

Prozessierung von HRMS-Daten mittels Python

Gesine Kuhnen^{1,2}, Lisa-Carina Class^{1,3}, Sascha Rohn^{2,3}, Jürgen Kuballa¹

¹GALAB Laboratories GmbH, Am Schleusengraben 7, 21029 Hamburg, Deutschland

²Institut für Lebensmitteltechnologie und Lebensmittelchemie, Technische Universität Berlin, Gustav-Meyer-Allee 25, 13355 Berlin, Deutschland

³Hamburg School of Food Science, Institut für Lebensmittelchemie, Universität Hamburg, Grindelallee 117, 20146 Hamburg, Deutschland

Korrespondierende Autorin: gesine.kuhnen@galab.de

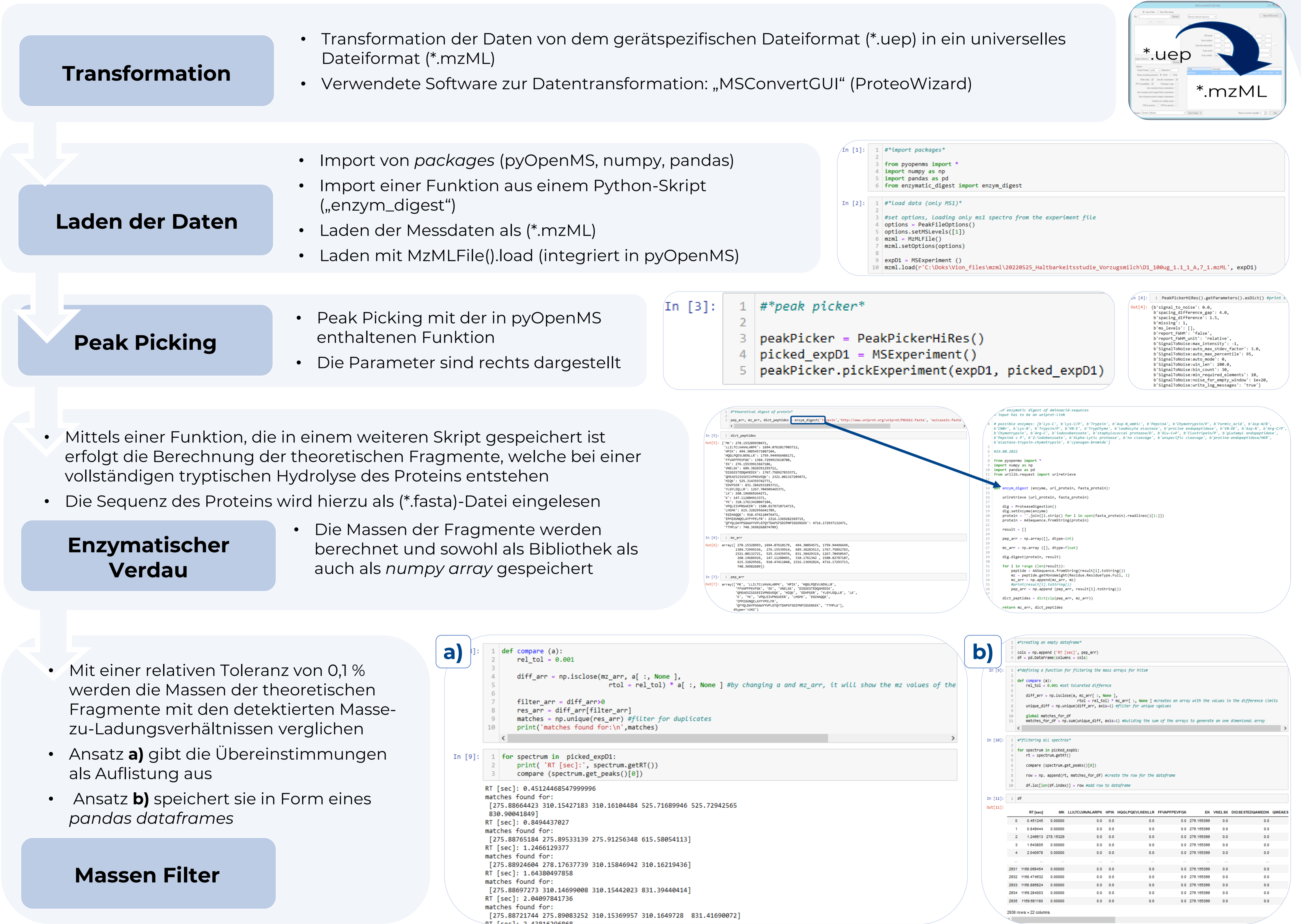
Einleitung

Die Auswertung von Messdaten ist ein elementarer Teil der Lebensmittelanalytik. Durch den Einsatz von Programmiersprachen kann eine universelle Anwendung und eine zumindest größtenteils von den gerätespezifischen Softwares unabhängige Auswertung ermöglicht werden. Gleichzeitig soll die Auswertung für ein breites Anwenderfeld zugänglich sein und eine hohe Verständlichkeit gesichert werden.

Die Programmiersprache Python ist vor allem im naturwissenschaftlichen Feld weit verbreitet. Python zeichnet sich durch eine vglw. einfache Syntax aus. Folge der Beliebtheit ist eine große Bandbreite sogenannter *packages* bzw. Bibliotheken, so auch bspw. spezifisch für massenspektrometrische und proteinbezogene Anwendungen.

Für die Daten-Prozessierung im Rahmen dieser Forschung wurden Datensätze aus Messungen von enzymatisch abgebauten Milchproteinen verwendet. Hierzu wurde Vorzugsmilch tryptisch mit der Methode nach GIANANTI¹ abgebaut. Die Proben wurden mittels hochauflösender Massenspektrometrie (HRMS, engl.: *high resolution mass spectrometry*) (Vion-IMS-qToF, Waters) untersucht und so die verwendeten Daten erzeugt.

