

Vom Massenspektrum zum humanen Interaktom

Bioinformatische Algorithmen visualisieren
komplexe zelluläre Kommunikationsnetzwerke

von Christian Münch

Nach der Entschlüsselung des menschlichen Genoms, die 2003 als abgeschlossen galt, ist die Aufgabe des 21. Jahrhunderts, das Proteom – die Gesamtheit aller Proteine einer Zelle – und dessen Interaktionen zu verstehen. Vor einigen Jahren wäre diese Sisyphusarbeit noch nicht möglich gewesen.

Jede menschliche Zelle enthält die genetische Information für über 20000 Proteine. Sie kreieren Zellen von erstaunlicher Vielfalt wie Muskel-, Nerven- oder Blutzellen. Die hoch spezialisierten Funktionen jedes Zelltyps werden ermöglicht durch ein charakteristisches Gemisch von Tausenden verschiedenen Proteinen, das so individuell und einzigartig ist wie ein Fingerabdruck.

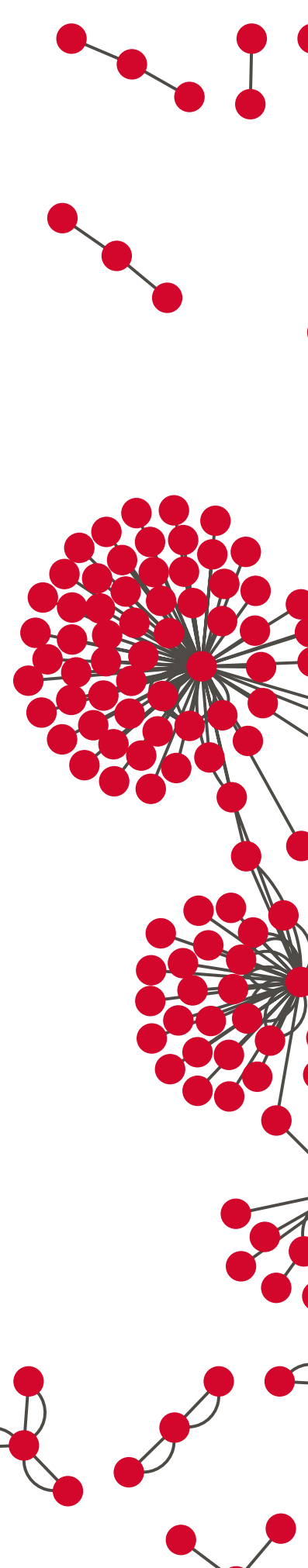
Winzige Veränderungen einzelner Proteine können genügen, um die gesamte Zelle zu beschädigen und schwerwiegende Krankheiten hervorzurufen. Ein Beispiel ist die Sichelzellanämie. Sie wird durch die Veränderung einer einzigen Aminosäure des Proteins Hämoglobin ausgelöst. Auch Krankheiten wie Alzheimer, Parkinson oder Krebs werden meist durch fehlerhaft oder nicht funktionierende Proteine ausgelöst. Aber selbst wenn einzelne Proteine für zelluläre Defekte verantwortlich gemacht werden können, sind üblicherweise die Mechanismen unklar, die zu veränderten Funktionen führen. Um das zu verstehen, muss man Veränderungen in der Zusammensetzung der Proteine, ihrer Mengenverhältnisse und im Zusammenspiel bei zellulären Abläufen untersuchen.

Noch vor wenigen Jahren war dies in den meisten Fällen unmöglich oder eine extreme Sisyphusarbeit. Erst die moderne Proteomanalyse mithilfe der quantitativen Massenspektrometrie gewährt Forschern heute den Überblick über die meisten zellulären Proteine. Sie erlaubt darüber hinaus die Analyse von Interaktionen zwischen verschiedenen Proteinen ebenso wie die von Modifikationen, die deren Aktivität steuern.

Massenspektrometer für die Lebenswissenschaften

Das erste Massenspektrometer baute vor ca. 100 Jahren der britische Physiker und Nobelpreisträger Joseph John Thomson an der University of Cambridge. Es wurde über viele Jahre vor allem zur Charakterisierung chemischer Verbindungen eingesetzt. Zahllose Veränderungen der Technik, insbesondere die MALDI-Methode, die in den 1980er Jahren von Michael Karas und Franz Hillenkamp an der Goethe-Universität entwickelt wurde, ermöglichten die Messung auch großer Moleküle, die in biologischen und medizinischen Fragestellungen eine Rolle spielen.

