

Neue Methoden zur Analyse von Fragmentierungsmassenspektren kleiner Moleküle¹

Franziska Hufsky²

Abstract: Die Identifikation von kleinen Molekülen im Hochdurchsatz spielt eine wichtige Rolle in vielen Forschungsbereichen. Massenspektrometrie (MS), insbesondere Fragmentierungsmassenspektrometrie, ist eine der am weitesten verbreiteten Techniken für deren Analyse. Die dadurch entstehende Menge an Daten ist jedoch kaum manuell zu analysieren. Wir stellen neue *de novo* Methoden zur automatisierten Analyse dieser Daten vor, die keine chemischen Kenntnisse über das zu untersuchende Molekül erfordern. Die zugrunde liegenden NP-schweren Probleme lösen wir mittels exakter aber trotzdem effizienter Algorithmen. Unsere Methoden ermöglichen die automatisierte Analyse von bisher vollkommen unbekannten Molekülen.

1 Einleitung

Kleine Moleküle, die als Zwischen- und Endprodukte aller chemischen Reaktionen innerhalb der Zellen lebender Organismen entstehen, werden als Metaboliten bezeichnet. Metaboliten umfassen eine breite Palette von Verbindungsklassen, unter anderem Zucker, Säuren, Basen, Lipide, hormonelle Steroide und viele andere [Is09]. Daher sind sie trotz ihrer geringen Größe in ihrer Struktur besonders vielfältig [Fe04]. Die biologischen und medizinischen Fragestellungen, die durch die Identifizierung dieser kleinen Moleküle adressiert werden, sind vielseitig. In der Diagnostik und Forensik wird auf kleine Moleküle, z.B. Drogenabbauprodukte, getestet [Wi07]; in der funktionellen Genomik hilft das Wissen über die Funktion von Metaboliten, um Rückschlüsse auf die Genfunktion zu ziehen [Fe04]; und in der Pharmazie basieren viele Medikamente auf kleinen Naturstoffen, unter anderem Antibiotika (z.B. Penicillin), Antiparasitika (z.B. Avermectine), Antimalariamittel (z.B. Chinin) oder Anti-Krebs-Medikamente (z.B. Taxol) [LV09]. Des Weiteren spielt die Aufklärung des Stoffwechsels verschiedener (neu entdeckter) Lebewesen eine große Rolle für viele weitere Forschungsgebiete, unter anderem der Klimaforschung.

Aufgrund der strukturellen Vielfältigkeit dieser kleinen Moleküle ist es derzeit nicht möglich, das gesamte Spektrum an Metaboliten mittels einer einzigen instrumentellen Analysemethode zu untersuchen. Stattdessen ergänzen sich verschiedene Technologien [Di11], insbesondere Kernspinresonanzspektroskopie (englisch: *nuclear magnetic resonance*, NMR) und Massenspektrometrie (MS). NMR-Spektroskopie ermöglicht die vollständige Strukturaufklärung unbekannter Moleküle, jedoch werden für dieses Verfahren große Mengen an Probensubstanz benötigt. MS ist hingegen wesentlich sensitiver

¹ Englischer Titel der Dissertation: "Novel Methods for the Analysis of Small Molecule Fragmentation Mass Spectra"

² Lehrstuhl für Bioinformatik, Institut für Informatik, Friedrich-Schiller-Universität Jena,
franziska.hufsky@uni-jena.de

und kann daher auch im Hochdurchsatz angewandt werden. Die dadurch entstehende Menge an Daten ist jedoch kaum manuell zu analysieren [Is09]. Vor der MS wird in der Regel eine Separationsmethode, z.B. Gas- oder Flüssigkeitschromatographie, angewandt, um die Moleküle voneinander zu trennen. Im Massenspektrometer werden die Moleküle ionisiert und entsprechend ihrem Masse-zu-Ladung-Verhältnis aufgetrennt. Das resultierende Massenspektrum ist ein zweidimensionaler Plot von Signalintensität gegen das Masse-zu-Ladung-Verhältnis (siehe Abb. 1(a)) und enthält in der Regel eine Reihe von Peaks – einen Peak je gemessenes Ion. Durch zusätzliche Fragmentierung der Moleküle gewinnt man Informationen über die Molekülmasse hinaus. Fragmentierung eines einfach geladenen Molekülions resultiert in einem einfach geladenen Fragment-Ion sowie einem neutralen Fragment (Neutralverlust), welches im Massenspektrometer nicht nachgewiesen werden kann.

