ratio (m/z), enabling an accurate measurement of atomic masses and chemical compounds. It is versatile and can detect the presence of specific elements or compounds in samples.

Due to its precision, mass spectrometry also helps to determine the structure of chemical compounds, such as the structure of proteins [21, 22]. Although samples tested by this method are destroyed, the technique is irreplaceable due to its accuracy and wide applicability.

Modern mass spectrometers analyze complex samples, making this technique crucial in scientific research. In the following chapters, we discuss its application in data analysis and advanced algorithms to aid in the interpretation of the results [8].

3.1. Construction

A mass spectrometer allows the statistical distribution of the atomic masses of elements and compounds in a sample to be obtained. It consists of an ionizer, an analyzer, and a detector, operating under high vacuum conditions to avoid collisions with air molecules, which would interfere with the measurements of [8].

The ionizer is the part of the spectrometer where the ionization of molecules takes place. Ionization methods are divided into those that cause a chemical compound to fragment into smaller fragments and those that give the molecules a charge without fragmenting them. Ionization can be positive or negative. In positive ionization, protons from dissociated acids [21] are most frequently present.

The analyzer splits the ion beam according to the mass-to-charge ratio (m/z), and the detector records the ions, allowing qualitative and quantitative analysis. Data from the detector are visualized as a mass spectrum, showing the signal intensity

as a function of m/z. The mass spectrum illustrates the characteristic peaks of the different ions, crucial for the identification of the substances under investigation [22].

Figure 1 shows example mass spectra of a metabolite from Axyris amaranthoides, illustrating differences between positive and negative ionization modes.

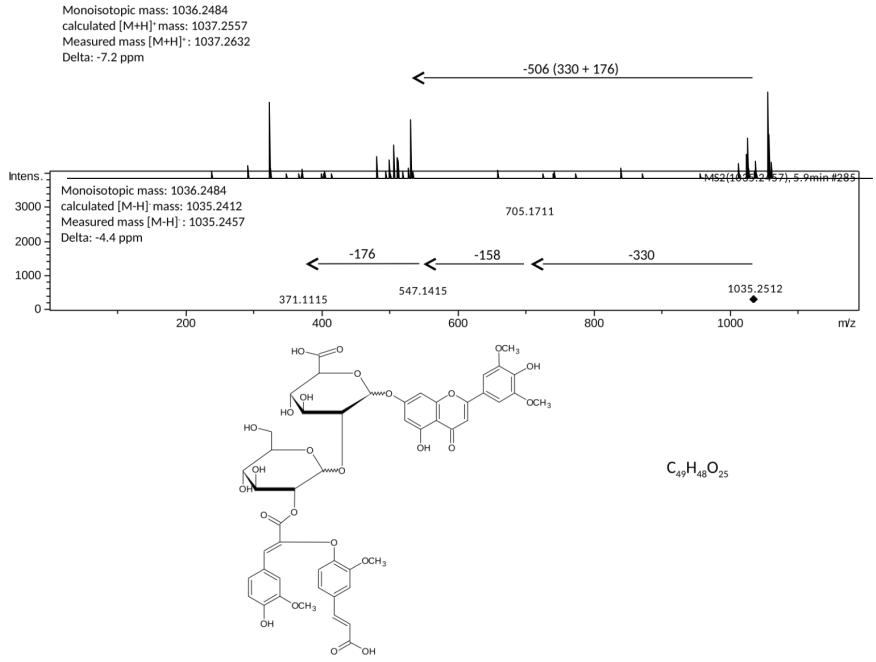

Figure 1. Example spectrum of a metabolite from the *Axyris amaranthoides* plant, taken in positive ionization mode (top spectrum) and negative ionization mode (bottom spectrum) of *MZAB16*.

Due to their high precision, modern mass spectrometers analyze complex samples, making this technique indispensable in scientific research. In the following chapters, we discuss the analysis and interpretation of spectrometric data and advanced algorithms that support these processes.

3.2. Mass spectra

A mass spectrum is a graphical representation of the dependence of the signal intensity on the mass-to-charge ratio (m/z), where mass is the mass number and charge is a multiple of the elemental charge of the ion [15]. Fragments of a certain mass form peaks in the spectrum, allowing the mass of individual fragments to be read off. For singly charged fragments, the value of the spectrum corresponds to the mass of the fragment.

The process of decomposition of a chemical compound into fragments recorded by the spectrometer is divided into several steps [13]. Figure 2 presents the proposed fragmentation steps based on the negative ionization spectrum, highlighting the main fragment ions detected.

