

Vollantrag auf Einrichtung eines LOEWE-Schwerpunktes an der Justus-Liebig-Universität Gießen (Federführung)
„AmbiProbe – Massenspektrometrische *in-situ*-Analytik für die Problembereiche Gesundheit, Umwelt, Klima und Sicherheit“

Projekt C.3: In-situ-massenspektrometrische Strukturaufklärung und Identifizierung

Projektleiter: Bernhard Spengler

Institut für Anorganische und Analytische Chemie, Justus-Liebig-Universität Gießen

Zusammenfassung

Ziel:

Entwicklung neuer Datenauswertungsverfahren, die *in-situ* auf der Basis hochgenauer Massenwerte Strukturinformationen liefern und Identifizierung ermöglichen.

Methoden:

Composition-based Sequencing (CBS) und Mass-based Classification (MBC).

Innovativer Aspekt:

- Vor-Ort-Analyse mit mobilen Geräten
- Sichere Identifizierung von Substanzklassen und individuellen unbekanntem Verbindungen
- Identifizierung und Charakterisierung unbekannter Verbindungen in Millisekunden
- Einbeziehung von Daten der Deuteriumaustausch-Kinetik in die Strukturanalyse

Verknüpfung mit anderen Schwerpunkt-Projekten:

Bereich A (Ionenquellen und Deuteriumaustausch), Projekt C1 (miniaturisiertes hochgenaues MS)

Ziele

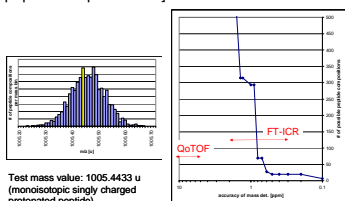
Während die kompositionsbasierte Strukturaufklärung für Peptide bereits erfolgreich etabliert ist, existieren bislang für andere Substanzklassen keine entsprechenden CBS-Algorithmen. Diese sollen im Rahmen des Projektes für eine Reihe relevanter Substanzklassen wie Mykotoxine und ihre Metaboliten entwickelt werden. Es ist dazu eine algorithmische Strategie aufzustellen, um eine allgemeine Lösung der kompositionsbasierten Strukturaufklärung für organische Moleküle zu finden. Eine grundlegende Substanzklasse, für die die CBS-Methode weiterentwickelt werden soll ist die der zyklischen Peptide, die im Zusammenhang mit Mykotoxinen von besonderer Bedeutung sind. Zyklische Peptide lassen sich mit herkömmlichen *de-novo*-Sequenzierungsverfahren nur äußerst unbefriedigend aufklären, da ihnen strukturbedingt Molekülanfang und -ende fehlen und damit ein Ansatzpunkt für typische wahrscheinlichkeitsbasierte Verfahren der Fragmentierungsweg-Analyse fehlt. Das CBS-Verfahren ist demgegenüber invariant bezüglich der Annahme bestimmter Fragmentierungsmechanismen und -wege und berücksichtigt stattdessen lediglich die möglichen Baustein-Zusammensetzungen beobachteter Fragmentionen. Der Anwendungsbereich der Zyklopeptid-Mykotoxine ist somit ein zentrales Zielgebiet der CBS-Entwicklungen dieses Projektes.

Neben der kompositionsbasierten Strukturaufklärung ist insbesondere bei Molekülen im unteren Massenbereich auch ohne Fragmentationsanalyse eine aussagekräftige Substanzklassenanalyse möglich. Die massenbasierte Klassifizierung (MBC) ermöglicht zum Beispiel die unmittelbare Unterscheidung von phosphorylierten und nicht-phosphorylierten Peptiden ohne Fragmentationsanalyse [Spengler 2008]. Die MBC-Methode wird im Rahmen des Projektes für die Substanzklassen der zyklischen Peptide, der verschiedenen weiteren Mykotoxine und der Explosivstoffe entwickelt. Es werden im Rahmen des Projektes Softwaresysteme entwickelt, die eine effiziente Umsetzung der komplexen Algorithmen beinhalten und eine anwendungsorientierte Nutzung *in-situ* ermöglichen.

Vorarbeiten

Für den Fall der *de-novo*-Peptidsequenzierung (Bestimmung der Aminosäuresequenz von unbekanntem Peptiden) ist das CBS-Verfahren bereits etabliert und erfolgreich eingesetzt worden [Spengler 2004, Thieu 2006, Römpf 2007, Spengler 2007]. Anders als bei herkömmlichen Verfahren der *de-novo*-Sequenzierung sind beim CBS-Verfahren keine weiteren Substanzkenntnisse oder Datenbanken erforderlich, so dass eine Analyse unmittelbar und universell erfolgen kann. Die erforderlichen Massengenauigkeiten liegen im oberen ppb-Bereich und können von (stationären) FT-ICR-Massenspektrometern oder von zu entwickelnden, mobilen MR-TOF-Systemen erreicht werden.

Das MBC-Verfahren ist bislang für die Unterscheidung von unmodifizierten und monophosphorylierten Peptiden beschrieben [Spengler 2008] und als Web-basiertes Online-System verfügbar (www.peptidecomposer.com).