Die Identifizierung unbekannter Metaboliten ist noch immer die Hauptproblematik bei der Analyse von (Fragmentierungs-)Massenspektren. Eine erheblich Zahl von Metaboliten ist bis heute unbekannt und somit in keiner Datenbank gelistet. Konventionelle Methoden, basierend auf Spektrenvergleich oder bereits bekannten Strukturformeln, stoßen daher häufig an ihre Grenzen [HSB14]. Die vollständige Strukturaufklärung eines bisher unbekannten Metaboliten ist besonders schwierig aufgrund deren hoher physikalischer und chemischer Vielfaltigkeit. Im Gegensatz zu Biopolymeren, wie z.B. Proteinen oder Glykanen, sind Metaboliten nicht aus sich wiederholenden Bausteinen zusammengesetzt. Zudem ist es in der Regel unmöglich Strukturinformation über Metaboliten aus Genomdaten zu gewinnen. Die vollständige Strukturaufklärung wird daher immer mittels mehrerer unabhängiger Methoden durchgeführt. Die Informationen, die man mittels MS gewinnt, helfen jedoch, den Suchraum stark einzuschränken, indem sie Hinweise auf die Struktur und Verbindungsklasse des Moleküls geben. In dieser Arbeit [Hu14] stellen wir computergestützte Methoden zur Analyse von Fragmentierungsmassenspektren unbekannter kleiner Moleküle, die in keiner Datenbank zu finden sind, vor.

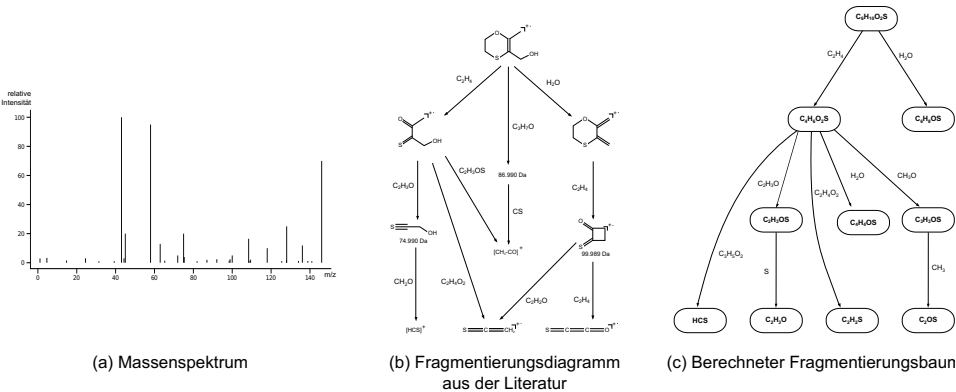


Abb. 1: (a) Fragmentierungsmassenspektrum, (b) das zugehörige Fragmentierungsdiagramm aus der Literatur [NV86], sowie (c) der mittels unserer Methode berechnete Fragmentierungsbaum eines Dihydro-1,4-Oxathiins.

2 Fragmentierungsbäume

MS-Experten evaluieren Fragmentierungsmassenspektren, indem sie Diagramme zeichnen, welche die im Massenspektrometer stattgefundenen Fragmentierungsereignisse manuell nachvollziehen (siehe Abb. 1(b)). Dafür muss jedoch die Molekülstruktur bereits bekannt sein. Die von uns entwickelte Methode extrahiert die Fragmentierungsschritte direkt aus den Daten, ohne Vorwissen bezüglich der Molekülstruktur. Wir modellieren diese Fragmentierungskaskaden mittels *Fragmentierungsbäumen* (siehe Abb. 1(c)) [BR08]. In einem Fragmentierungsbaum werden Knoten mit den Molekülformeln der Fragmente und Kanten mit den entsprechenden Fragmentierungsereignissen annotiert. Die Wurzel des Baumes enthält die Molekülformel des vollständigen Moleküls.