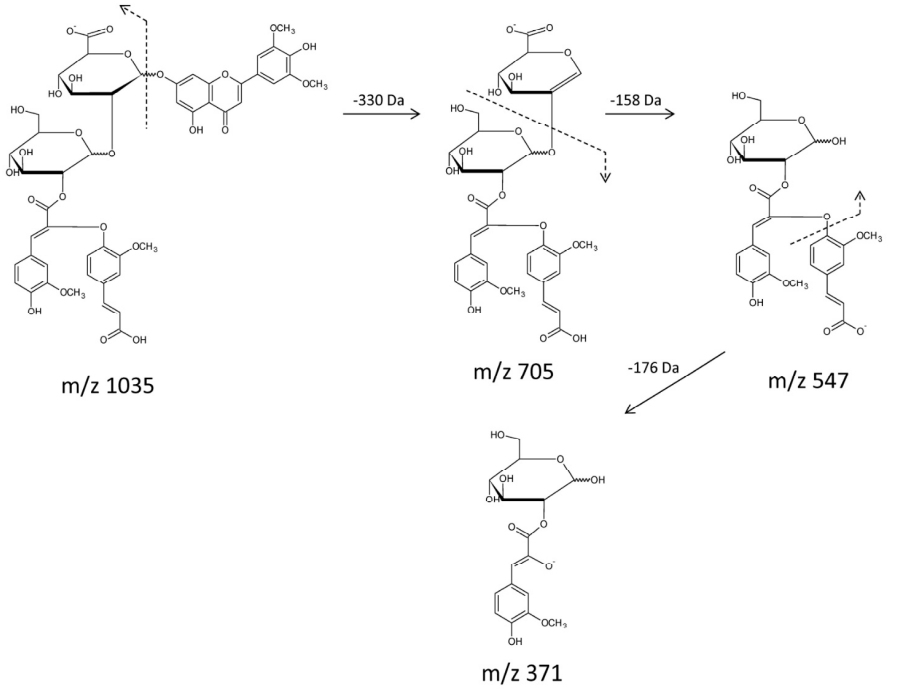

Figure 2. The proposed fragmentation steps based on the spectrum in negative ionization mode of the metabolite shown in Figure 1 which are described in [13].

Each fragment can be described by a summary formula which can be determined from its mass from the spectrum (Figure 1). Figure 3 shows the sum formulas of the fragments in Figure 2.

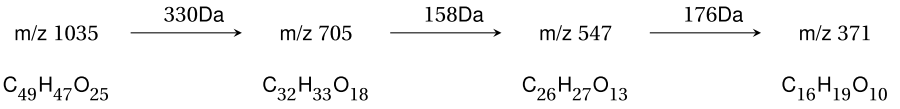

Figure 3. The summation patterns of the fragments shown in Figure 2 which are described in [13].

4. Definition of the problem

Database studies of [10] show that it is not possible to unambiguously identify the sum formula of a metabolite based on the error in its mass measurement. In this work, this problem will be extended to the identification of all possible sum formulas of a metabolite. Examples of methods for the identification of the sum formula are available in the literature [4, 5, 18, 16]. The problem of identifying potential sum formulas will be formally presented as Problem 1.

Problem 1. *Let the data be a set P of ordered pairs (m_i, w_i), where $m_i \in \mathbb{Q}^+$ and $w_i \in \mathbb{Z}^+$, $i = 1, 2, \ldots, k$, number $M \in \mathbb{Q}^+$ and precision $\delta \in \mathbb{Q}^+$. Find the set of vectors $x = (x_1, x_2, \ldots, x_i, \ldots, x_k)$ of integer nonnegative values, where each vector in this set satisfies the assumption:*

$$M - \delta \leq \sum_{i=1}^{k} m_i x_i \leq M + \delta, \tag{1}$$

$$\sum_{i=1}^{k} w_i x_i = 2n, \quad n \in \mathbb{N}, \tag{2}$$

$$\sum_{i=1}^{k} (w_i - 2) x_i + 2 \geq 0, \tag{3}$$

$$\sum_{i=1}^{k} w_i x_i - 2 \max\{w_i \mid x_i > 0\} \geq 0, \tag{4}$$

In the definition of the problem 1, the set P consists of pairs where m_i is the mass of the i-th element and w_i is its valence, determining the number of chemical bonds of an element. Elements can have different valences depending on the reaction conditions, e.g. nitrogen and phosphorus can occur as multiple pairs with the same m_i, but different w_i.

The vector x determines the number of pairs occurrences in a chemical compound of mass M and meets the conditions (1) - (4). The condition (5) ensures that the mass of the selected elements is within the range of $M \pm \delta$. Conditions (1) - (4) concern the structure of a compound that is not an ion. Chemical compounds must form a whole, and atoms must form all possible bonds.

