



# Design einer Hochdurchsatz-Proteomik-Plattform

*Eine Plattform für wissenschaftliche und klinische Studien*

■ Die individuelle Antwort des Körpers auf Krankheit und Therapie stellt eine große Herausforderung dar. Patienten mit ähnlichen klinischen Parametern können sehr verschiedene Verläufe aufweisen und sehr unterschiedlich auf dieselbe Therapie reagieren. Meist liegt der Grund in kleinen physiologischen Unterschieden, die gemeinsam von Umwelteinflüssen, Ernährung, genetischem Hintergrund, und individueller (Krankheits-) Historie beeinflusst werden. In der Medizin der Zukunft werden diese Parameter eine Vorhersage über notwendige und Erfolg versprechende Maßnahmen liefern können. Am Beispiel von COVID-19 wird hier gezeigt, dass durch die Hochdurchsatz-

Proteomik bereits jetzt eine neue Generation prädiktiver und prognostischer Tests ermöglicht wird.

## Hochdurchsatz-Proteomik in der Universitätsmedizin

Die Anwendungsgebiete der klinischen Proteomik erstrecken sich von der Identifizierung von Marker-Peptiden/Proteinen bis hin zur Diagnostik. Proteine geben uns Auskunft über den funktionellen Status des Organismus beziehungsweise der Probe, sind im Vergleich zu anderen Biomolekülen sehr stabil und können sensitiv und kostengünstig mittels LC-MS bestimmt werden. Durch die große Anzahl an

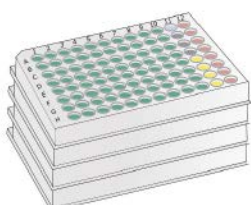
Einflussfaktoren ist die Identifizierung von individuellen prognostischen und prädiktiven Signaturen herausfordernd. Mit großen explorativen Kohorten und Tausenden Patienten können die Erfolgsaussichten, mit Einbezug der Faktoren, erhöht werden. Obwohl sich die Massenspektrometrie wie auch Affinitätsreagenz-basierte Technologien mittlerweile für großangelegte Untersuchungen eignen, gibt es nur begrenzt Studien mit einer Größe von mehreren Hundert Patienten. Ein Grund dafür liegt im fehlenden Zugang zu Facility-Strukturen, die spezifisch dafür ausgelegt wurden. Dies betrifft die Messkapazität, aber auch die spezialisierte Methodik und Organisation,

um Projekte in adäquatem Zeit- und Kostenrahmen umzusetzen. Die Charité in Berlin hat auf diesen Bedarf reagiert und 2019 mit der Core Facility „High Throughput Mass Spectrometry“ einen zentralen Forschungsdienstleister ins Leben gerufen. Ebenso wurde über die BMBF-Förderung des Nationalen Forschungskerns „MSTARS – Multimodale klinische Massenspektrometrie für die Untersuchung von Therapieresistenz“ die Infrastruktur geschaffen, um mechanistische Ansätze mit denen der Hochdurchsatzanalytik zu kombinieren.

## Design einer Hochdurchsatz LC-MS-Plattform

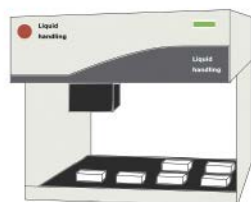
Die robuste und exakte Analyse von Hunderten bis Tausenden von Proteomik-Proben wird erst durch die Automatisierung von Laborprozessen, kurzen analytischen Laufzeiten, neuer Aufzeichnungsverfahren in der Massenspektrometrie sowie maschinellem Lernen in der Rohdatenanalyse ermöglicht. Alle Prozesse müssen dabei einem hohen Maß an Standardisierung und Dokumentation unterworfen werden. Der Durchsatz in der Facility der Charité liegt mit zwei Massenspektrometern bei 384 Proben pro Tag. Bei Laufzeiten von 7,5 Minuten pro Probe werden 300 Proteine in Plasma und 4.000 Proteine in Zelllysaten identifiziert. Die notwen-

### Proben im 96-well Format



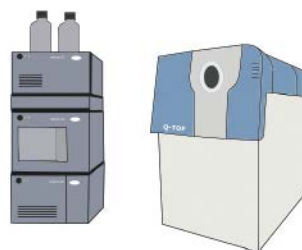
- 15 QC-Proben
- 80 Studienproben
- Blank

### Semiautomatisierte Probenbearbeitung



384 Proben/Tag

### Hochfluss-Flüssigkeitschromatographie und Scanning SWATH-MS



7,5-Minuten Messung/Probe



DIA-NN

Abb. 1: Workflow der Hochdurchsatz-Proteomik: Die Proben werden in 96-well Mikrotiterplatten arrangiert. Der Pipettierroboter bearbeitet parallel 4 Platten mit anschließend semi-automatischer Festphasenextraktion. Pro Tag können damit 384 Proben vorbereitet und an zwei Massenspektrometern gemessen werden.