Die Berechnung von Fragmentierungsbäumen basiert auf einem graphentheoretischen Modell. Der sogenannte Fragmentierungsgraph ist ein gerichteter, azyklischer Graph. Für alle Peaks im Spektrum enthält dieser Graph jede mögliche Molekülformel, welche die Masse dieses Peaks erklärt, als Knoten. Die Knoten sind gefärbt, sodass alle Knoten, die den selben Peak erklären, die selbe Farbe erhalten. Zwei Knoten werden mittels einer gerichteten Kante verbunden, wenn der Nachfolgerknoten eine Teilmolekülformel des Vorgängerknotens enthält. Die Kante wiederum wird mit dem zugehörigen Neutralverlust beschriftet. Da diese Beziehung zwischen Knoten transitiv ist, ist somit auch der gesamte Fragmentierungsgraph transitiv. Um innerhalb des Graphen die wahrscheinlichste Erklärung der Fragmentierungsprozesse zu finden, werden die Knoten und Kanten gewichtet. Wir betrachten jeden Teilbaum, welcher in einer möglichen Erklärung des Molekülionenpeaks wurzelt, als hypothetischen Fragmentierungsbaum und suchen den Teilbaum mit maximaler Summe der Gewichte, wobei jede Farbe maximal einmal verwendet werden darf. Indem wir Bäume betrachten, wird jedes Fragment durch einen einzigen Fragmentierungsprozess erklärt; durch das Beschränken auf Bäume, in denen jede Farbe maximal einmal verwendet werden darf, wird jeder Peak nur von einem einzigen Fragment erklärt. Dass ein Peak von mehreren Fragmenten erklärt wird, kommt in der Praxis nur extrem selten vor und wird hier daher nicht betrachtet, um im generellen Fall das mehrfache Zählen von Peaks zu verhindern. Das eben beschriebene Problem wurde als MAXIMUM COLORFUL SUBTREE Problem formalisiert [BR08] und ist NP-schwer [FGN06] sowie APX-schwer [DFV09] sogar auf Binärbäumen. Des Weiteren gibt es keine Approximationsalgorithmen mit Approximationsgüte, die durch eine Konstante beschränkt wird [DFV09, Si11]. Zur Lösung des Problems wurden bereits mehrere exakte und heuristische Algorithmen entwickelt [BR08, Ra12b].

3 Auswertung von EI-Fragmentierungsspektren

Gaschromatographie mit Massenspektrometrie-Kopplung ist eine der ältesten und wichtigsten Verfahren zur Analyse kleiner Moleküle. Die dabei am häufigsten verwendete Ionenquelle ist die Elektronenstoßionisation (EI), welche zusätzlich zur Fragmentierung des Moleküls führt. Bei instabilen Molekülionen kann es bis zu einer vollständigen Fragmentierung kommen, sodass der Peak des Moleküls im Massenspektrum kaum sichtbar ist oder gar fehlt. Die Fragmentierung kleiner Moleküle durch Elektronenstoßionisation ist

bereits gut verstanden; die manuelle Interpretation dieser Fragmentierungsmassenspektren ist jedoch umständlich, zeitaufwendig und erfordert Expertenwissen (siehe z.B. Abb. 3). Automatische Methoden für die Analyse von EI-Massenspektren beschränken sich derzeit auf Datenbanksuche und regelbasierte Ansätze. Wir präsentieren eine automatisierte Methode zur Rekonstruktion von Fragmentierungsmustern kleiner Moleküle aus hoch-aufgelösten EI-Massenspektren mit hoher Massengenauigkeit.

Bei der Auswertung von EI-Spektren muss man beachten, dass weder die Molekülformel des vollständigen Moleküls, noch dessen Masse bekannt ist und ferner der Molekülionenpeak häufig nicht im Spektrum vorhanden ist. Für unsere Analysen gehen wir davon aus, dass der Molekülionenpeak vorhanden ist, wenn auch möglicherweise von sehr geringer Intensität. Mittels automatischer Erkennung der echten Signale im Spektrum und deren Auswertung berechnen wir Fragmentierungsbäume für EI-Spektren [Hu12]. Die Methode ermöglicht zusätzlich die Identifizierung des Moleküls des untersuchten Metaboliten und dessen Molekülformel. Neben einer Liste von bekannten, häufig vorkommenden Neutralverlusten verwendet unser Methode keine chemischen Vorkenntnisse und arbeitet vollkommen unabhängig von Datenbanken. Auf der anderen Seite benötigt unsere Methode hohe Massengenauigkeit der Messungen.

Da die Molekülformel des vollständigen Moleküls unbekannt ist, können die möglichen Erklärungen der Fragmente im Fragmentierungsgraph nicht auf Teilformeln der Molekülformel beschränkt werden, was zu einer drastischen Vergrößerung des Fragmentierungsgraphen führt. Dieser Effekt verstärkt sich, je größer das Molekül und je mehr Peaks im Spektrum vorhanden sind. Um den optimalen Teilbaum unter all den möglichen Teilbäumen zu finden, gehen wir daher in zwei Schritten vor (siehe