Allen massenspektrometrischen Verfahren ist gemeinsam, dass sie das Verhältnis von Masse zu (elektrischer) Ladung eines Moleküls messen und die Moleküle einer Mischung aufgrund dieses Verhältnisses sortieren. Als Ergebnis erhält man ein Bild mit Strichen, die verschiedene Masse-zu-Ladung-Verhältnisse repräsentieren. (Abb. 1) Diese simpel wirkenden Schwarz-Weiß-Grafiken enthalten komplexe Informationen, die erst prozessiert werden müssen, um für den Menschen leicht verständliche Daten und Bilder zu generieren. Besonders deutlich wird dies am Beispiel einer sogenannten Interaktionsanalyse (siehe Abb. 1), mit deren Hilfe wir in unseren Labors am Frankfurter Institut für Biochemie II schon viele Kommunikations-Netzwerke von Proteinen aufklären konnten. Ziel einer solchen Analyse ist es, alle Partner, mit denen ein bestimmtes Protein innerhalb einer Zelle interagiert, zu erfassen.

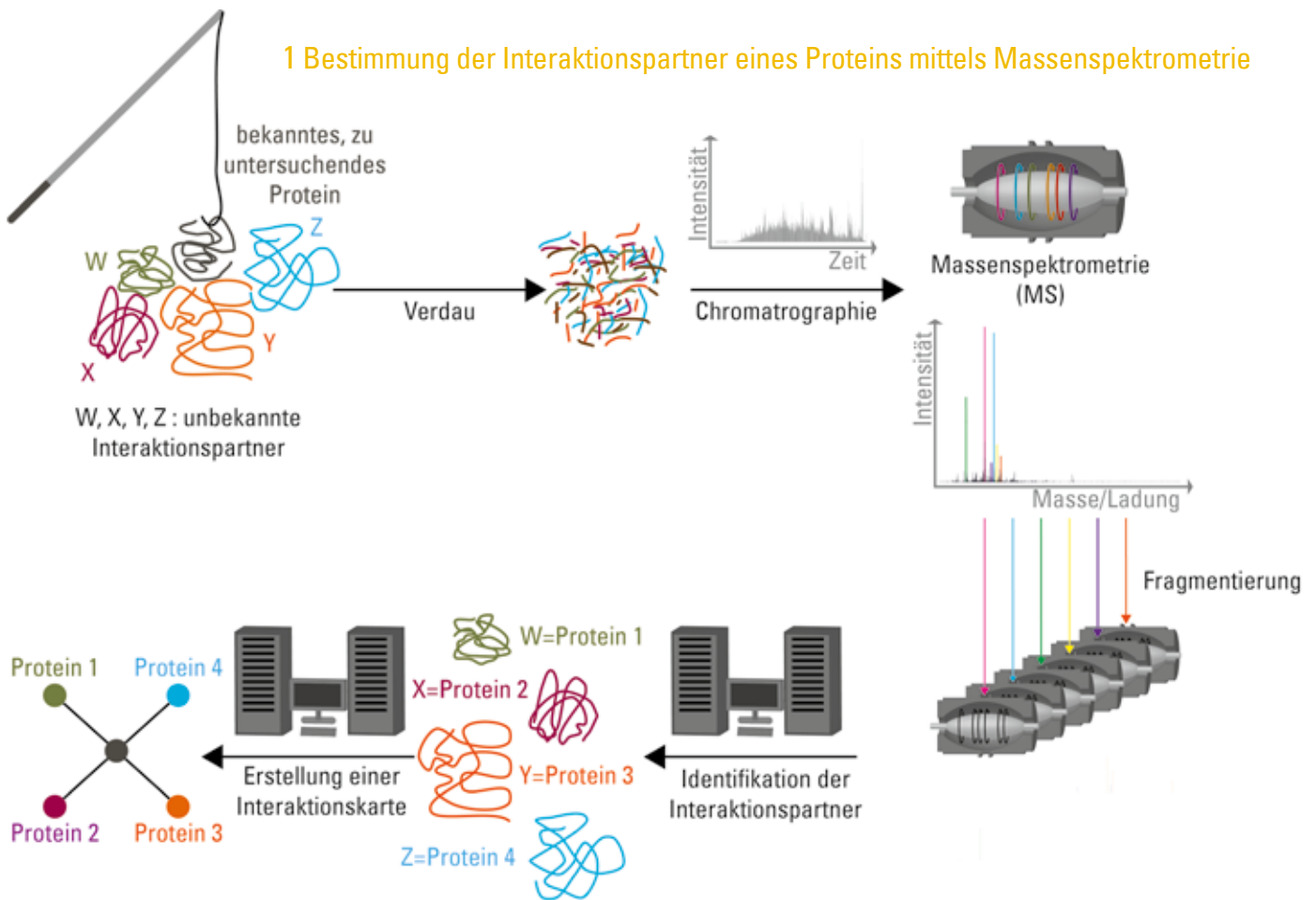




AUF DEN PUNKT GEBRACHT

- Die Funktion einer Zelle wird von der Interaktion ihrer Proteine gesteuert. Dabei hat jeder Zelltyp ein spezifisches Proteom.
- Um zu verstehen, mit welchen Partnern ein bestimmtes Protein in der Zelle interagiert, wird es mit diesen Partnern aus der Zelle isoliert, in kleinere Peptide aufgebrochen und das Gemisch im Massenspektrometer gemessen.
- Mithilfe bioinformatischer Algorithmen können die komplexen Interaktionen von Proteinen visualisiert werden.
- Weltweit arbeiten Wissenschaftler und große Konsortien daran, das komplette menschliche Interaktom zu entschlüsseln, um besser vorhersagen zu können, welche Auswirkungen die Veränderung einzelner Proteine oder Proteingemeinden auf die Funktion der Zelle hat.

1 Bestimmung der Interaktionspartner eines Proteins mittels Massenspektrometrie