Condition (2) ensures the uniformity of the sum of valencies, preventing the presence of unpaired electrons and the formation of ions. The condition (3) requires the structure to be completely connected. Condition (4) requires that the sum of the valence of all elements, except the largest, be sufficient to form all bonds.

For example, the summary formula C_2H_5, C_2H_8 and PH_3, where phosphorus is pentavalent, does not adequately meet the assumptions of (2) - (4).

A vector belongs to the set of solutions if it satisfies all the assumptions. Condition (1) ensures that the mass calculated from the vector is close to the given mass. For

a large mass of a chemical compound, all vectors satisfy condition (4), which follows from the inequality.

$$(\max_i\{w_i\} + 1) \cdot \max_i\{m_i\} < M - \delta, \tag{5}$$

The maximum value in a set of elements is denoted by $\max_i\{w_i\}$, and the maximum mass is denoted by $\max_i\{m_i\}$.

Vector x can be transformed into a sum formula by adding up the number of pairs occurrences with the same masses but different values.

5. Algorithms

The algorithms receive as input a set P representing the acceptable elements that can be present in the compounds sought, the mass M of the metabolite concerned and the precision δ. The error, which is understood as the difference between the mass of the metabolite and the mass of the selected solution, is minimized. An admissible solution is a vector of element counts.

5.1. SCIP Solver

SCIP (Solving Constraint Integer Programs) [1] is a programming platform for integer programming with constraints [2].

SCIP was chosen because of its ability to exhaustively enumerate all solutions that satisfy complex chemical constraints, providing a comprehensive reference for comparison with heuristic methods. However, its computational cost increases rapidly with the number of elements and the size of the molecule.

The problem 1 can be formulated as a mathematical programming task and solved using a solver. Since the experimental data analyzed in the next section satisfy the condition (4), it is not included in the formulation of the problem 1 as a linear integer programming task. A problem without this constraint can be represented as an error minimization, which is understood as the difference between the mass of the chemical compound under study and the mass of the compound with a given integer formula. The problem can be written as the following integer linear programming task:

Problem 2. *For a given set P of ordered pairs (m_i, w_i), where $m_i \in \mathbb{Q}^+$ and $w_i \in \mathbb{Z}^+$, $i = 1, 2, \ldots, k$, number $M \in \mathbb{Q}^+$ and the precision $\delta \in \mathbb{Q}^+$ d assuming that the restrictions are met:*

$$\sum_{i=1}^{n} w_i x_i - 2n = 0 \tag{6}$$

$$\sum_{i=1}^{n} w_i x_i - \sum_{i=1}^{n} x_i + 2 \geq 0 \tag{7}$$

$$\sum_{i=1}^{n} m_i x_i + d \geq M \tag{8}$$

$$\sum_{i=1}^{n} m_i x_i - d \leq M \tag{9}$$

$$d \leq \delta \tag{10}$$

Where $x1, x2, ..., xn$ are numbers with integer nonnegative values.

The SCIP also makes it possible to obtain, in addition to the optimal solution, all admissible solutions satisfying the given conditions.

5.2. Tabu Search

A tabu search algorithm has been designed to solve the problem 1. Its main goal is to avoid local minima by remembering and avoiding previous moves (forbidden moves - tabu). Here are the most salient features of this algorithm for the problem 1:

Tabu search was selected as a practical alternative to exhaustive search, capable of efficiently exploring large solution spaces and identifying most valid candidates within reasonable computation times. The trade-off is that some solutions may be missed compared to the exact enumeration.

- **Input:** Set of ordered pairs (m_i, w_i), number M, accuracy δ, calculation time t
- **Output:** A set of pairs $(x, |\sum_i m_i w_i|$ meeting the conditions of the Problem 1

The process of finding a solution is initiated by preparing a vector of elementary movements. These movements involve substitutions that make minimal changes to the total mass while maintaining the total mass value. The procedure ensures that the total mass is consistent with the search. The resulting interval includes all suitable solutions for a given mass.

The procedure calculates additional properties, including density, to characterize the solutions obtained. This allows for the determination of the minimum and maximum bond number values for a given mass. The tabu search algorithm is run for all bond numbers within the specified interval, keeping the total mass number. The tabu list stores elemental moves and directions of change to prevent elemental substitution in an undesirable way.

Before determining the next neighborhood solution, the feasibility of the elementary move is checked. The function checks that the subtraction of the appropriate number of primes will not result in a negative number of these primes in the solution and thus the satisfiability of the constraints of problem (1) - (4).

