

# Ein Genom – verschiedene Proteome



Nicht nur das Genom, sondern vor allem die jeweils vorhandene Proteinausstattung – das Proteom – bestimmen das Aussehen und den Zustand eines biologischen Organismus.



# Proteomanalyse – ein Weg zur Funktionsanalyse von Proteinen

Friedrich Lottspeich\*

Genomics – die Sequenzierung ganzer Genome – wird mit ständig steigender Geschwindigkeit vorangetrieben. Viele Projekte sind begonnen, 20 bereits abgeschlossen. In wenigen Jahren wird auch das Genom des Menschen vollständig sequenziert sein. Gerade wurde ein neuer Meilenstein mit der Entschlüsselung des Genoms des Rundwurmes (*Caenorhabditis elegans*) gesetzt, das mit 97 Millionen Basenpaaren das bisher größte sequenzierte Genom ist. Von seiner Analyse verspricht man sich Aufschluß über die rund 19000 gefundenen Gene. Doch von einer „Entschlüsselung“ dieser Daten im eigentlichen Sinn des Wortes ist man noch sehr weit entfernt. Was können wir aus den Genomdaten lernen? Was sind diese Daten wert? Kein Zweifel, die Genomdaten und deren Analyse haben die ungeheure Komplexität der Natur erst richtig aufgedeckt. An die Stelle des Dogmas „ein Gen – ein Protein – eine Funktion“ trat ein Verständnis von verschachtelten Regulationsnetzwerken. Wir haben gelernt, mit großen Datenmengen umzugehen. Datenbankstrukturen wurden aufgebaut, hochspezialisierte Pro-

gramme zum Data-mining sind oder werden entwickelt. Auch die klassische Proteinchemie profitierte von den Genomprojekten. Nun muß nicht mehr jede Aminosäure eines großen Proteins analysiert werden – in vielen Fällen ein hoffnungsloses Unterfangen –, sondern es genügt, kleine Bereiche eines Proteins zu bestimmen und daraus durch Vergleich mit bekannten Sequenzen in Datenbanken auf eine Verwandtschaft mit anderen Proteinen zu schließen. Aber aus der DNA-Sequenz kann nicht alles abgeleitet werden. Was ist die Funktion der Genprodukte, wie sehen die aktiven Proteine aus? Die Antworten darauf sind oft nicht in der Primärsequenz enthalten. Wie sieht es auf der Ebene der mRNA aus? Neueste Entwicklungen bei cDNA-Chips lassen große Hoffnung aufkommen, Veränderungen der mRNA schnell und kostengünstig analysieren zu können. Aber kann die Menge der mRNA auch die komplexen Zusammenhänge charakterisieren, die eine bestimmte Stoffwechselsituation oder einen pathologischen Status ausmachen? Zum Teil sicher; so wird sich in einigen Fällen ein Befund mit

einem veränderten mRNA-Muster korrelieren lassen. Aber immer dann, wenn posttranslationale Modifikationen, Interaktionen, Abbau- und Transportphänomene die Funktionen eines Proteins bestimmen, kann die mRNA keine Aussage mehr liefern. Hier hilft nur, sich auf die Ebene der Proteine zu begeben und deren Art, Modifikationen und vor allem deren Konzentration (Menge pro Kompartiment, Zelle etc.) direkt zu untersuchen. Proteomics – die quantitative Analyse der zu einem bestimmten Zeitpunkt und unter bestimmten Bedingungen in einem Organismus vorhandenen Proteine, ist ein Schlüssel zur Funktionsanalyse. Proteomics wird in nächster Zeit sowohl die Grundlagenforschung, z.B. bei der Aufklärung von Reaktions- und Regulationsnetzwerken, als auch die angewandte Forschung, z.B. bei der Suche und Auswahl von Targets zur Entwicklung von Medikamenten, bestimmen.