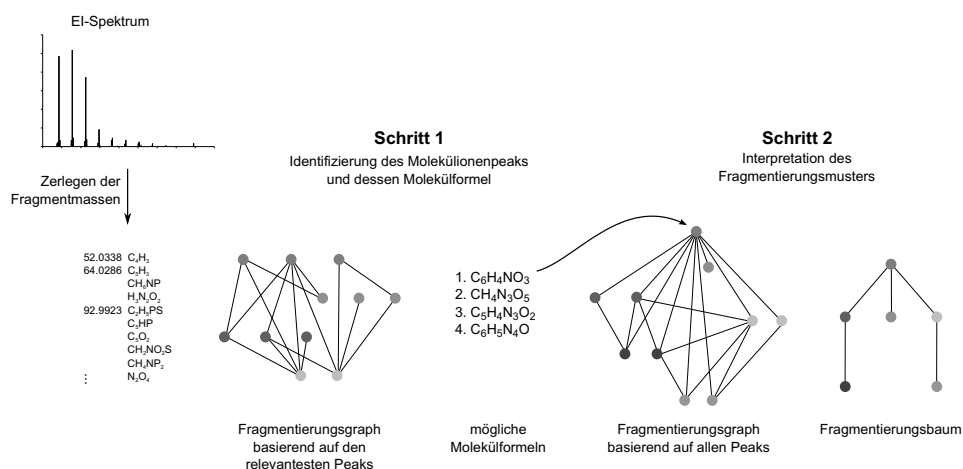


Abb. 2: Methode zur Auswertung von EI-Fragmentierungsspektren in zwei Schritten: Als Eingabe verwenden wir ein EI-Spektrum mit hoher Massengenauigkeit. Für jeden Peak werden alle möglichen Molekülformeln, welche diesen Peak erklären könnten, berechnet. In *Schritt 1* versuchen wir den Molekülionenpeak zu identifizieren und dessen Molekülformel zu bestimmen. In *Schritt 2* versuchen wir die beste Erklärung für das Fragmentierungsspektrum zu finden.

