



Doctoral Thesis

Peptides and peptide networks in human cerebrospinal fluid

Author(s):

Lamerz, Jens

Publication Date:

2004

Permanent Link:

<https://doi.org/10.3929/ethz-a-004861240> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

Diss. ETH No. 15623

**Peptides and Peptide Networks in
Human Cerebrospinal Fluid**

A Dissertation Submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH
For the Degree of
Doctor of Natural Sciences

Presented by
JENS LAMERZ
Dipl. Biochemist, University of Hannover, Germany
Born August 31st, 1974
Citizen of Hannover, Germany

accepted on the recommendation of
PROF. DR. L. SCAPOZZA, examiner
PROF. DR. R. CRAMERI, co-examiner
PROF. DR. G. FOLKERS, co-examiner

2004

Summary

Peptides represent an important class of biomolecules. They function as hormones, neuropeptides and cytokines; they organize and control thousands of biological processes. Peptidomics[®] is the technology for the comprehensive qualitative and quantitative description of peptides in a biological sample. This approach profiles the peptidic content from human body fluids or tissues by chromatography and subsequent mass spectrometry. The resulting profiling data represents static snapshots of the analyzed peptidome and reveals several thousand individual peptide signals per sample, originating from maturation and degradation of bioactive peptides. However, how these peptides are related, whether and how static snapshots can be assembled by statistical methods to reflect the dynamic processes of peptides remains to be answered. Furthermore, the underlying structure of peptide relations has to be addressed and peptides, that are characteristic for a disease, have to be identified. In this work, different statistical methods were applied to Peptidomics[®] profiling data from human cerebrospinal fluid (CSF) and extended to address these questions. As an important result, the concept of Correlation-Associated Networks (CAN), was introduced. CAN automatically relate peptide signals by means of correlation analysis. Examples for peptide-peptide relations found by CAN are posttranslational modifications of peptides, simultaneous clearance processing of bioactive proteins, common secretion of peptides by neuroendocrine cells as well as ubiquitin-mediated degradation. In the course of maturation and degradation of a bioactive peptide, many intermediate products are generated; many of these peptidic intermediates can be automatically grouped by CAN. These postulated relations are visualized as networks. In addition, novel peptides that were related to bioactive peptides by CAN represent potential therapeutic or diagnostic targets themselves.

Comparative studies of the topology of the networks analyzed underlying structures of CAN in human CSF: It was found, that few peptides dominate the overall connectivity in strictly selective networks. These peptides are often precursors or nodal points of further processing events. Peptides that are derived from the same precursor can be efficiently grouped by CAN. The concept of CAN was extended by a basic recognition algorithm of probable cleavage sites for peptidases and proteases in CSF. Moreover, the mass accuracy of mass spectrometry is used to interpret mass differences between peptide pairs as posttranslational modification, generating proposals for yet unknown peptide signals. This combination results in a

promising bioinformatic model that predicts the peptide sequence of network members with high accuracy, if they are related to an already identified peptide.

On the basis of this model, identification of peptide coordinates can be prioritized and a rapid overview of the peptide content of a novel sample source can be obtained.

This work also investigated methods of statistical analysis for the discovery of peptidic biomarkers by comparing samples from two groups, e.g. samples from healthy and diseased individuals. Peculiarities of mass spectrometric quantification and requirements of statistical analysis had to be considered to achieve an accelerated and automated statistical workflow. This workflow is exemplified by a study that comprehensively analyzes the peptide profile in CSF from Alzheimer patients and control individuals in order to detect reliable and robust biomarkers for Alzheimer's disease (AD). Amino acid sequencing identified nine possible marker candidates with remarkable robustness related to the pathophysiology of AD. They are fragments of the neuroendocrine protein VGF and C3f, a peptide derived from the complement factor C3. A marker panel of VGF peptides and C3f should allow differentiation of samples from AD patients and from control samples of individuals with non-AD dementia or without cognitive impairment.

In summary, the Peptidome of human CSF samples was analyzed from different perspectives by means of novel statistical methods. Novel, potential biomarker for Alzheimer disease were discovered, which may contribute to improve the diagnosis of this disease or stimulate new therapeutic approaches. CAN provide a very promising model to describe peptide-peptide relationships of biological sources. We expect that CAN will prove to be a supporting module for peptide identifications in Peptidomics[®] approaches and an important tool for the discovery of novel bioactive, therapeutic and diagnostic peptides.

Zusammenfassung

Peptide bilden eine wichtige Klasse der Biomoleküle. Sie agieren als Hormone, Neuropeptide, Cytokine und kontrollieren Tausende von Prozessen. Peptidomics[®] ist die Technologie zur umfassenden qualitativen und quantitativen Beschreibung von Peptiden in einer biologischen Probe. Dieser Ansatz charakterisiert die Peptid-Zusammensetzung von human Körperflüssigkeiten oder Geweben durch Chromatographie und anschließende Massenspektrometrie. Die daraus resultierenden Daten ergeben statische Momentaufnahmen des untersuchten Peptidoms und präsentieren je Probe einige tausend unterschiedliche Peptide, die bei der Reifung und anschließendem Abbau der bioaktiven Peptide entstehen. Angesichts dieser Datenmenge bleiben die Fragen offen, wie diese Peptide miteinander in Beziehung stehen und ob und wie mit Hilfe von statistischen Verfahren die dynamischen Prozesse der Peptide wiedergegeben werden können. Zudem soll die grundlegende Struktur der Peptid-Beziehungen adressiert und Peptide, die für bestimmte Erkrankungen charakteristisch sind, identifiziert werden.