**Stichwörter:** Analytische Methoden • Elektrophorese • Massenspektrometrie • Proteine • Proteomforschung

## 1. Einleitung

Die weltweit durchgeführten Genomprojekte haben die Sequenzierung und die nachfolgende funktionelle Analyse von gesamten Genomen zur Zielsetzung. Die Sequenzanalysen mehrerer Genome einfacher Organismen, wie Bakterien

(*Escherichia coli*, *Bacillus subtilis*, *Helicobacter pylori*, *Haemophilus influenzae* u.v.a.) und auch der Hefe *Saccharomyces cerevisiae* als Vertreter der Eukaryonten, sind schon fertiggestellt.<sup>[1]</sup> Auch große Genome, wie das des Menschen oder einer Pflanze, *Arabidopsis thaliana*, wurden in weltweit konzertierten Aktionen in Angriff genommen, wobei das Humangenomprojekt voraussichtlich in wenigen Jahren abgeschlossen sein wird. Aber das vollständige Genom gibt nur einen relativ statischen Überblick über das funktionelle Potential eines Organismus und beschreibt nicht das ungeheuer dynamische Geschehen, das in jedem lebenden Orga-

[\*] Dr. F. Lottspeich  
Max-Planck-Institut für Biochemie  
D-82152 Martinsried  
Fax: (+49) 89-85-78-28-02  
E-mail: lottspei@biochem.mpg.de

nismus stattfindet. Beispielsweise enthält jede somatische Zelle des in Abbildung 1 gezeigten Schmetterlings und seiner Raupe die identische genetische Information. Die Umsetzung dieser genetischen Information, d.h. die Expression der verschiedenen Gene in Proteine, läuft aber in den verschiedenen Entwicklungsstadien eines Organismus sowie in verschiedenen Zelltypen und unter verschiedenen Umweltbedingungen unterschiedlich ab. Dies erst führt zur enormen individuellen phänotypischen Vielfalt der Natur.

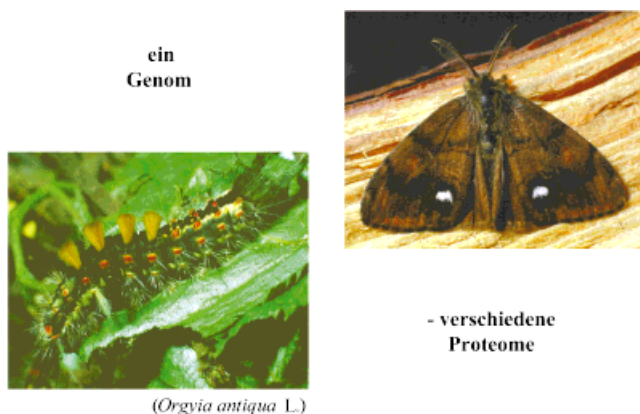


Abbildung 1. Raupe und Falter von *Orgyia antiqua* L. (mit freundlicher Genehmigung von TOPLAB GmbH, Martinsried).

Bei der Umsetzung der genetischen Information in Proteine – die Werkzeuge und Werkstoffe einer Zelle – gibt es Regelmechanismen, welche die Mengenverhältnisse der einzelnen Proteine zueinander äußerst genau einstellen. Kleinste Störungen dieser feinjustierten Proteinexpression können zu erheblichen biologischen Konsequenzen führen. Daher hat die Natur ein komplexes (und sehr robustes) Regelnetzwerk entwickelt, das auf allen Ebenen der Informationsumsetzung – von der Transkription über die Translation bis hin zur Ebene der Proteine – mit vielfältigen Rückkopplungen und Kontrollpunkten ausgestattet ist. So wird die Expression der Proteine durch Transkriptions- und Translationsfaktoren geregelt, die selbst Proteine sind und daher ihrerseits wieder gleichen Regulationsmechanismen von Synthese und Proteinabbau unterworfen sind. Zusätzlich wird das Regulationsnetzwerk noch viel komplizierter, da durch die posttranslationalen Modifikationen (z. B. Phosphorylierung, Glycosylierung, Pro-

zessierung etc.) Proteine strukturell verändert werden und damit in ihrer biologischen Aktivität und Funktion moduliert werden können. Die Kenntnis der posttranslationalen Modifikationen eines Proteins ist daher ein wesentlicher Faktor für das Verstehen seiner Wirkungsweise und somit für eine Funktionsanalyse unerlässlich. Diese endogenen molekularen Zusammenhänge werden ihrerseits von exogenen Parametern wie Temperatur, Kulturbedingungen, Streß etc. in erheblichem Maße beeinflusst.