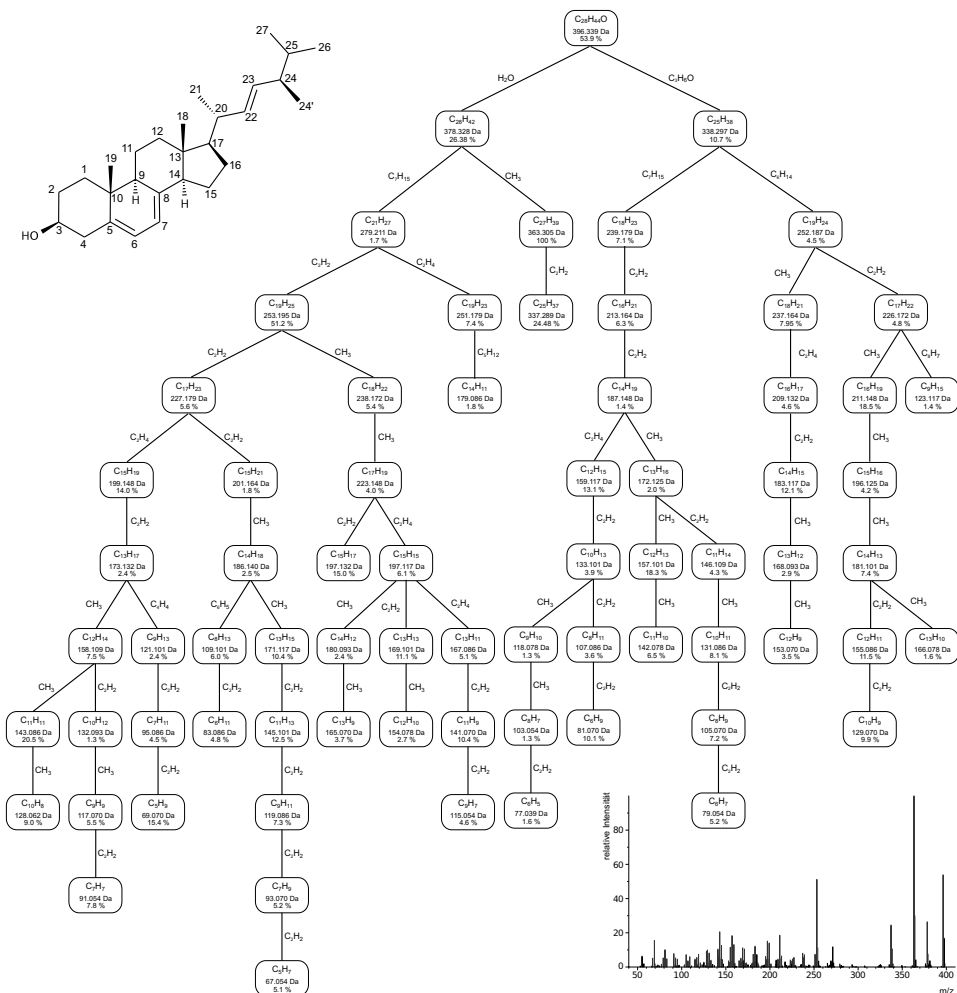


Abb. 3: Berechneter Fragmentierungsbaum für Ergosterol. Die Knoten enthalten die Fragmentformeln, die Massen der Fragmente (in Da) sowie die Intensitäten der zugehörigen Peaks im Spektrum. In der Anwendung ist die Identität und Strukturformel des Moleküls unbekannt. Der Baum dient als Grundlage für die weitere Auswertung des Massenspektrums. Die manuelle Interpretation des Massenspektrums wäre ohne diese Grundlage umständlich, zeitaufwendig und ohne Kenntnis der Molekülstruktur und entsprechendem Expertenwissen nicht möglich.

Abb. 2): (1) Zuerst versuchen wir den Molekülionenpeak zu identifizieren und dessen Molekülformel zu bestimmen. Dafür bauen wir einen Fragmentierungsgraph, welcher nur eine Teilmenge der Peaks aus dem Spektrum enthält. Diese Peaks erscheinen für die Bestimmung der Molekülformel besonders interessant. Jeder dieser Peaks ist potentiell der Molekülionenpeak. Die möglichen Molekülformeln der potentiellen Molekülionenpeaks werden geordnet entsprechend der Summe der Kantengewichte des optimalen Teilbaums, der in diesem Knoten wurzelt. (2) Ist die Molekülformel des vollständigen Moleküls bekannt, so wird erneut ein Fragmentierungsgraph gebaut, welcher in dieser Molekülformel wurzelt und mögliche Fragmentmolekülformeln für alle im Spektrum vorhandenen Peaks enthält. Aus diesem Graph wird nun der Fragmentierungsbaum berechnet, welcher die gemessenen Daten am besten erklärt.

Evaluiert haben wir unsere Methode sowohl auf simulierten als auch auf gemessenen Daten unter verschiedenen Gesichtspunkten. Die Identifizierung des Molekülionenpeaks und der Molekülformeln wurde auf EI-Spektren von 50 Molekülen unterschiedlicher Größe aus verschiedenen Verbindungsklassen ausgewertet. Das Molekülion wurde in 88 %, die Molekülformel in 78 % der Fälle richtig identifiziert. Die Methode funktioniert sogar für Spektren, in denen das Molekülion kaum sichtbar ist. Um die Qualität der Fragmentierungsbäume zu evaluieren, haben wir außerdem für 22 Moleküle, deren Fragmentierungswege in der Literatur beschrieben sind, Massenspektren simuliert. Massenspektrometrie-Experten bestätigten, dass die auf diesen Daten berechneten Bäume und die Fragmentierungsbäume der 50 gemessenen Moleküle, sehr gut mit dem veröffentlichten Fachwissen sowie bekannten Fragmentierungsmechanismen übereinstimmen. Alle wichtigen Fragmente und Fragmentierungsreaktionen der verschiedenen Verbindungsklassen konnten selbst in den komplexesten EI-Spektren unseres Datensatzes identifiziert werden (siehe Abb. 3).

4 Vergleich mittels Fragmentierungsbaumalignments

Im nächsten Schritt können die Fragmentierungsbäume zur Klassifizierung und Identifizierung von unbekannten Molekülen verwendet werden. Mittels Baumvergleich lassen sich strukturelle und chemische Ähnlichkeiten zu anderen, bereits bekannten Molekülen feststellen. Hierfür werden paarweise lokale Alignments von Fragmentierungsbäumen verwendet [Ra12a]. Lokales Baumalignment ist eine Generalisierung des lokalen Sequenzalignment-Problems. Das lokale Baumalignment zweier Fragmentierungsbäume enthält die Teilbäume, in denen ähnliche Fragmentierungskaskaden stattgefunden haben (siehe Abb. 4).