In dieser Arbeit wurden verschiedene statistische Verfahren auf Daten des Peptidoms von humanem Liquor Cerebrospinalis (CSF) angewandt und ausgebaut, um diese Fragen zu beantworten. Als ein wichtiges Ergebnis wurde das Konzept der *Correlation-Associated Networks* (CAN) eingeführt. CAN setzen Peptidsignale durch Korrelationsanalyse automatisch miteinander in Verbindung. Beispiele für Peptid-Peptid-Beziehungen, die von CAN erkannt werden können, sind posttranslationale Modifikationen, gleichzeitige Spaltung von bioaktiven Proteinen, gemeinsame Sezernierung von Peptiden durch neuroendokrine Zellen als auch Ubiquitin-vermittelter Abbau. Im Laufe der Reifung und des Abbaus eines bioaktiven Peptides werden viele peptidische Zwischenprodukte gebildet; viele dieser Zwischenprodukte werden automatisch von CAN gruppiert. Die postulierten Beziehungen zwischen den Peptiden können als Netzwerke visualisiert werden. Zudem können neuartige Peptide, die durch CAN mit bioaktiven Peptiden verbunden sind, selber diagnostische *targets* darstellen.

Vergleichende Studien der Topologie dieser Netzwerke untersuchten die grundlegenden Strukturen der CAN in CSF. Es stellte sich heraus, dass in streng selektiven Netzwerken einige wenige Peptide das allgemeine Verbindungsverhalten des Peptid-Netzwerkes dominieren. Diese Peptide sind häufig Ausgangsstoffe oder Knotenpunkte für weitere Prozessierungsschritte. Peptide, die vom gleichen Vorläuferprotein abstammen, können durch CAN effektiv gruppiert werden. Das Konzept der CAN wurde mit einem einfachen

Algorithmus zur Erkennung von möglichen Schnittstellen von Proteasen und Peptidasen in humanem CSF erweitert. Zudem wird die Massenrichtigkeit der Massenspektrometrie bei der Messung der Peptidsignale dazu genutzt, Massendifferenzen zwischen verwandten Peptiden als posttranslationale Modifikation zu interpretieren und damit Vorschläge für unbekannte Peptide zu generieren. Diese Kombination liefert ein vielversprechendes bioinformatisches Modell, das die Peptid-Sequenz von Mitgliedern eines Netzwerkes voraussagt, wenn sie mit einem bereits identifizierten Peptid verbunden sind.

Auf der Basis dieses Modells kann die Identifikation von Peptiden priorisiert und ein schneller Überblick über die Peptid-Zusammensetzung erzielt werden.

Während der Doktorarbeit wurden weitere statistische Verfahren auf ihre Eignung für die Entdeckung von peptidischen Biomarkern erprobt. Besonderheiten bei der Quantifizierung mit Massenspektrometrie und Voraussetzungen der statistischen Datenanalyse wurden berücksichtigt um einen beschleunigten und automatisierten Arbeitsablauf zu entwickeln. Als Beispiel für diesen Arbeitsablauf wird eine Studie präsentiert, die das CSF Peptid-Profil von Alzheimer- und Kontrollpatienten analysiert, um verlässliche und robuste Biomarker für Alzheimer Demenz (AD) zu finden. Neun potentielle Biomarker mit bemerkenswerter Robustheit und Nähe zur AD Pathophysiologie wurden durch Massenspektrometrie identifiziert. Es handelt sich um Fragmente des neuroendokrinen Proteins (VGF) und C3f, ein aus dem Komplementfaktor C3 abgeleitetes Peptid. Eine Marker-Kombination aus C3f und VGF-Peptiden sollte eine diagnostische Unterscheidung der Alzheimer Patienten von Nicht-Alzheimer Dementen und Patienten ohne Demenz ermöglichen.

Somit wurde das Peptidom von humanem CSF durch neue bioinformatische Verfahren aus mehreren Blickwinkeln untersucht. Neuartige, mögliche Biomarker für die Alzheimer Demenz wurden entdeckt, die möglicherweise die Diagnose dieser Krankheit verbessern oder sogar neue Therapie-Ansätze anregen. CAN stellen ein sehr vielversprechendes Modell dar, um Peptid-Peptid- Beziehungen von biologischen Proben zu entdecken und zu beschreiben. Wir erwarten, dass sich CAN weiterhin als ein unterstützendes Modul zur Peptid-Identifikation und als wichtiges Werkzeug für die Entdeckung neuer bioaktiver, und diagnostischer oder gar therapeutischer Peptide erweisen wird.