

Proteomics und Massenspektrometrie

◆ Auch im Jahr 2 nach der Veröffentlichung des humanen Genoms schwanken die Schätzungen über die Anzahl der menschlichen Gene noch erheblich (35 000 bis 70 000). Die genaue Lage der neu entdeckten Gene im Genom ist unklar und die Funktionen der meisten Gene liegen noch im Dunkeln. Zunehmend rücken die Proteine als eigentliche Effektoren der Zelle in das Zentrum der Forschung. Ähnlich wie bei der Sequenzierung des humanen Genoms ist das Ziel, die Gesamtheit der Proteine, das Proteom, zu analysieren. Dieses Unterfangen ist allerdings ungleich komplexer: Das Proteom ist zellspezifisch und unterliegt zellinternen großen Schwankungen, etwa durch den Zellzyklus oder als Antwort auf äußere Einflüsse.

Fortschritte in der Massenspektrometrie (MS) revolutionierten die systematische Analyse von Proteinen, deren Identifizierung und Modifikationsanalyse. Ihren Anfang nahm diese Entwicklung durch bahnbrechende Arbeiten, welche zu neuen Volatisierungs- und Ionisierungsverfahren führten: der Matrix-assistierten Laser-Desorption/Ionisierung (MALDI)^{1,2} und der Elektrosprayionisierung (ESI).³ Bei MALDI wird der Analyt in einer festen Matrix aus einer niedermolekularen, meist aromatischen Verbindung durch Laser-Beschuss in die Gasphase überführt und ionisiert. Der genaue Mechanismus wird noch diskutiert.^{4,5} Dieses Verfahren ist so schonend, dass Peptide und kleine Proteine in aller Regel intakt beobachtet werden können. Bei ESI wird der Analyt in einer flüssigen Matrix (z.B. einem sauren Gemisch aus Wasser und organischem Lösungsmittel) aus einer Kapillare durch Anlegen einer elektrischen Spannung vernebelt. Eine Reihe von Studien konnte im vergangenen Jahr zum Verständnis des Mechanismus beitragen, der den sanften Übergang des Analyten in die Gasphase erklärt.⁶ Beide Techniken haben inhärente

Stärken, welche zu bevorzugten Anwendungsbereichen führen. Gegenwärtig weiter verbreitet ist ESI.

Identifizierung von Proteinen

◆ Von einem gereinigten Protein reichen generell 50 bis 100 fmol aus, um es mit Massenspektrometrie zu identifizieren, d.h. mit einem Eintrag in einer Sequenzdatenbank zu verknüpfen. Für gewöhnlich trennt Gelelektrophorese das Protein im letzten Reinigungsschritt von anderen Proteinen. Die Masse des Proteins lässt sich bei diesen Mengen nicht genau bestimmen und wäre wegen auftreten der Modifikationen ohnehin nicht hinreichend zur Identifikation. Statt dessen verdaut eine spezifische Protease das zu analysierende Protein in Peptide. Typischerweise wird hierzu Trypsin verwendet, welches die Peptidbindung spezifisch C-terminal zu Lysin und Arginin hydrolysiert. Die Massen dieser Peptide werden im Massenspektrometer bestimmt und gleich einem Fingerabdruck mit einem theoretischen Verdau aller Datenbankeinträge im Computer verglichen.⁷⁻¹¹ Ist diese Methode nicht erfolgreich, so können – etwas aufwendiger – einzelne Peptide in einem Tandem-Massenspektrometer isoliert und fragmentiert werden (Abbildung 1).¹² Aus den Differenzen der Fragmentmassen lässt sich ein Teil der Peptidsequenz ablesen und kann zusammen mit der Peptidmasse als peptide sequence tag zur Datenbanksuche verwendet werden.¹³ Alternativ können alle Fragmentmassen direkt zur Datenbanksuche eingesetzt werden.¹⁴⁻¹⁶

Der Peptidmassen-Fingerabdruck wird üblicherweise mit MALDI gekoppelt an einen Flugzeitanalysator (MALDI-TOF, TOF von time of flight) mit einem Fehler von weniger als 20 ppm bestimmt. Der Probenvorbereitung kommt große Bedeutung zu.¹⁷ Spezielle Probenhalter zur Kristallisation des Matrix-Probengegemisches beeinflussen die Messung positiv.¹⁸ Der geringe zeitliche Aufwand und die Robustheit von MALDI-TOF-MS gestatten, dass durch Automatisierung und den Einsatz

von Robotern ein Durchsatz von theoretisch mehreren Tausend Proben am Tag möglich ist.¹⁹