The tabu search algorithm runs for a set amount of time. Browse through the solution space and save those that meet all the conditions. The functions of the algorithm check the conditions of the problem accordingly and finally write the solutions that meet the conditions into the resulting file.

Next, we evaluate these algorithms using real spectrometric data from plant metabolites.

6. Research

The actual spectrometric data of selected metabolites of the plant Axyris amaranthoides [13] (Table 1 lists the monoisotopic masses of the 29 metabolites analyzed in this study) were studied in a computational experiment.

lp.	summary formula	monoisotopic mass [Da]	Lp.	summary formula	monoisotopic mass [Da]
1	$C_{22}H_{20}O_{12}$	476.0954	16	$C_{38}H_{36}O_{22}$	844.1698
2	$C_{22}H_{20}O_{13}$	492.0904	17	$C_{39}H_{40}O_{21}$	844.2062
3	$C_{23}H_{22}O_{13}$	506.1050	18	$C_{39}H_{38}O_{22}$	858.1854
4	$C_{28}H_{28}O_{18}$	652.1275	19	$C_{40}H_{40}O_{22}$	872.2011
5	$C_{28}H_{28}O_{19}$	668.1224	20	$C_{39}H_{38}O_{23}$	874.1803
6	$C_{29}H_{32}O_{18}$	668.1588	21	$C_{40}H_{40}O_{23}$	888.1960
7	$C_{29}H_{30}O_{19}$	682.1381	22	$C_{42}H_{42}O_{24}$	930.2066
8	$C_{30}H_{32}O_{19}$	696.1537	23	$C_{43}H_{44}O_{25}$	960.2171
9	$C_{34}H_{34}O_{21}$	778.1592	24	$C_{43}H_{44}O_{26}$	976.2120
10	$C_{36}H_{34}O_{20}$	786.1643	25	$C_{44}H_{46}O_{26}$	990.2277
11	$C_{37}H_{34}O_{20}$	798.1643	26	$C_{45}H_{50}O_{26}$	1006.2590
12	$C_{37}H_{34}O_{21}$	814.1592	27	$C_{45}H_{48}O_{27}$	1020.2383
13	$C_{38}H_{38}O_{20}$	814.1956	28	$C_{49}H_{48}O_{25}$	1036.2484
14	$C_{38}H_{36}O_{21}$	828.1749	29	$C_{49}H_{48}O_{26}$	1052.2433
15	$C_{39}H_{38}O_{21}$	842.1905			

Table 1. Data set analyzed. Actual monoisotopic masses expressed in Daltons of 29 compounds, measured with a mass spectrometer.

The set of available elements in the compounds analyzed was limited to those shown in Table 2 - this table details the elemental composition, monoisotopic masses and valences used in our computational analysis.

The experiment was carried out on a PC with an Intel Core i5-4590 3.30 GHz Haswell processor, 16 GB of DDR3 1600 MHz RAM, and an SSD, running the Linux operating system.

Figure 4 shows the dependence of the number of solutions found as a function of the monoisotopic mass (expressed in Daltons) of the selected metabolites. The calculations were performed multiple times for the same data set, but with different sets of elements in the instance. The sets of elements considered were defined above the respective graphs. In creating these sets, only the most common elemental valences in organic chemistry were considered, with valence II always assumed for sulfur (for simplicity).