Um den Fragmentierungsbaum eines unbekannten Metaboliten mit einer großen Fragmentierungsbaumdatenbank vergleichen zu können, sind schnelle Algorithmen zur Lösung des Baumalignmentproblems erforderlich. Leider ist das Alignieren von ungeordneten Bäumen, zu denen die Fragmentierungsbäume zählen, nicht nur NP-schwer, sondern auch MAX SNP-schwer [JWZ95]. Dies bedeutet, es gibt kein Polynomzeit-Approximationsschema für dieses Problem, außer $P = NP$ [Ar98]. Wir haben drei exakte Algorithmen zur Berechnung von Baumalignments entwickelt: Zwei dieser Algorithmen

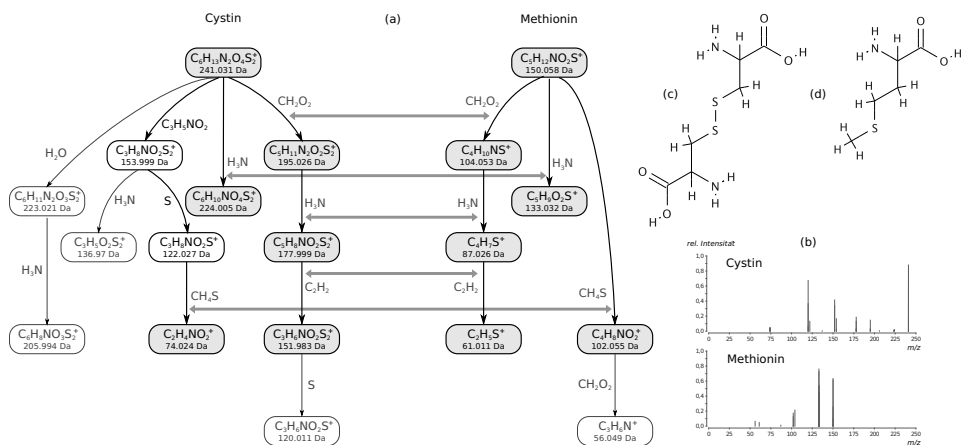


Abb. 4: Optimales Fragmentierungsbaumalignment (a) und Fragmentierungsmassenspektren für Cystin und Methionin (a). Die Massenspektren haben keine gemeinsamen Peaks. Die Molekülstrukturen von Cystin (c) und Methionin (d) werden zur Berechnung des Baumalignments nicht benötigt. Das Alignment enthält die gemeinsamen Fragmentierungswege.

basieren auf dynamischer Programmierung (DP), wobei der zweite Algorithmus auf dem ersten aufbaut, indem er die dünne Besetzung der DP Tabelle zu seinen Gunsten nutzt. Der dritte Algorithmus basiert auf ganzzahliger linearer Programmierung (englisch: *integer linear programming*, ILP).

Die Auswertung unserer Algorithmen auf drei Datensätzen hat gezeigt, dass mittels DP tausende Alignments in wenigen Minuten berechnet werden können, selbst für schwierige Instanzen. Der verbesserte DP Ansatz dominiert die klassische DP, was für einen der Datensätze sogar zu einer elffachen Beschleunigung führte (siehe Abb. 5). Überraschenderweise sind die beiden DP Algorithmen *wesentlich* schneller als das ILP. Nichtsdestotrotz hat das ILP das Potenzial, die für die DP-basierten Algorithmen schwierig zu lösenden Instanzen zu berechnen.

Fragmentierungsbaumalignments können unter anderem dazu verwendet werden um eine Reihe bekannter und/oder unbekannter Verbindungen, basierend auf deren Fragmentierungsbaumähnlichkeit, zu clustern (siehe Abb. 6), und um strukturelle Eigenschaften und/oder Verbindungsklassen vorherzusagen [Ra12a]. Basierend auf einer Reihe von Referenzverbindungen haben wir gezeigt, dass sich das Clustering mittels Fragmentierungsbaumähnlichkeiten mit den echten Verbindungsklassen deckt und konnten daraufhin für unbekannte Metaboliten aus isländischem Mohn die Verbindungsklassen und Strukturinformationen ableiten, um die nachfolgende NMR-Analyse zur vollständigen Strukturaufklärung zu vereinfachen.