An Grenzen stößt das Verfahren bei der Analyse von Proteingemischen oder wenn die Proteinsequenz lediglich in Fragmenten in der Datenbank vorliegt. Letzteres trifft auf die meisten humanen Proteine zu, da deren kodierender Bereich im Genom zwar sequenziert, in der Regel aber in viele kleine Exons unterteilt ist, deren Lage und Grenzen nicht immer vorhersagbar sind.²⁰ Durch cDNA-Sequenzierung sowie eine verbesserte computergestützte Genvorhersage wird dieses Problem in der nächsten Zukunft an Bedeutung verlieren. Zusätzlich wird MALDI in neuen Entwicklungen im Instrumentendesign erfolgreich mit Tandem-Massenspektrometern gekoppelt, wodurch von einigen Peptiden in jeder Probe unterstützend zum Peptidmassen-Fingerabdruck Fragmentinformationen gewonnen werden können. Dies steigert einerseits die Selektivität der Datenbanksuche, andererseits verbessert es auch die Analyse einfacher Mischungen.

Im Routinebetrieb wurden Fragmentinformationen von Peptiden bislang hauptsächlich an Geräten erhalten, die ESI und Tandem-MS kombinieren. Da ESI konzentrationsempfindlich ist, erlauben verlängerte Messzeiten durch geringere Flussgeschwindigkeiten eine höhere Empfindlichkeit, was zur Entwicklung von Nano-ESI mit Flüssen von 5 bis 20 nL·min⁻¹ führte.²¹ Von 100 fmol eines Peptidgemisches (aufkonzentriert in einem Mikroliter) lassen sich mittels Nano-ESI einige Dutzend Fragmentspektren aufnehmen.

Zunehmend werden allerdings auch komplexere Proben massenspektrometrisch analysiert. Peptidgemische werden dabei aus Proteinkomplexen, Zellkompartimenten oder Zellfraktionen durch Verdau gewonnen und mit Mikro-HPLC aufgetrennt, die direkt an ein ESI-Massenspektrometer gekoppelt ist.^{22,23} Die Flussgeschwindigkeiten betragen etwa 200 nL·min⁻¹, wobei die Empfindlichkeit dieser Methode durch die Aufkonzentrierung der einzelnen