1 Ein bekanntes, zu untersuchendes Protein wird »herausgefischt«, um interagierende Proteine zu finden. Das Proteingemisch wird verdaut, durch Chromatographie fraktioniert und in kleinen Portionen im Massenspektrometer gemessen. Hierfür wird erst das gesamte Gemisch im Orbitrap-Massenspektrometer (MS) gemessen, um die Massen der Peptide zu bestimmen. Danach werden die Massen mit der höchsten Intensität getrennt angereichert, im Massenspektrometer weiter aufgebrochen und diese Fragmente erneut im Orbitrap-Massenspektrometer gemessen (MS/MS). Mittels der Massen der intakten Peptide und deren Fragmenten können anhand von Computeralgorithmen die Proteine im Gemisch und die Interaktionspartner des zu untersuchenden Proteins ermittelt werden. Diese werden dann in ein Interaktionsnetzwerk zusammengefügt, um das visuelle Verständnis zu vereinfachen.

Bioinformatik ermöglicht Zuordnung von Proteinen

Zunächst isolieren wir aus Zellen das zu untersuchende Protein zusammen mit all seinen Interaktionspartnern. Da wir schonende Verfahren anwenden, bleiben die Interaktionspartner an das Protein gebunden, jedoch enthält das Gemisch üblicherweise auch noch Tausende anderer Proteine. Im Gemisch sind diejenigen Proteine am häufigsten, die mit dem zu untersuchenden Protein am häufigsten in Kontakt treten. Dann wird das Proteingemisch mit dem Enzym Trypsin in viele kleine Abschnitte zerlegt, die ca. ein Dutzend Aminosäuren enthalten und Peptide genannt werden. Dieser Arbeitsschritt ähnelt der Verdauung im Dünndarm.

Weil die Peptide kurz sind und ihre Masse weniger variabel ist, können wir die Proteine, denen sie entstammen, bei der folgenden Analyse leichter erkennen. Das Peptidgemisch wird dann aufgrund seiner unterschiedlichen chemischen Eigenschaften vorsortiert (durch Chromatographie) und in kleinen Portionen in das Massenspektrometer injiziert. Um die Peptide im elektromagnetischen Feld des Massenspektrometers voneinander trennen zu können, werden sie ionisiert. So kann man das Verhältnis von Masse zu Ladung für alle Peptide messen.