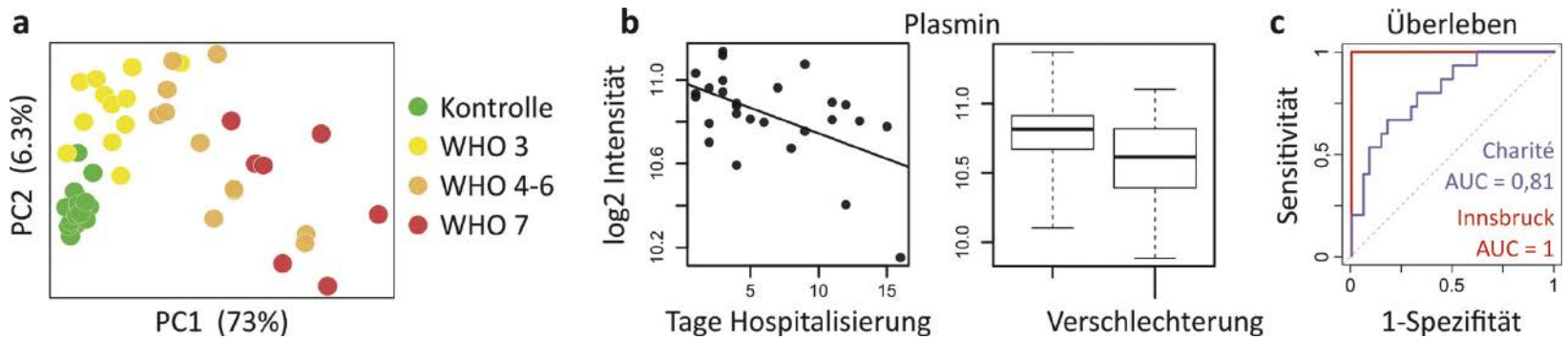


Abb. 2 a) Patienten mit unterschiedlichem COVID-19 Schweregrad können auf Grundlage des Serumproteoms unterschieden werden (PCA mit 29 Proteinen). b) Plasmaproteine wie Plasmin sind prädiktiv für die Dauer der Hospitalisierung und einer Verschlechterung des Krankheitsverlaufs bei WHO-3-Patienten. c) Vorhersage von Überleben oder Tod bei kritisch kranken Patienten, ab dem ersten Probenahmezeitpunkt auf kritischer Behandlungsstufe (WHO-Grad 7). Diese Abbildung basiert auf Grafiken zweier Preprint-Artikel.

digen technischen Entwicklungen, die diese Analytik ermöglichen, werden im Folgenden beschrieben.

Der Ursprung der größten Varianz in der Proteomik stellt weiterhin die Probenvorbereitung dar. Pipettierroboter verringern die technische Varianz, verhindern menschliche Fehler und erlauben es einer Person, vier Platten (384 Proben) routinemäßig pro Tag zu bearbeiten. Proteinreiche Biofluide, darunter Serum/Plasma, Liquor oder Synovialflüssigkeit, werden mit einem Urea-basierten In-Lösung-Verdau behandelt und über Festphasenextraktion entsalzen. Ein ähnlicher Durchsatz wird für frisches und fixiertes Gewebe sowie mit Detergenzien behandelte Proben mit dem SP3-Protokoll erzielt.

Die chromatografische Trennung der Peptide geschieht in Gradienten von fünf und weniger Minuten. Wir haben uns für deutlich höhere Flussraten (800  $\mu\text{l}/\text{min}$ ) als üblich in der Proteomik (Nano- und Kapillarfluss) entschieden, um Wasch- und Equilibationszeiten zu verkürzen sowie die Robustheit zu erhöhen. Mit neuesten Technologien für stationäre Phasen (z. B. Core-Shell-Partikel) werden effiziente chromatografische Trennleistung mit hohen Kapazitäten erreicht. Die Injektion von 5–10  $\mu\text{g}$  Peptid kompensiert den Sensitivitätsverlust.