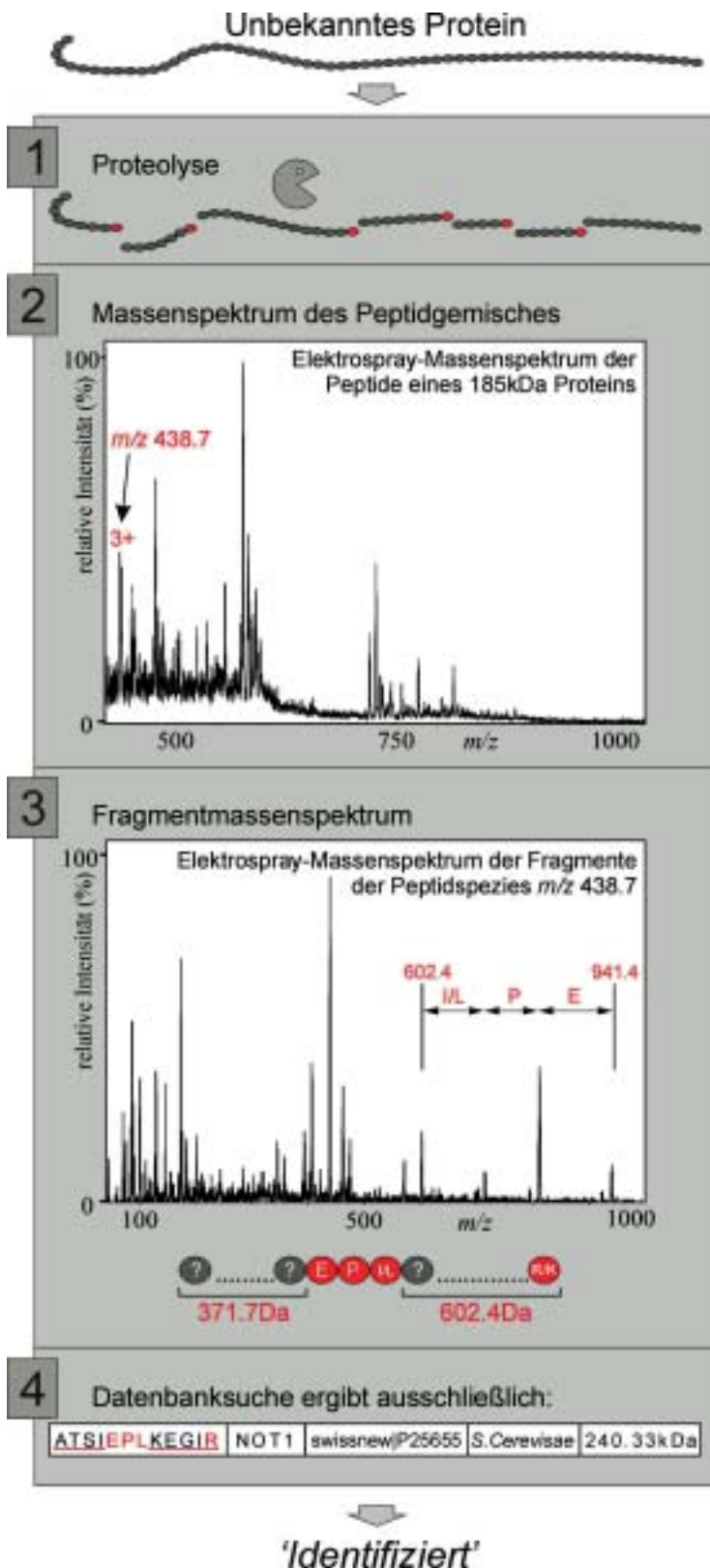


Abb. 1. (1) Ein gereinigtes Protein unbekannter Identität wird proteolytisch in Peptide zerlegt. (2) In einem Tandem-Massenspektrometer werden erst die Massen der erhaltenen Peptide bestimmt, um dann (3) einzelne Peptidspezies zu fragmentieren. Einige Fragmentmassen unterscheiden sich gerade um die Masse bestimmter Aminosäuren, wodurch ein Teil der Peptidsequenz ausgelesen werden kann. (4) Mit dieser Sequenz, deren Position im Peptid und der Peptidmasse kann die passende Proteinsequenz per Datenbanksuche gefunden und so das Protein identifiziert werden.

Peptidspezies oft besser als bei der Nano-ESI ist. Von mindestens ebenso großer Bedeutung ist, dass sich die Analyse automatisieren lässt. Außerdem können mehr Peptide beobachtet werden, da die wechselseitige Signalunterdrückung reduziert wird. Durch zweidimensionale HPLC, basierend auf ionischen Wechselwirkungen in der ersten Dimension (Kationenaustauscher) und hydrophoben in der zweiten (Umkehrphase), konnten in einer einzigen Analyse nahezu 1500 Proteine von Bäckerhefe identifiziert werden.²⁴⁾ Dies war eine beeindruckende Demonstration der Technik, die bei der Untersuchung von Zellkompartimenten und damit der räumlichen Zuordnung von Proteinen in der Zelle ihre Anwendung finden wird.

Quantifizierung

◆ Mit dem analytischen Zugang zu komplexen Proteingemischen eröffnen sich Forschungsgebiete, die quantitative Aussagen erfordern. So können medizinisch relevante Fragen angegangen werden, z. B. die nach dem Unterschied zwischen Krebszellen und gesundem Gewebe. Trotz der technischen Fortschritte ist die Komplexität der menschlichen Zelle für eine direkte Analyse aller Proteine zu groß; es muss auf die eine oder andere Art vorfraktioniert werden. Zudem ist eine vergleichende Analyse nur möglich, wenn verlässliche Methoden zur Quantifizierung entwickelt werden. Mit dem Einsatz von stabil-isotopenmarkierten Reagentien ist man dabei einen großen Schritt voran gekommen.^{25,26)} Probe A wird mit einem leichten Agens markiert und Probe B mit einem schweren, die sich beide lediglich durch den Austausch von stabilen Isotopen (z. B. ^{15}N gegen ^{14}N) unterscheiden. Anschließend werden die Proben vereinigt und gemeinsam fraktioniert, verdaut und analysiert. Markierte Peptide geben sich im Massenspektrum als Signaldubletts zu erkennen, deren Flächenverhältnis die relativen Mengen des entsprechenden Proteins in den Ausgangsproben A und B angibt. Durch aminosäure-