Figure 5 shows an analogous relationship for the results obtained in the tab al-

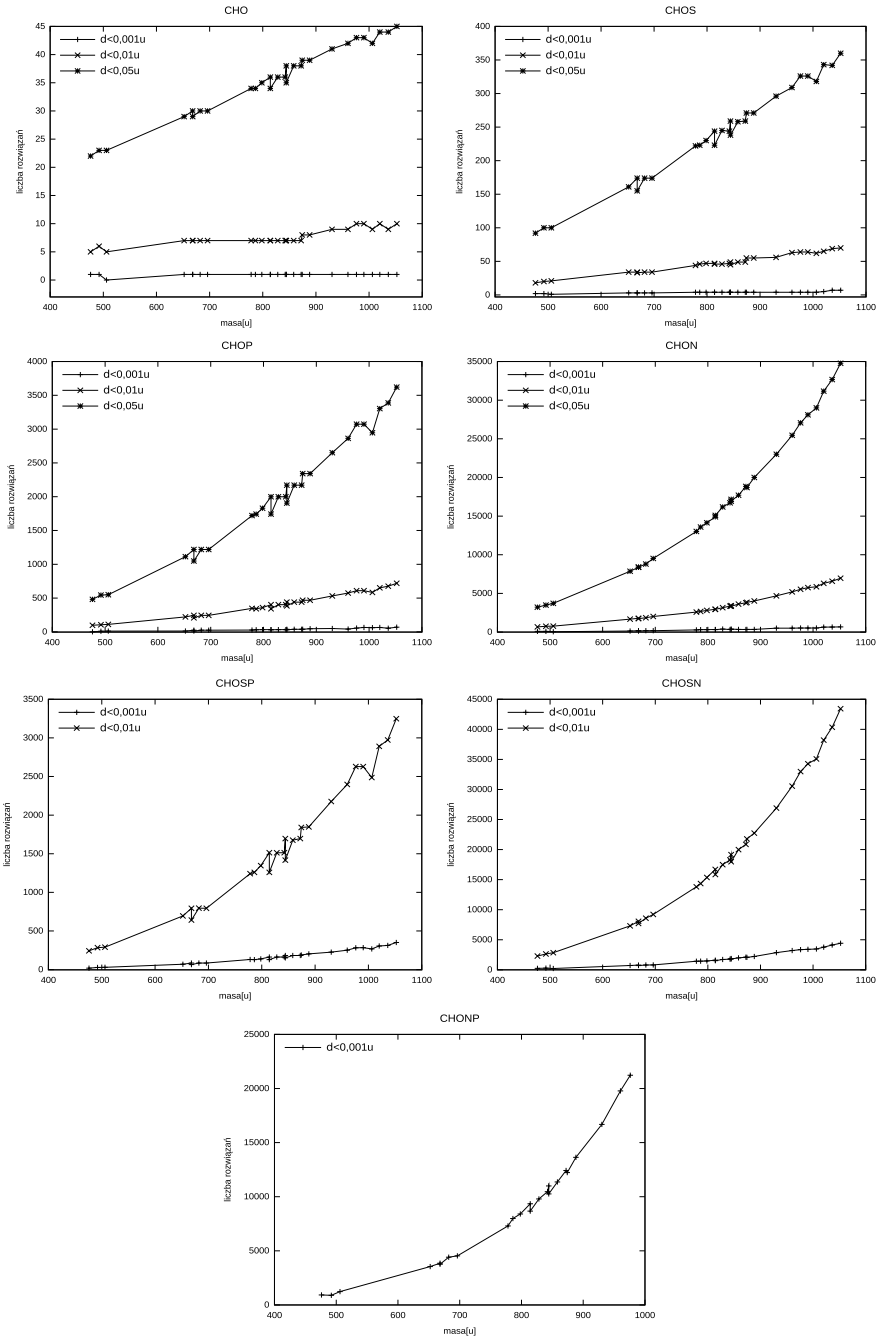

Figure 4. Results obtained from the SCIP analysis.

lp.	name	symbol	monoisotopic mass [Da]	valence
1	hydrogen	H	1.0078250	I
2	carbon	C	12.0000000	I, II, III, IV
3	nitrogen	N	14.0030740	I, II, III, IV, V
4	oxygen	O	15.9949146	I, II
5	phosphorus	P	30.9737620	I, II, III, IV, V
6	sulphur	S	31.9720712	I, II, III, IV, V, VI

Table 2. Selected elements were considered in data analysis with their monoisotopic masses and valences.

gorithm experiment - the number of solutions found depending on the monoisotopic mass (expressed in Daltons) of the selected metabolites. The calculations were repeated for the same data set but with different sets of elements in the instance. The graphs are information about the considered set of elements for the instance.

Figure 6 shows graphs showing the dependence of the percentage of missing solutions in the results obtained with the tabu search algorithm compared to the exact results obtained with the SCIP solver.

The above analysis demonstrates that the constraints formulated in the problem 1 allow the determination of a large number of potential sum formulas for the single metabolite under investigation. In addition, the results show that the number of these formulas is very large, highlighting the complexity of the problem and the potential diversity of solutions.

7. Discussion

Our results confirm that the challenge of identifying the sum formula persists even with high mass accuracy, as previously reported [17, 19, 11]. The SCIP-based exact algorithm provides comprehensive candidate sets, but computational demands increase rapidly with molecular size and the number of allowed elements. The tabu search heuristic offers a practical alternative, efficiently exploring the solution space and recovering most valid candidates, though some may be missed compared to exhaustive search. Compared to established tools such as Rdisop and MZmine 2, our methods are distinguished by their formal mathematical framework and flexibility in accommodating custom chemical constraints. Although deep learning approaches such as FIDDLE show promise in further narrowing candidate lists and ranking solutions, they require extensive spectral libraries and may not generalize to all compound classes.