Alle diese komplex vernetzten Vorgänge sind sowohl räumlich als auch zeitlich genau geregelt. Es hat sich gezeigt, daß selbst aufwendige molekularbiologische Methoden zur Funktionsanalyse einzelner Gene (z. B. „gene-disruption“, „Knock-out“-Techniken) sehr oft keine eindeutigen oder auf molekularer Ebene brauchbaren Resultate liefern. Auch aus der Menge einer mRNA kann man im allgemeinen nicht auf die Menge des entsprechenden aktiven Proteins schließen. Daher besteht ein großer Bedarf an komplementären, auf Proteine ausgerichteten Strategien, die – gemeinsam mit molekularbiologischen Ansätzen – das Problem der umfassenden Funktionsanalyse von Genen und Genprodukten bewältigen können.

## 2. Von der Proteinanalytik zur Proteomanalyse

Noch bis 1950 war unklar, ob ein bestimmtes Protein eine definierte und eindeutige kovalente Struktur hat oder aus einem sehr heterogenen Gemisch von Aminosäurepolymerketten besteht. Erst die Arbeiten von Pehr Edman und Frederic Sanger zur Sequenzanalyse von Proteinen und Peptiden bewiesen, daß ein bestimmtes Protein eine eindeutige und einheitliche Struktur hat.<sup>[2]</sup> Die folgenden drei Jahrzehnte waren von dem Bestreben geprägt, einzelnen Proteinen bestimmte biologische Funktionen zuzuordnen. Fast immer wurde ein Protein anhand seiner biochemischen Aktivität in reiner Form isoliert, bevor seine kovalente Struktur (Aminosäuresequenz und Modifikationen) detailliert aufgeklärt werden konnte. Waren diese Informationen bekannt, suchte man nach möglichen Interaktionspartnern und analysierte diese im Detail, was dann wiederum als Ausgangspunkt für die Suche nach weiteren interagierenden Molekülen diente. Der Nachteil dieser einsichtigen, funktionsbasierten und sehr erfolgreichen Strategie war vor allem



Friedrich Lottspeich, geboren 1947, studierte Chemie an der philosophischen Fakultät der Universität Wien. 1978 promovierte er bei Pehr Edman am Max-Planck-Institut für Biochemie in Martinsried über die Primärstruktur von Fibrinogen. Nachdem er am selben Institut als wissenschaftlicher Mitarbeiter tätig war, wurde er 1984 zum Leiter der selbstständigen Arbeitsgruppe „Mikrosequenzierung“ am Genzentrum der Universität München berufen. Nach seinen Habilitationen an den Universitäten München und Innsbruck nahm er 1990 einen Ruf als Leiter der Proteinanalytik am Max-Planck-Institut für Biochemie in Martinsried an. Er ist Autor von über 550 Originalpublikationen und Herausgeber mehrerer wissenschaftlicher Bücher. Seine Forschungsschwerpunkte sind methodische und anwendungsorientierte Arbeiten zur Proteinstrukturaufklärung und zur Proteomanalyse.

der relativ große Materialverbrauch und die Langwierigkeit der Isolierungsarbeiten, bei denen immer auf die Erhaltung der biologischen Aktivität Rücksicht genommen werden mußte.

Die Methoden der Proteomanalytik wurden kontinuierlich weiterentwickelt und in ihrer Empfindlichkeit erheblich gesteigert. Ein wichtiger Meilenstein dieses methodischen Fortschrittes war die Entwicklung der zweidimensionalen (2D) Gelelektrophorese im Jahr 1975.<sup>[3]</sup> Schon damals wurde das große Potential dieser hochauflösenden Trennmethode erkannt, die eine Vielzahl einzelner Proteine selbst aus so komplexen Gemischen wie Geweben, Zellen oder Körperflüssigkeiten trennen konnte. Es wurden Proteinmuster von normalen und pathologischen Zuständen verglichen, jedoch hatten zu dieser Zeit die gefundenen Unterschiede ausschließlich diagnostischen Wert, da die identifizierten Proteine wegen der damals fehlenden oder zu unempfindlichen Analysemethoden nicht weiter charakterisiert werden konnten. Erst die Ausarbeitung von speziellen Probenvorbereitungstechniken<sup>[4]</sup> sowie die Weiterentwicklung der Proteinsequenzierung<sup>[5]</sup> ermöglichten ab Mitte der 80er Jahre eine Analyse von Proteinen, die durch 2D-Gelelektrophorese getrennt worden waren. Allerdings geriet die Proteinchemie Anfang der 80er Jahre durch die spektakulären Erfolge und das hohe Entwicklungspotential der Gentechnik deutlich aus dem Fokus des wissenschaftlichen Interesses. Molekularbiologische Ansätze zur Lösung biologischer Fragestellungen wurden bevorzugt und führten auch oft in erstaunlich kurzer Zeit zum Erfolg. Immer neue gentechnische Methoden wurden entwickelt und verbessert, und ihre Anwendung erreichte in den Projekten zur Sequenzierung vollständiger Genome einen bisherigen Höhepunkt.<sup>[1]</sup>