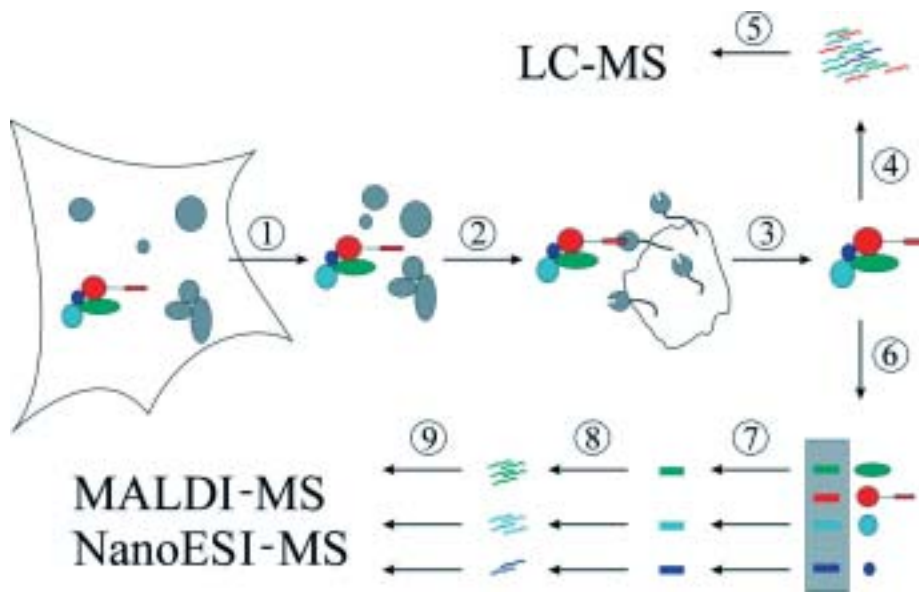


Abb. 2. (1) Zellen werden lysiert, in denen ein Protein mit einem Protein-affinity tag versehen wurde, um seine Isolierung zu vereinfachen und seine Interaktionspartner zu identifizieren. (2) Der Protein-affinity tag bindet an eine zugefügte spezifische Matrix und ermöglicht so die Reinigung des Zielkomplexes. (3) Der eluierte Komplex kann (4) direkt proteolytisch in Peptide zerlegt werden. (5) LC-MS analysiert die Peptide und identifiziert die Proteine. (6) Alternativ können die einzelnen Komplexkomponenten durch Gelelektrophorese getrennt, (7) isoliert und (8) proteolytisch zerlegt werden. MALDI-MS und Nano-ESI-MS ergibt die Identität der Proteine.

spezifische Agentien, die einen affinity tag (z.B. Biotin) tragen, können die markierten Peptide aus der Probe isoliert werden. Dadurch kann die Probenkomplexität reduziert werden.²⁷⁾ Diese Technik, die bislang bei cysteinhaltigen Peptiden eingesetzt wurde, spielt dann eine Rolle, wenn die Analyse von Proben nicht durch die zur Verfügung stehende Materialmenge begrenzt ist, sondern durch ihre Komplexität.

Proteinvarianten

Die besprochenen Verfahren reduzieren die Proteinidentifikation auf die leichter zu bearbeitende Peptididentifikation. Dabei wird das Protein ausschnittsweise sequenziert und die experimentell nicht bestimmten Bereiche durch einen Datenbankeintrag ermittelt. Obschon ein Protein massenspektrometrisch vollständig sequenziert werden kann, findet eine solche Analyse wegen der dafür benötigten Materialmengen und des Arbeitsaufwandes routinemäßig nicht statt. Aus der Vielzahl möglicher Proteinvarianten, die z.B. durch alternatives Spleißen oder Proteinmodifikationen aus einem Gen hervorgehen

können, folgt damit zwar eine gewisse Unsicherheit, welches Protein im Detail vorliegt. Das zugehörige Gen aber ist einwandfrei zuzuordnen.²⁸⁾