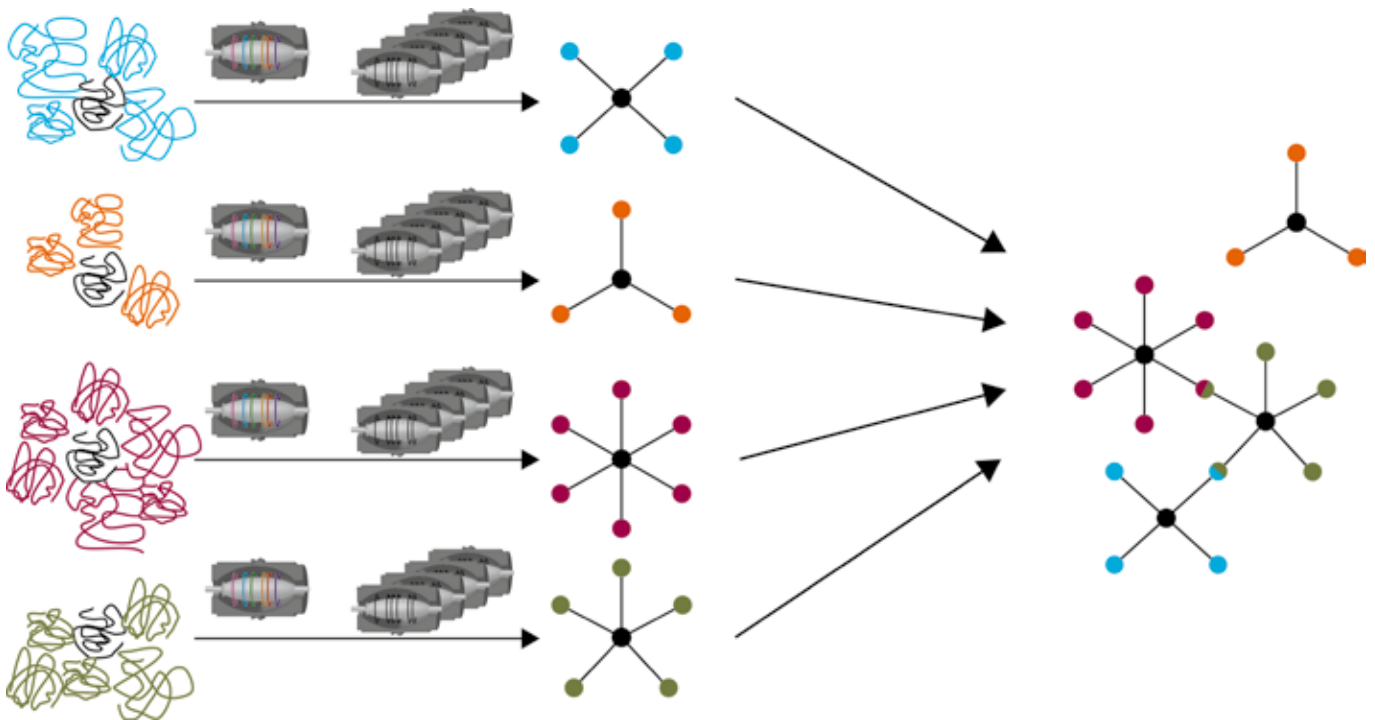
Wir verwenden für die Messung der Masse eine sogenannte Orbitrap.

Bei der Messung eines komplexen Proteingemischs kommen schnell 100 000 Bilder zusammen, deren manuelle Analyse unmöglich wäre. Abhilfe schaffen leistungsstarke Computer mit dazugehöriger spezialisierter Analyse-Software. Mit der Entwicklung der Bioinformatik hat sich auch hier Revolutionäres getan: Moderne Algorithmen berechnen die Zusammensetzung des jeweiligen Peptids aus den im Orbitrap gemessenen Werten, gleichen die Peptidsequenz mit existierenden Datenbanken ab und ermitteln so, zu welchem Protein das jeweilige Peptid gehört und wie häufig es im Gemisch vorhanden war. Am Ende dieser hoch komplizierten Rechengänge steht eine Liste mit der Häufigkeit aller gemessenen Proteine, aus denen sich die Interaktionspartner ermitteln und grafisch darstellen lassen. (Abb. 1)