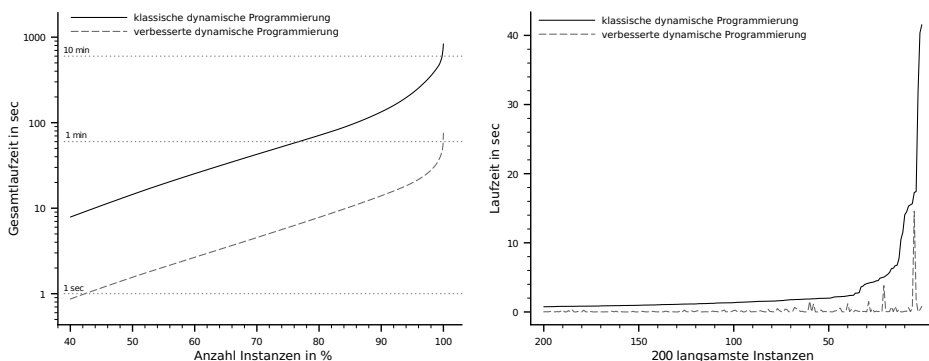


Abb.5: Laufzeiten für den schwierigsten der drei Datensätze mit 5151 Alignments. Links: Gesamtlaufzeit für x % der Instanzen, für welche die Berechnung des Alignments schneller lief als bei den verbleibenden Instanzen. Dabei wurden die Instanzen für jeden Algorithmus getrennt sortiert. Rechts: Individuelle Laufzeiten der 200 langsamsten Instanzen für den klassischen DP-Algorithmus. Instanzen wurden nach ihrer Laufzeit mittels klassischer DP sortiert. Die klassische DP wird für alle Instanzen von der verbesserten DP geschlagen.

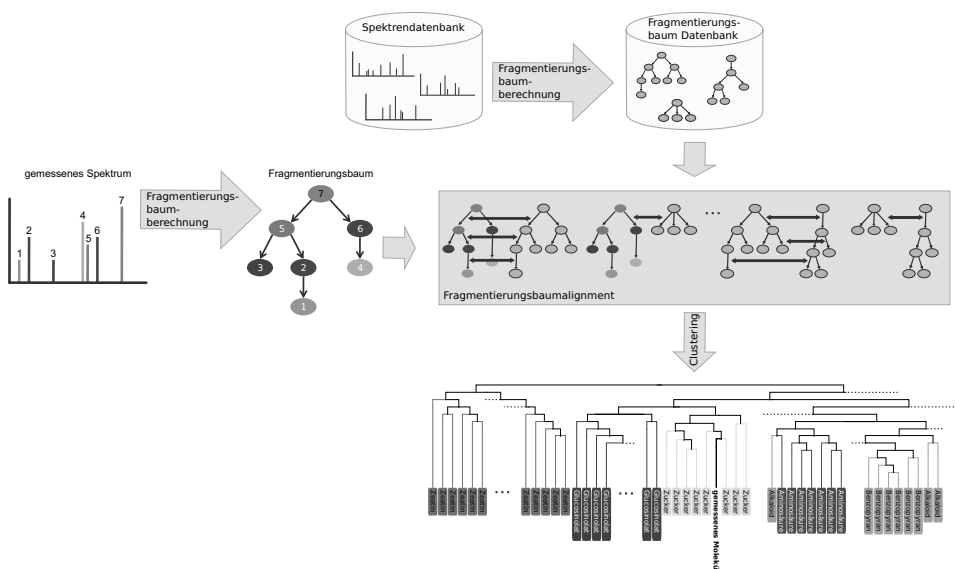


Abb. 6: Fragmentierungsbaumalignments zur Klassifizierung von unbekannten Molekülen. Aus dem gemessenen Spektrum wird ein Fragmentierungsbaum berechnet und gegen eine Datenbank von Fragmentierungsbäumen aligniert. Die Moleküle werden basierend auf den berechneten Fragmentierungsbaumähnlichkeiten geclustert. Ähnliche Moleküle (z.B. aus derselben Verbindungsklasse) clustern zusammen. Die Klasse des unbekannten Moleküls kann aus dem Cluster, in den es fällt, erschlossen werden.

5 Zusammenfassung

Wir haben neue *de novo* Methoden zur automatisierten Analyse von Fragmentierungsmassenspektren kleiner Moleküle entwickelt, die keine Kenntnis des zu untersuchenden Moleküls voraussetzen. Sowohl die Berechnung als auch der Vergleich von Fragmentierungsbäumen sind regelfreie Ansätze, die weder auf Spektren- noch auf Strukturdatenbanken basieren. Damit ermöglichen sie die Analyse von bisher vollkommen unbekannten Molekülen.