Die selektive Beobachtung von Sequenzvielfalt ist schwierig. Allerdings erlauben spezielle massenspektrometrische Experimente die selektive Beobachtung von modifizierten Peptiden in einem Peptidgemisch. Bei einem dieser Experimente, dem Vorläuferionenscan (precursor ion scan), werden beispielsweise nur jene Peptide beobachtet, die ein Fragment mit einem bestimmten Massen-Ladungsverhältnis (m/z) hervorbringen.^{29,30)} Massenspektrometrie misst generell m/z , die Masse aber ist meist zugänglich: Hochauflösende Geräte ergeben isotopenaufgelöste Signale, aus deren Differenz die Ladungszahl bestimmt werden kann. Für die Untersuchung der Tyrosinphosphorylierung, der in vielen Signalübertragungskaskaden in der menschlichen Zelle eine zentrale Rolle zukommt, kann z.B. das Immoniumion von Phosphotyrosin (m/z 216,043) beim Vorläuferionenscan Verwendung finden.³¹⁾ Ein sich stetig vergrößernder Satz von Proteinmodifikationen lässt sich auf ähnlichem Weg untersuchen.

Protein-Protein-Wechselwirkungen und Multiproteinkomplexe

Die Funktion und genaue Zusammensetzung der meisten Produkte des humanen Genoms und der anderen sequenzierten Genome sind unbekannt. Bekannt ist allerdings, dass viele Proteine in Proteinkomplexen wirken. Nach der Devise „schuldig durch Assoziation“ kann neuen Proteinen ein Funktionsbereich zugeordnet werden, wenn sie als Komponente eines bestimmten Proteinkomplexes nachgewiesen werden können.³²⁾ Dieser Strategie folgend werden Proteinkomplexe isoliert und deren Komponenten massenspektrometrisch identifiziert.

Proteinkomplexe lassen sich sehr rein gewinnen, wenn ein Antikörper gegen eine der Komplexkomponenten erhältlich ist. Die Herstellung von Antikörpern ist allerdings aufwendig und nicht immer erfolgreich.

Allgemeinen Zugriff auf Proteinkomplexe erlaubt die elegante molekularbiologische Technik der Protein-affinity tags. Dabei wird der kodierenden DNA eines Proteins die DNA eines anderen Proteins oder Peptids angehängt, welches hohe Affinität zu einer bestimmten festen Matrix aufweist. Das so entstandene Fusionsprotein bindet einerseits an die spezifische feste Matrix und andererseits alle Interaktionspartner des ursprünglichen Proteins (Abbildung 2). Auf diesem Weg lassen sich relativ einfach und allgemein übertragbar Proteinkomplexe isolieren und anschließend analysieren. Wichtig ist hierbei nicht nur die Ausbeute, sondern auch die Reinheit des isolierten Komplexes. Deshalb wird diese Technik weiterentwickelt, z.B. durch die Kombination von zwei verschiedenen Protein-affinity tags, was eine Zweistufenreinigung erlaubt.³³⁾

Durch systematische Markierung von Proteinen kann das Netzwerk der Protein-Protein-Interaktionen aufgedeckt und können sogar neue Proteinkomplexe aufgespürt werden, deren nachfolgende biologische Untersuchung neue Funktionen in der Zelle offen legen kann. Eine solche

Studie des Protein-Interaktionsnetzes wurde gerade in Bäckerhefe durchgeführt,^{34,35)} wobei es interessanterweise nur zu einer geringen Überschneidung mit den Ergebnissen vorangegangener Untersuchungen (mit anderen Techniken) kam.³⁶⁾ Dies unterstreicht die Komplexität der Protein-Protein-Wechselwirkungen.

Ein nächster Schritt bei der Charakterisierung der Protein-Protein-Wechselwirkungen ist die Analyse der Nachbarschaftsverhältnisse und der Kontaktflächen in den Proteinkomplexen. Auch diese Untersuchung kann massenspektrometrisch angegangen werden, benötigt allerdings größere Mengen an gereinigtem Proteinkomplex als zur Proteinidentifizierung notwendig sind. Zudem befindet sich die Technik weitestgehend noch im Entwicklungsstadium.