Our findings highlight the importance of integrating multiple sources of information-accurate mass, isotopic patterns, fragmentation data, and chemical rules-to improve formula assignment. Future work could explore hybrid strategies that combine the strengths of constraint-based optimization, heuristic search, and machine learning, as

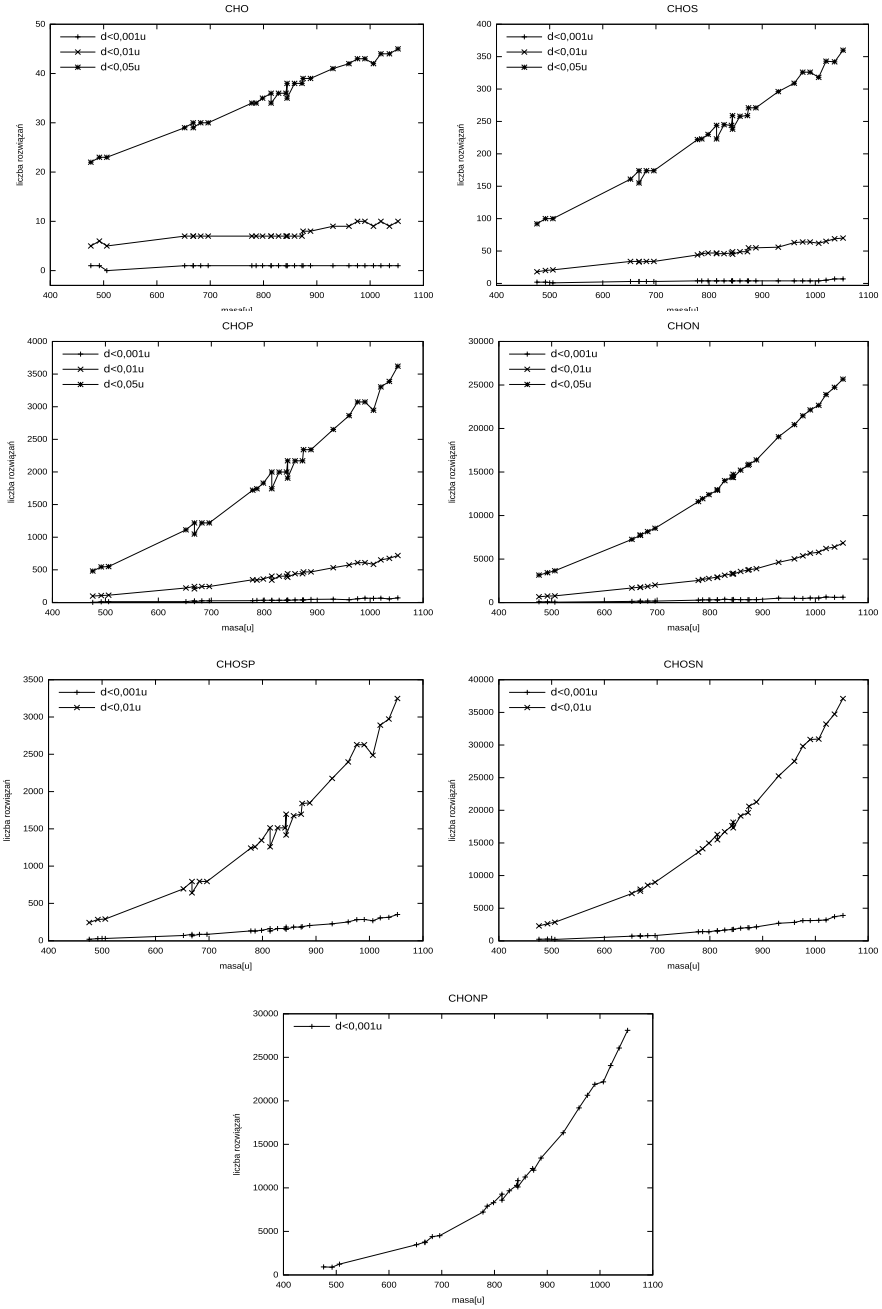

Figure 5. Results obtained from the analysis of the Tabu search algorithm.