Literaturverzeichnis

- [Ar98] Arora, Sanjeev; Lund, Carsten; Motwani, Rajeev; Sudan, Madhu; Szegedy, Mario: Proof verification and the hardness of approximation problems. *J ACM*, 45(3):501–555, 1998.
- [BR08] Böcker, Sebastian; Rasche, Florian: Towards *de novo* identification of metabolites by analyzing tandem mass spectra. *Bioinformatics*, 24:I49–I55, 2008. *Proc. of European Conference on Computational Biology (ECCB 2008)*.
- [DFV09] Dondi, Riccardo; Fertin, Guillaume; Vialette, Stephane: Maximum Motif Problem in Vertex-Colored Graphs. In: *Proc. of Symposium on Combinatorial Pattern Matching (CPM 2009)*. Jgg. 5577 in *Lect Notes Comput Sci*. Springer, Berlin, S. 221–235, 2009.
- [Di11] Dieterle, Frank; Riefke, Björn; Schlotterbeck, Götz; Ross, Alfred; Senn, Hans; Amberg, Alexander: NMR and MS methods for metabonomics. *Methods Mol Biol*, 691:385–415, 2011.
- [Fe04] Fernie, Alisdair R; Trethewey, Richard N; Krotzky, Arno J; Willmitzer, Lothar: Metabolite profiling: From diagnostics to systems biology. *Nat Rev Mol Cell Biol*, 5(9):763–769, 2004.
- [FGN06] Fellows, Michael R.; Gramm, Jens; Niedermeier, Rolf: On the parameterized intractability of motif search problems. *Combinatorica*, 26(2):141–167, 2006.
- [HSB14] Hufsky, Franziska; Scheubert, Kerstin; Böcker, Sebastian: New kids on the block: Novel informatics methods for natural product discovery. *Nat Prod Rep*, 31(6):807–817, 2014.
- [Hu12] Hufsky, Franziska; Rempt, Martin; Rasche, Florian; Pohnert, Georg; Böcker, Sebastian: *De Novo Analysis of Electron Impact Mass Spectra Using Fragmentation Trees*. *Anal Chim Acta*, 739:67–76, 2012.
- [Hu14] Hufsky, Franziska: *Novel Methods for the Analysis of Small Molecule Fragmentation Mass Spectra*. Dissertation, Friedrich-Schiller-Universität Jena, Jena, Germany, 2014.

- [Is09] Issaq, Haleem J; Van, Que N; Waybright, Timothy J; Muschik, Gary M; Veenstra, Timothy D: Analytical and statistical approaches to metabolomics research. *J Sep Sci*, 32(13):2183–2199, 2009.
- [JWZ95] Jiang, Tao; Wang, Lusheng; Zhang, Kaizhong: Alignment of trees: An alternative to tree edit. *Theor Comput Sci*, 143(1):137–148, 1995.
- [LV09] Li, Jesse W-H; Vederas, John C: Drug discovery and natural products: End of an era or an endless frontier? *Science*, 325(5937):161–165, 2009.
- [NV86] Nevalainen, Vesa; Vainiotalo, Pirjo: Electron impact induced fragmentation of dihydro-1,4-oxathiines. 1. 2,3-substituted 5,6-dihydro-1,4-oxathiines. *Org Mass Spectrom*, 21(9):543–548, 1986.
- [Ra12a] Rasche, Florian; Scheubert, Kerstin; Hufsky, Franziska; Zichner, Thomas; Kai, Marco; Svatoš, Aleš; Böcker, Sebastian: Identifying the unknowns by aligning fragmentation trees. *Anal Chem*, 84(7):3417–3426, 2012.
- [Ra12b] Rauf, Imran; Rasche, Florian; Nicolas, François; Böcker, Sebastian: Finding Maximum Colorful Subtrees in practice. In: *Proc. of Research in Computational Molecular Biology (RECOMB 2012)*. Jgg. 7262 in *Lect Notes Comput Sci*. Springer, Berlin, S. 213–223, 2012.
- [Si11] Sikora, Florian: Aspects algorithmiques de la comparaison d'éléments biologiques. Dissertation, Université Paris-Est, 2011.
- [Wi07] Wishart, David S.: Proteomics and the human metabolome project. *Expert Rev Proteomics*, 4(3):333–335, Jun 2007.



Franziska Hufsky erhielt 2010 ihr Diplom in Bioinformatik von der Friedrich-Schiller-Universität (FSU) Jena, Deutschland. 2014 erlangte sie ihren Ph.D. im Bereich der algorithmischen Massenspektrometrie zur Identifizierung unbekannter kleiner Moleküle am Lehrstuhl für Bioinformatik (FSU Jena) unter der Leitung von Prof. Sebastian Böcker. Während dieser Zeit wurde sie durch ein Stipendium der *International Max Planck Research School* für chemische Ökologie in Jena unterstützt. Gefördert durch ein Postdoktoranden-Stipendium der Carl-Zeiss-Stiftung arbeitet sie derzeit zudem in der Gruppe von Prof.

Manja Marz am Lehrstuhl für RNA Bioinformatik und Hochdurchsatzanalyse (FSU Jena) an einem Projekt, welches die beiden Arbeitsgruppen und somit Metabolom- und Transkriptomanalyse vereint.