Prozessierungsschritte



Zusammenfassung & Ausblick

Die große Bandbreite an *packages* ermöglicht das Bearbeiten diverser Fragestellungen mittels Python. PyOpenMS² bietet Funktionen für proteinbasierte Fragestellungen und vereinfacht so die Umsetzung dieser. Vorallem das Lesen der (*.mzML)-Daten wird hierdurch stark vereinfacht. Oben dargestellt ist das Filtern nach bestimmten Massen eines Proteins. Hierzu wird das Protein im (*.fasta)-Format eingelesen, die theoretisch entstehenden Fragmente werden ermittelt und die zugehörigen Massen berechnet. Durch Vergleich der theoretischen Massen und der gemessenen Masse-zu-Ladungsverhältnisse wurden die Übereinstimmungen ermittelt.

Im weiteren Vorgehen sollen mehrfach-geladenen Fragmente miteinbezogen werden. Um eine eindeutige Zuordnung der Fragmente zu gewährleisten sollen daraufhin dann die MS2-Spektren in die Auswertung miteinbezogen werden. Die Fragmentierung in b- und y-Fragmente werden hierbei auch theoretisch berechnet und mit den gemessenen Spektren verglichen werden.

Eine eindeutige Identifizierung von Peptiden in der Probe ist ohne die beiden genannten Erweiterungen des Ansatzes noch nicht möglich. Zur Suche nach bestimmten Peptiden kann das vorgestellte Skript verwendet werden, die Identifizierung des Peptids kann separat unter Einbezug des zugehörigen MS2-Spektrums erfolgen.

Literatur:

- [1] P. Giansanti et al., Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat. Protoc.* **2016**, 11, 993–1006, doi:10.1038/nprot.2016.057.
- [2] H.L. Röst et al., pyOpenMS: A Python-based interface to the OpenMS mass-spectrometry algorithm library. *Proteomics* **2014**, 14, 74–77, doi:10.1002/pmic.201300246.