Funktionale Zusammenhänge erschließen sich erst durch Visualisierung

Mithilfe dieser Methode können wir sämtliche Interaktionen eines Proteins in einer Zelle darstellen. Um zu verstehen, wie das Protein in das gesamte zelluläre Kommunikationsnetzwerk integriert ist, reicht dies allerdings oft nicht aus. Hierzu bedarf es Grafikprogramme, die einzelne

2 Erstellung von Interaktionsnetzwerken



PROTEINE

Proteine, auch Eiweiße genannt, bestehen aus unterschiedlich langen Ketten von Aminosäuren. Insgesamt gibt es 20 verschiedene Aminosäuren, welche in einzigartiger Kombination und Länge jedes Protein einer Zelle formen. Die Information zur Abfolge der Aminosäuren liegt auf der DNS im Zellkern. Daraus ergeben sich die dreidimensionale Struktur des Proteins und, oft im Komplex mit anderen Proteinen, seine Funktion. Durch enzymatische Verdauung im Magen-Darm-Trakt, aber auch bei der Vorbereitung unserer Proben für die Massenspektrometrie, entstehen kleine Proteinstückchen mit wenigen bis mehreren Dutzend Aminosäuren. Diese nennen sich Peptide. Da wir in unseren Experimenten bestimmte Verdauungsenzyme verwenden, von denen wir genau wissen, wo sie schneiden, können wir genau vorhersagen, mit welchen Aminosäuren unsere Peptide enden. Dies erleichtert die bioinformatische Analyse der Massenspektrometrieproben.

Interaktions-Netzwerke zu einem Gesamtbild zusammenfügen. Die Visualisierung ist ein entscheidender Schritt zum Verständnis der komplexen Interaktionen zwischen einzelnen Proteinen. Durch sie lassen sich die berechneten Hierarchien innerhalb von Netzwerken erkennen und interpretieren.

Oft zieht die bioinformatische Auswertung einen Rattenschwanz weiterer massenspektrometrischer Analysen nach sich – etwa, wenn das uns interessierende Protein mit anderen Proteinen in Kontakt tritt, über deren eigenes Interaktions-Netzwerk bislang nichts oder nur wenig bekannt ist. Seit einigen Jahren gibt es sogar Bestrebungen, für alle Proteine einer Zelle alle Interaktionen zu kartieren. Dabei handelt es sich um eine der größten biomedizinischen Herausforderungen seit der Sequenzierung des humanen Genoms.

Für solche erweiterten Protein-Interaktionsnetzwerke wird aus mehreren Millionen Zellen das erste Protein mit seinen Interaktionspartnern gereinigt und per Massenspektrometrie untersucht, aus einem zweiten Satz von Zellen das nächste usw. (Abb. 2) Für erweiterte Interaktionsnetzwerke ergeben sich so bereits Projekte von mehreren Monaten Dauer. An der Kartierung aller zellulären Interaktionsnetzwerke arbeiten weltweit zahlreiche Wissenschaftler und

2 Zur Darstellung von Proteininteraktionsnetzwerken wird die Analyse der Interaktionen einzelner Proteine (in schwarz) unabhängig voneinander durchgeführt. Dazu durchlaufen verschiedene Proteine die in Abbildung 1 dargestellten Schritte. Dabei entsteht für jedes Protein ein eigenes Interaktionsnetzwerk (hier in unterschiedlichen Farben), welches dann mit anderen zusammengeführt werden kann. Einzelne Proteine können mit mehreren Proteinen interagieren und so zweifarbig dargestellte Punkte schaffen, welche zwei schwarze Punkte miteinander vernetzen. Das gezeigte Netzwerk stellt die Interaktionen von vier Proteinen dar, kann aber durch die Analyse weiterer Proteine erweitert werden (s. S. 81) bis es alle Proteine einer Zelle darstellt.

große Konsortien. Der benötigte Arbeitsaufwand beträgt viele Jahre, trotz des hohen personellen und maschinellen Einsatzes bei der Probenherstellung, der Messung und der umfangreichen bioinformatischen Analysen. Millionen Massenspektrometrie-Messungen müssen zu Interaktionskarten einzelner Proteine verarbeitet und diese dann zu gigantischen Netzwerken zusammengestellt werden. Die Visualisierung dieser immer komplexeren Bilder stellt eine weitere große Herausforderung dar.