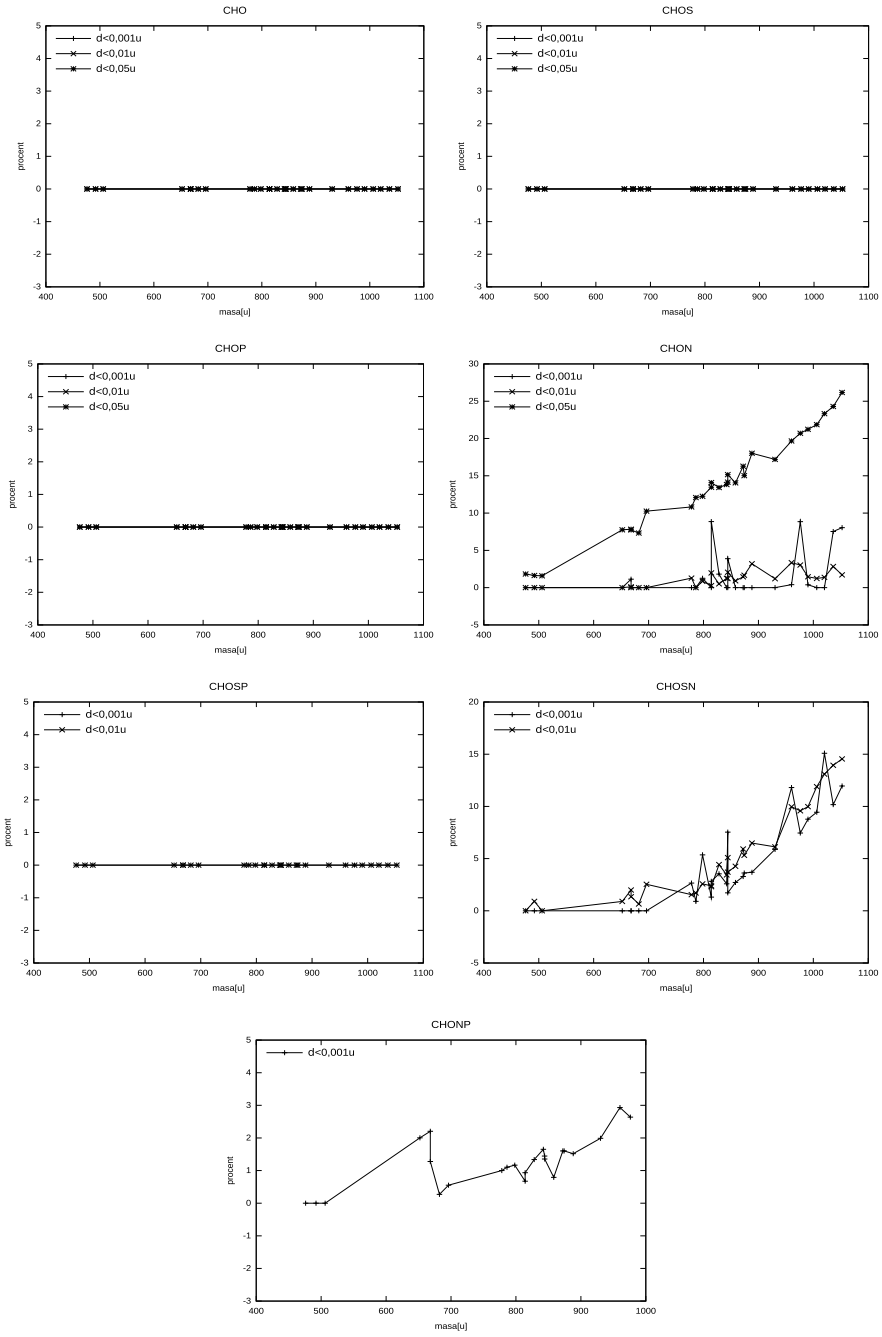

Figure 6. Results of a comparison of the results obtained from testing the SCIP and Tabu algorithms.

well as the inclusion of retention time and other orthogonal data.

8. Conclusions

In this study, we addressed the NP-hard problem of sum formula identification from mass spectrometry data by developing and benchmarking both exact (SCIP) and heuristic (tabu search) algorithms. Our approach formally encodes chemical plausibility constraints and demonstrates robust performance on real metabolomic datasets.

The analysis demonstrated that measurement error alone is insufficient to unambiguously identify a unique sum formula. The large number of possible formulas corresponding to a given mass highlights the need for additional constraints or complementary analyzes to narrow down the candidate set, a challenge similar to that observed in clinical lipidomics research, where integrated analytical approaches have proven to be most effective [14].

Compared to existing methods, our algorithms offer a transparent and flexible framework for candidate generation and filtering. However, as with previous work, we find that unique formula assignment remains challenging for larger molecules, underscoring the need for multimodal data integration and further methodological innovation. Future research should focus on:

- Incorporating machine learning models, such as deep neural networks, to improve candidate ranking and reduce false positives.
- Integrating additional experimental data types (e.g., retention time, MS/MS fragmentation) for more confident formula assignment.
- Extending the approach to broader classes of compounds and larger datasets.

Our work contributes to the ongoing development of computational tools for metabolomics and analytical chemistry, supporting more accurate and efficient compound identification from complex mass spectrometry data.