Es wurde ein Bewußtsein für die Komplexität von natürlichen Regulationsnetzwerken geschaffen, und der Umgang mit einer großen Zahl von komplexen Proben in der Molekularbiologie erzeugte einen hohen Entwicklungsdruck zur Erarbeitung automatisierter und schneller Analysetechniken. Diese „High-throughput“-Methoden produzierten eine große Menge an Daten, was wiederum die Entwicklung von Datenbanken und der dazugehörigen Software stimulierte und die Geburtsstunde der Bioinformatik bedeutete.

Gegen Ende der 80er Jahre wurden aber immer mehr Stimmen laut, welche die Meinung vertraten, daß mit molekularbiologischen Methoden allein die Vielfalt und Komplexität von biologischen Vorgängen wohl nicht erklärt werden könne. Es wurde klar, daß das mRNA-Muster zwar in vielen Fällen Aufschluß über an- oder abgeschaltete Gene und Genfamilien geben kann, daß aber die Menge einer mRNA keine Schlüsse auf die Menge des entsprechenden aktiven Proteins zuläßt. Sowohl auf mRNA-Ebene wie auch auf Proteinebene zerstören nicht kalkulierbare Vorgänge (mRNA-Abbau, Translationskontrolle, Proteinabbau und posttranslationale Modifikationen) die strenge Korrelation zwischen RNA- und Proteinmenge.<sup>[6]</sup> Es setzte sich immer mehr die Ansicht durch, daß man für ein vollständigeres Bild von biologischen Abläufen die Proteine miteinbeziehen müsse, da diese die in der Zelle wirkenden Akteure sind. Gerade zu diesem Zeitpunkt konnten auch zwei neue massenspektrometrische Techniken, die Elektrospray-Ionisa-

tions-Massenspektrometrie (ESI-MS)<sup>[7]</sup> und die matrix-assistierte Laserdesorptions/Ionisations-Massenspektrometrie<sup>[8]</sup> (MALDI-MS) ihre ersten Erfolge in der Proteomanalytik feiern. Mit der Verbesserung dieser MS-Techniken und den enorm schnell anwachsenden Sequenzdaten aus den Genomprojekten war die Grundlage für eine vollständig neue, empfindliche und vergleichsweise schnelle Proteomanalytik geschaffen, mit der man in der Lage war, eine große Zahl von Proteinen zu untersuchen. Jetzt wurde die alte Idee der differentiellen Analyse von komplexen Proteingemischen von neuem aufgegriffen, die heute in der Proteomanalyse ihre Anwendung findet.

Der Begriff Proteom wurde 1994 auf dem 2D-Gelelektrophorese-Treffen in Siena zu ersten Mal als „das Proteinäquivalent zu einem Genom“ von Marc Wilkins gebraucht und wurde wegen seiner „Griffigkeit“ und Analogie zum Ausdruck „Genom“ schnell akzeptiert. In der Folgezeit wurde diese Definition etwas schärfer gefaßt, da klar wurde, daß der bloße Nachweis und die Auflistung aller möglichen Proteine und Proteinvarianten, die von einem Genom produziert werden können, relativ wenig Informationswert hat. Auch die bloße Positionierung jedes möglichen Proteins in einem gegebenen Trennraum – also die reine Kartierung – wird aus verschiedenen Gründen trotz eines enormen Aufwandes nur sehr wenig praktischen Nutzen bringen:

- Für die biologische Aktivität und Wirkung eines Proteins sind seine Menge und seine posttranslationalen Modifikationen von entscheidender Bedeutung;
- es existiert kein biologischer Zustand, in dem alle im Genom codierten Proteine eines Organismus exprimiert sind;
- auch bei sehr guten Trennsystemen kommen oft mehrere Proteine in einer Position zu liegen, so daß im aktuellen Experiment doch wieder die Identität des interessierenden Proteins analytisch überprüft werden muß;
- technische Probleme der Proteomanalyse und kleine Unterschiede in der Probenvorbereitung und den Analysenprotokollen verhindern die Vergleichbarkeit von „Proteomkarten“ aus verschiedenen Laboratorien.