Ein Beispiel aus unseren eigenen Labors verdeutlicht die schnell anwachsende Komplexität: Kürzlich waren wir an einer vergleichsweise kleinen Interaktionsstudie für drei Proteine beteiligt, die eine wichtige Rolle in zellulären Stressantworten spielen (Yang et al., Nature 2016). Für diese drei Proteine fanden wir 119 Interaktionspartner. Um das gesamte Netzwerk besser erfassen und visualisieren zu können, wurden aus den 119 Interaktionspartnern 58 ausgewählt und auch deren Interaktionen untersucht. So entstand letztendlich ein Netz-



Der Autor

Dr. Christian Münch, Jahrgang 1982, studierte Biochemie in Tübingen und München, gefolgt von einer Promotion am Laboratory of Molecular Biology und der University of Cambridge (UK). Während dieser Zeit entwickelte er ein starkes Interesse an der Untersuchung von Proteinen in der Zelle, insbesondere im Kontext von Proteinfehlfaltung und Proteinqualitätskontrolle, die eine wichtige Rolle bei neurodegenerativen Erkrankungen spielen. Von 2012 bis 2016 forschte er an der Harvard Medical School in Boston (USA) und ist seit Dezember 2016 Gruppenleiter am Institut für Biochemie II an der Goethe-Universität. Dort setzt er seine Forschung zur Qualitätskontrolle in Proteinen fort und untersucht dabei Mitochondrien, welche hauptsächlich in der zellulären Energiegewinnung und im Metabolismus wichtig sind. Seit 2017 wird er durch das Emmy Noether-Programm der Deutschen Forschungsgemeinschaft gefördert.

ch.muench@em.uni-frankfurt.de

werk rund um drei spannende Proteine, welches aus 197 Proteinen und 341 Interaktionen besteht. Mithilfe dieser Ergebnisse können wir verstehen, wie die zelluläre Antwort auf spezifischen Stress reguliert wird, der z.B. durch Fehlfunktionen und Erkrankungen verursacht wird.

Eine Aufgabe von gigantischem Ausmaß

Das komplette menschliche Interaktom zu erfassen, nimmt gigantische Ausmaße an: Eine kürzlich publizierte Zwischenauswertung des Netzwerks enthielt fast 6000 Proteine, welche über 56000 Interaktionen ausführten (Huttlin et al., Nature 2017). Damit konnten die Autoren die Funktionen Hunderte kaum untersuchter Proteine erklären, neue »Proteingemeinden« definieren und dadurch auch Gruppen von Proteinen bestimmen, die eine Rolle bei spezifischen Krankheiten spielen.

Neben der bioinformatischen Auswertung und der grafischen Visualisierung sind auch bei der technologischen Entwicklung der Massenspektrometrie in den vergangenen Jahren enorme Fortschritte zu verzeichnen. Wesentliche Vorteile ergaben sich vor allem aus der Verkürzung der Messzeit, der stark verbesserten Sensitivität und neuen Methoden. Kontinuierlich werden leistungsstärkere Massenspektrometer entwickelt, womit auch die quantitative Messung von sehr komplexen Gemischen, beispielsweise aller Proteine einer Zelle, möglich wird. So können wir spezifisch Veränderungen des nahezu kompletten zellulären Proteoms untersuchen, beispielsweise durch bestimmte Krankheiten oder durch Behandlung der Zellen mit existierenden oder neu entwickelten Medikamenten. Kombiniert mit den sich entwickelnden neuen Erkenntnissen bei der Untersuchung des menschlichen Interaktoms, können wir auf diese Weise die Prozesse und Reaktionen in Zellen viel besser verstehen. Künftig werden wir Veränderungen im Proteom viel genauer vorhersagen können und damit zum Beispiel Therapie-relevante Vorhersagen treffen können. ●