Nicht-kovalente Komplexe wie Ribosomen wurden im Massenspektrometer vermessen, wobei die partielle Dissoziation von Komplexen bei der Messung Aussagen über ihre Stabilität und Subkomplexe erlaubt.³⁷⁾ Chemisches Quervernetzen kann nicht-kovalente Interaktionen in kovalente Bindungen überführen, welche dann unter den Analysebedingungen stabil sind. Damit können die Nachbarschaftsverhältnisse in einem Proteinkomplex untersucht werden.³⁸⁾ Wenn die vernetzten Peptide nach dem Verdau der Proteine identifiziert werden können, sind sogar Aussagen über die Kontaktbereiche der Proteine möglich.³⁹⁾ Kontaktflächen lassen sich auch durch den Austausch von aciden Protonen gegen Deuteronen untersuchen, wobei bestimmte Bereiche durch Proteinbindung maskiert werden.⁴⁰⁾

Die Funktionszuordnung von Proteinen erweitert das Verständnis von Mechanismen in der Zelle. Das ist von großem Interesse für die Grundlagenforschung. Aber auch die Pharmaforschung ist stark abhängig von dieser Information. Es geht dabei um die Auswahl von Arzneimitteltargets: Ausgehend von den Zehntausenden neu entdeckten menschlichen und pathogenen Genen sollen Proteine gefunden werden, welche in medizinisch relevante Zellprozesse

involviert sind. Außerdem kann ein Bindungspartner eines bereits erfolgreich bearbeiteten Arzneimitteltargets günstigere Eigenschaften aufweisen. Schließlich ist es möglich, das sich die Zusammensetzung von Multiproteinkomplexen krankheitsbedingt ändert und diese Feststellung zum Verständnis der Pathogenese beiträgt. Auf diesen Gebieten liegen wohl die Schwerpunkte von massenspektrometrischen Proteinanalysen und Protein-Protein-Interaktionsstudien.

Die Bedeutung der Proteomanalyse wird international unterstrichen durch eine Reihe wohlfinanzierter Firmengründungen. Sie setzen den Trend fort, der sich bei der Genomanalyse abgezeichnet hat: die zunehmende Bedeutung der Industrie gegenüber der akademischen Forschung bei biologischen Großprojekten.

Juri Rappsilber, Matthias Mann
Protein Interaction Laboratory
University of Southern Denmark, Odense
Rappsilber@PIL.SDU.DK
Mann@PIL.SDU.DK