Acknowledgment

The author would like to express sincere gratitude to Prof. dr hab. inż. Piotr Formanowicz for his invaluable guidance, insightful suggestions, and constructive feedback during the initial stages of this article's preparation. His expertise and support were essential in shaping the direction of the research.

References

[1] Achterberg T., Scip: Solving constraint integer programs. Mathematical Programming Computation, 1(1):1–41, July 2009.

[2] Achterberg T., Berthold T., Koch T., Wolter K., Constraint integer programming: A new approach to integrate CP and MIP. Lecture Notes in Computer Science, 5015:6–20, 2008.

[3] Blazewicz J., Hammer P.L., Lukasiak P., Predicting secondary structures of proteins, IEEE Engineering in Medicine and Biology, 24, pp.88-94, 2005.

[4] Bocker S., Liptak Z.. A Fast and Simple Algorithm for the Money Changing Problem. Algorithmica, 48(4):413–432, 2007.

[5] Bocker S., Liptak Z., Martin M., Pervukhin A., Sudek H.. DECOMP–from interpreting Mass Spectrometry peaks to solving the Money Changing Problem. Bioinformatics, 24(4):591–593, Feb 2008.

[6] Borowski M., 2011, Models and algorithms for peptide sequence identification, Wydawnictwo NAKOM, Poznan Monographs in Computing and Its Applications (Edition 1, Volume 14), 2011.

[7] DiMaggio P.A., Floudas C.A., De Novo Peptide Identification via Tandem Mass Spectrometry and Integer Linear Optimization, Analytical Chemistry, 2007

[8] Gross J.H.. Mass Spectrometry. Springer Berlin Heidelberg, 2011.

[9] Hong Y., Li S., Ye Y., Tang H., FIDDLE: a deep learning method for chemical formulas prediction from tandem mass spectra, bioRxiv, 2024

[10] Kind T., Fiehn O.: Metabolomic database annotations via query of elemental compositions. Mass accuracy is insufficient even at less than 1 ppm. BMC Bioinformatics, 7(1), 2006, 1–10.

[11] Kuehl D., Wang Y., The role of spectral accuracy in mass spectrometry, Spectroscopy Supplements Special Issues 04/01, 2007.

[12] Lukasiak P., Blazewicz J., Milostan M., Some operations research methods for analyzing protein sequences and structures, Annals of Operations Research, 175, pp. 9-35, 2010.

[13] Marczak Ł., Znajdek-Awiżeń P., Bylka W. The use of mass spectrometric techniques to differentiate isobaric and isomeric flavonoid conjugates from axyris amaranthoides. Molecules, 2016, 21(9):1229.

[14] Marczak L., Idkowiak J., Tracz J., Stobiecki M., Perek B., Kostka-Jeziorny K., Tykarski A., Wanic-Kossowska M., Borowski M., Osuch M., Formanowicz D. and Luczak M., Mass Spectrometry-Based Lipidomics Reveals Differential Changes in the Accumulated Lipid Classes in Chronic Kidney Disease, Metabolites, 11(5):275, April 2021, DOI: 10.3390/metabo11050275

[15] McNaught A. D., Wilkinson A. IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Blackwell Scientific Publications, Oxford, 1997.

[16] Meringer M., Reinker S., Zhang J., Muller A.. MS/MS Data Improves Automated Determination Molecular Formulas by Mass Spectrometry. MATCH Communications in Mathematical and Computer Chemistry, 65(2):259–290, June 2011.

[17] Neumann, Steen, Anton Pervukhin and Sebastian Böcker, Mass decomposition with the Rdisop package, Chemistry, Physics, 2006, 2008, 2009.

[18] Patiny L., Borel A.. ChemCalc: a building block for tomorrow's chemical infrastructure. J Chem Inf Model, 53(5):1223–1228, May 2013.

[19] Pluskal t., Uehara T., Yanagida M., Highly Accurate Chemical Formula Prediction Tool Utilizing High-Resolution Mass Spectra, MS/MS Fragmentation, Heuristic Rules, and Isotope Pattern Matching, Analytical Chemistry, 2012

[20] Remes et al., The Role of Spectral Accuracy in Mass Spectrometry, Chromatography Online, 2007.

[21] Silberring J., Kraj A., Drabik A. Proteomika i metabolomika. Wydawnictwo Uniwersytetu Warszawskiego, Warszawa, 2010.

[22] Silberring J., Kraj A. Proteomika. Wydział Chemii Uniwersytetu Jagiellońskiego, Kraków, 2004.

Received 06.05.2025, Accepted 26.05.2025