Zur Messung der in sehr kurzer Zeit eluierender Peptide (3 Sekunden FWHM) erfordert es massenspektro-

metrischer Aufzeichnungsmethoden, die ausreichend Messpunkte pro Peak generieren. Datenunabhängige Verfahren (DIA) zeichnen sich hierbei aus, da sie alle verfügbaren Ionen in sequenziellen Isolationsfenstern fragmentieren und damit der Bias zu höchst-abundanten Peptiden verringert wird und mehr Ionen identifiziert werden. Quadrupol – Time Of Flight Geräte sind hervorragend geeignet die benötigten hohen Scanraten ( $> 30$  Hz) bei einer Auflösung von über 30.000 zu leisten. Es wurde gemeinsam mit Sciex die Scanning SWATH-Technologie auf dem Triple TOF 6600 entwickelt. Diese generiert durch kontinuierliches Verschieben/Scannen des Quadrupole-Fensters eine zusätzliche Dimension, in der jedem MS2-Signal eine Precursor-Masse zugeordnet wird. Peptide und Interferenzen werden damit besser unterschieden und 70% mehr Vorläuferionen in komplexen Proben identifiziert als mit konventionellen DIA Methoden.

In kurzen Gradienten eluieren Tausende von Peptiden in wenigen Minuten und erzeugen hochkomplexe Daten, bei denen eine Vielzahl an Ionen co-fragmentiert wird. Dies führt zu Multiplexspektren, deren Entmischung ausgefeilte Algorithmen erfordert. Die eigens entwickelte Software-Suite DIA-NN verwendet ein Ensemble tiefer neuronaler Netze, um echte Signale von Rauschen zu unterscheiden, sowie eine Reihe von Quantifizierungsalgorithmen, die den

Effekt von Interferenzen durch co-fragmentierende Peptide minimiert. Die „Verfeinerung“ öffentlicher Spektrallbibliotheken anhand projektspezifischer Daten erhöht die Konsistenz der Peptid- und Proteinidentifikation. DIA-NN ist in der Lage, mehrere Tausend Proben pro Tag auf einem herkömmlichen PC zu analysieren, wodurch eine schnelle Datenverarbeitung für die Proteomik mit hohem Durchsatz ermöglicht wird.

### Klinische Marker des Krankheitsverlaufs von COVID-19-Infektion

Der stark unterschiedliche Krankheitsverlauf von COVID-19-Patienten unterstreicht den Bedarf an schnell zu entwickelnden Tests, um die Behandlungen besser und individueller gestalten zu können sowie Ressourcen optimal einzusetzen. Mit der beschriebenen Hochdurchsatz-LC-MS Plattform konnte schnell auf die Pandemie reagiert werden. Im Serum der ersten 31 Patienten an der Charité Berlin wurden 37 Proteine identifiziert, die mit dem Schweregrad der Erkrankung assoziiert waren. 27 Proteine wurden in einer zweiten Kohorte in Plasma verifiziert. Sie spielen eine Rolle im Komplementsystem, der Akutphase und generellen Entzündungsreaktion. Ohne krankheitsspezifisch zu sein, liefern sie Aussagen über den Schweregrad und erlauben es Patienten zu stratifizieren. Um prognostische Aus-

sagen treffen zu können, wurden 139 Patienten (687 Proben) über den Zeitraum ihrer Hospitalisierung überwacht. Eine Validierung der Ergebnisse fand in einer unabhängigen Kohorte der Medizinischen Universität Innsbruck statt. Die zeitliche Änderung der Proteine erlaubte die Vorhersage der Verschlechterung des Patientenstatus, der Länge der Hospitalisierung sowie des Überlebens beziehungsweise den Tod von schwer kranken Patienten. Die Untersuchung rekonvaleszenter und unter Long-COVID leidender Patienten wird fortgesetzt. Ebenso wird die Impfantwort und der Effekt von immunmodulierenden Medikamenten untersucht. Das Ziel ist es, den Krankheitsprozess mechanistisch besser zu verstehen und in Zukunft anhand prognostischer und prädiktiver Modelle klinische Entscheidungen zu unterstützen. ■■

### Autoren:

Dr. Michael Mülleler, Core Facility High Throughput Mass Spectrometry  
Charité – Universitätsmedizin Berlin,  
Dr. Katharina Janek, Vadim Demichev und  
Prof. Dr. Markus Ralser, Institut für Biochemie,  
Charité – Universitätsmedizin Berlin,  
<https://biochemie.charite.de>  
Dr. Christoph Messner, The Francis Crick Institute,  
Molecular Biology of Metabolism Laboratory,  
London, Großbritannien,  
[www.crick.ac.uk](http://www.crick.ac.uk)