- 1) M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, *Int. J. Mass Spectrom. Ion Proc.* 1987, 78, 53–68.
- 2) M. Karas, F. Hillenkamp, *Anal. Chem.* 1988, 60, 2299–3201.
- 3) J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science* 1989, 246, 64–71.
- 4) M. Karas, M. Glückmann, J. Schäfer, J. Mass Spectrom. 2000, 35, 1–12.
- 5) R. Knochenmuss, A. Stortelder, K. Breuker, R. Zenobi, *J. Mass Spectrom.* 2000, 35, 1237–1245.
- 6) P. Kebarle, *J. Mass Spectrom.* 2000, 35, 804–817.
- 7) W. J. Henzel, T. M. Billeci, J. T. Stults, S. C. Wong, *Proc. Natl. Acad. Sci. USA* 1993, 90, 5011–5015.
- 8) P. James, M. Quadroni, E. Carafoli, G. Gonnet, *Biochem. Biophys. Res. Commun.* 1993 195, 58–64.
- 9) M. Mann, P. Højrup, P. Roepstorff, *Biol. Mass Spectrom.* 1993, 22, 338–345.
- 10) D. J. C. Pappin, P. Højrup, A. J. Bleasby, *Curr. Biol.* 1993, 3, 327–332.
- 11) J. R. Yates, S. Speicher, P. R. Griffin, T. Hunkapiller, *Anal. Biochem.* 1993, 214, 397–408.
- 12) M. Wilm, A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis, M. Mann, *Nature* 1996, 379, 466–469.
- 13) M. Mann, M. Wilm, *Anal. Chem.* 1994, 66, 4390–4399.
- 14) J. K. Eng, A. L. McCormack, I. J. R. Yates, *J. Am. Soc. Mass Spectrom.* 1994, 5, 976–989.
- 15) J. Qin, D. Fenyó, Y. Zhao, W. W. Hall, D. M. Chao, C. J. Wilson, R. A. Young, B. T. Chait, *Anal. Chem.* 1997, 69, 3995–4001.
- 16) D. N. Perkins, D. J. Pappin, D. M. Creasy, J. S. Cottrell, *Electrophoresis* 1999 20, 3551–3567.
- 17) M. Kussmann, U. Lassing, C. A. Sturmer, M. Przybylski, P. Roepstorff, *J. Mass Spectrom.* 1997, 32, 483–493.
- 18) M. Schürenberg, C. Lübbert, H. Eickhoff, M. Kalkum, H. Lehrach, E. Nordhoff, *Anal. Chem.* 2000, 72, 3436–3442.
- 19) W. P. Blackstock, *Trends Biotechnol.* 2000, 18 (Suppl.), 12–17.
- 20) M. G. Reese, G. Hartzell, N. L. Harris, U. Ohler, J. F. Abril, S. E. Lewis, *Genome Res.* 2000, 10, 483–501.
- 21) M. Wilm, M. Mann, *Anal. Chem.* 1996, 68, 1–8.
- 22) J. S. de Wit, C. E. Parker, K. B. Tomer, J. W. Jørgenson, *Anal. Chem.* 1987, 59, 2400–2404.
- 23) D. F. Hunt, R. A. Henderson, J. Shabanowitz, K. Sakaguchi, H. Michel, N. Sevilir, A. L. Cox, E. Appella and V. H. Engelhard, *Science* 1992, 255, 1261–1263.
- 24) M. P. Washburn, D. Wolters, J. R. Yates, 3rd, *Nature Biotechnol.* 2001 19, 242–247.
- 25) Y. Oda, K. Huang, F. R. Cross, D. Cowburn, B. T. Chait, *Proc. Natl. Acad. Sci. USA* 1999, 96, 6591–6596.
- 26) H. W. Lahm, H. Langen, *Electrophoresis* 2000, 21, 2105–2114.
- 27) S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, R. Aebersold, *Nature Biotechnol.* 1999, 17, 994–999.
- 28) J. Rappsilber, M. Mann, *Trends Biol. Sci.* 2002, im Druck.
- 29) M. J. Huddleston, R. S. Annan, M. T. Bean, S. A. Carr, *J. Am. Soc. Mass Spectrom.* 1993, 4, 710–717.
- 30) M. Wilm, G. Neubauer, M. Mann, *Anal. Chem.* 1996, 68, 527–533.
- 31) H. Steen, B. Küster, M. Fernandez, A. Pandey, M. Mann, *Anal. Chem.* 2001, 73, 1440–1448.
- 32) A. I. Lamond, M. Mann, *Trends Cell Biol.* 1997, 7, 139–142.
- 33) G. Rigaut, A. Shevchenko, B. Rutz, M. Wilm, M. Mann, B. Seraphin, *Nature Biotechnol.* 1999, 17, 1030–1032.
- 34) Y. Ho et al., *Nature* 2002, 415, 180–183.
- 35) A.-C. Gavin et al., *Nature* 2002, 415, 141–147.
- 36) A. J. Walhout, S. J. Boulton, M. Vidal, *Yeast* 2000, 17, 88–94.
- 37) A. A. Rostom, P. Fucini, D. R. Benjamin, R. Juenemann, K. H. Nierhaus, F. U. Hartl, C. M. Dobson, C. V. Robinson, *Proc. Natl. Acad. Sci. USA* 2000, 97, 5185–5190.
- 38) J. Rappsilber, S. Siniosoglou, E. C. Hurt, M. Mann, *Anal. Chem.* 2000, 72, 267–275.
- 39) J. Vater, K. Heinze, B. Friedrich, B. Kalbitz, A. Blokesch, K.-D. Irrgang, B. Thiede, J. Salnikow, *Ber. Bunsenges. Phys. Chem.* 1996, 100, 2107–2111.
- 40) J. G. Mandell, A. Baerga-Ortiz, S. Akashi, K. Takio, E. A. Komives, *J. Mol. Biol.* 2001, 306, 575–589.