Test mass value: 1005.4433 u (monoisotopic singly charged protonated peptide)

FT-ICR

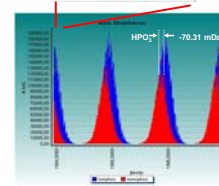
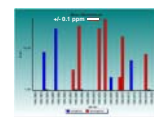
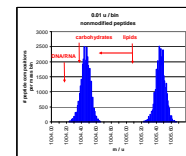
Einleitung

Die Auswertung hochgenauer Massenwerte von Molekül- und Fragmentionen ermöglicht die schnelle und sichere Strukturaufklärung und Identifizierung unbekannter Stoffe. Dabei können neuartige Strategien eingesetzt werden, die auf der Bestimmung der Zusammensetzung von Vorläuferionen und Fragmentionen beruhen (zusammensetzungsbasierte Sequenzierung/Strukturaufklärung, composition-based sequencing, CBS). Hochgenaue Massenbestimmungen erlauben die Aufstellung überschaubarer Listen von möglichen Elementar- bzw. Baustein-Zusammensetzungen unterschiedlicher Substanzklassen, jeweils für Vorläuferionen und jedes individuelle Fragmention. Solche Bausteine größerer Moleküle sind z.B. die natürlichen Aminosäuren in Peptiden oder Proteinen oder die Nucleotide in Nucleinsäuren.

Methoden

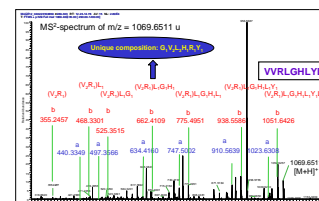
Die bislang publizierte CBS-Methode beruht auf der erschöpfenden Aufzählung von theoretisch möglichen Aminosäurekombinationen und der Berechnung und Auflistung der resultierenden möglichen Peptidmassen. Die für größere Peptide zunehmend langen Listen von möglichen Aminosäurekombinationen einer Zielmasse werden durch Verknüpfung mit entsprechenden Baustein-kombinations-Listen der erzeugten Produktionen reduziert, bis eine eindeutige Zusammensetzung gefunden ist bzw. die Liste sich nicht weiter einschränken lässt.

Die verbleibende(n) Peptidzusammensetzung(en) wird (werden) durchpermutiert und die zugehörigen theoretischen Fragmentmuster mit dem beobachteten Produktionenspektrum verglichen und bewertet gereiht. Das MBC-Verfahren ist für Peptide und deren Derivate beschrieben [Spengler 2008]. Es beruht auf der Auswertung lediglich der exakten Vorläuferionen-Massen und vergleicht diese mit vollständigen Listen von möglichen Vertretern bestimmter Substanzklassen (z.B. unmodifizierte Peptide und phosphorylierte Peptide). Beide Verfahren sind im Rahmen des Schwerpunkt-Projektes auf zusätzliche Substanzklassen zu erweitern.



Arbeitspakete

- AP 1: Definition von Algorithmen für die kompositionsbasierte Strukturaufklärung verschiedener Substanzklassen.
- AP 2: Entwicklung von Softwaremodulen für CBS und MBC.
- AP 3: Kopplung der Datenauswertungsmodulare an FT-ICR-, Mini-QIT- und MR-TOF-Daten.
- AP 4: Anwendung auf Substanzklassen der Mykotoxine, der bekanntesten Explosivstoffe und der zyklischen Peptide.



Literatur

Roempp A, Dekker L, Taban I, Jenster G, Boogerd G, Bonfrer H, Spengler B, Heeren R, Sillevis Smitt P, Luidert T (2007): Identification of leptomeningeal metastasis-related proteins in cerebrospinal fluid of patients with breast cancer by a combination of MALDI-TOF, MALDI-FTICR and nanoLC-FTICR mass spectrometry. *Proteomics* 7, 474-481.

Spengler, B. (2004) De Novo Sequencing, Peptide Composition Analysis and Composition-based Sequencing: A new Strategy Employing Accurate Mass Determination by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 15, 703-714.

Spengler B (2007): Accurate Mass as a Bioinformatic Parameter in data-to-knowledge conversion: FT ICR for peptide *de novo* sequencing. *Eur. J. Mass Spectrom.* 13, 83-87.

Spengler B, Hester A (2008): Mass-based classification (MBC) of peptides: Highly accurate precursor ion mass values can be used to directly recognize peptide phosphorylation. *J. Am. Soc. Mass Spectrom.* 19, 1808-1812.

Thieu VA, Kirsch D, Flad D, Mueller C, Spengler B (2006): Direct protein identification from nonspecific peptide pools by high-accuracy MS data filtering. *Angew Chem Int Ed* 45, 3317-3319.

Arbeitsplan	1. Jahr	2. Jahr	3. Jahr
C3 - in-situ-Strukturaufklärung			
Definition von Algorithmen			
Entwicklung von Softwaremodulen			
Kopplung an FT-ICR und MR-TOF-Daten			
Anwend. auf Substanzklassen und Ionenquellen			