Aus diesen Gründen setzen sich heute immer mehr die im Folgenden aufgeführten Ansichten zu Proteomexperimenten durch:

Ein Proteom ist „das zu einem bestimmten Zeitpunkt und unter exakt definierten Randbedingungen quantitativ ermittelte Proteinmuster eines Organismus, einer Zelle, einer Organelle oder auch einer Körperflüssigkeit“.

Das Proteom reflektiert einen aktuellen Stoffwechselzustand der entsprechenden Zelle (oder des entsprechenden Organismus), der durch heute noch schwer analysierbare, vielfältige Interaktionen der verschiedenen Moleküle und unzählige Umgebungsparameter bestimmt wird. Das Proteom ist somit – im Unterschied zum Genom – ein hochdynamisches System, das durch kleine Änderung der „Umweltbedingungen“ (z. B. die Änderung der Kulturbedingungen für einen zur Produktion verwendeten Bakterienstamm oder die Zugabe eines Medikamentes zu einer Zellkultur etc.) charakteristisch verändert wird.

Eine Proteomanalyse ist nur im Zusammenhang mit einem subtraktiven Ansatz (Abbildung 2) sinnvoll, bei dem zwei

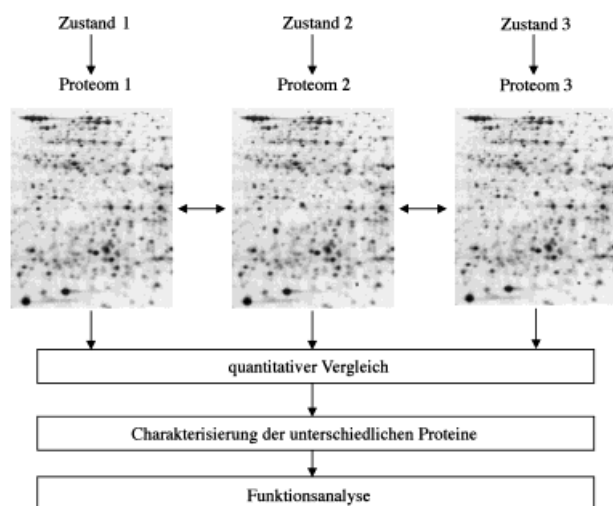


Abbildung 2. Schematische Darstellung des subtraktiven Ansatzes. Über den Vergleich von zwei Proteinmustern werden Unterschiede erkannt, die für die einzelnen Ausgangszustände charakteristisch sind.

oder mehrere wohldefinierte Zustände verglichen werden können. In den Proteinmustern dieser verschiedenen Zustände (z. B. einer Zelle vor und nach Zugabe eines Medikaments, oder beim Vergleich von Zellen aus normalen und pathologischen Zuständen etc.) werden Veränderungen einzelner Proteine beobachtet und quantitativ ausgewertet. Daher kann das Proteom als ein einzigartiger, hochempfindlicher Monitor für komplexe metabolische und regulatorische Zusammenhänge eines Organismus gesehen werden. Aus den gefundenen Veränderungen, die meist viele Proteine betreffen, sollen Rückschlüsse auf die beteiligten funktionellen Netzwerke (Stoffwechselwege, Regulationskaskaden etc.) gezogen werden und über eine Korrelation mit dem Phänotyp eine funktionelle Aussage erreicht werden. Man kann sich die Situation ähnlich wie bei einem Film vorstellen, der zwar auch aus einer Folge von statischen Einzelbildern besteht, aber einen wesentlichen Informationsgehalt aus der Dynamik der Veränderungen zwischen den einzelnen Bildern erhält. Die Standbilder eines Films sind dennoch in der Lage, einen Eindruck vom Filmgeschehen zu vermitteln, und dies um so besser, je sorgfältiger und charakteristischer die Auswahl der Szenen für die Standbilder erfolgte. So kann auch bei der Proteomanalyse über die Betrachtung der Veränderungen zwischen unterschiedlichen (per se statischen) Momentaufnahmen der quantitativen Proteinzusammensetzung eine sinnvolle Aussage erzielt werden. Wir stehen allerdings heute erst am Beginn eines Verstehens der „Bilder“ des Proteomfilms. Wir müssen lernen, die beobachteten Differenzen der Proteinmuster mit den biologischen Auswirkungen zu korrelieren. Je besser die Auswahl der einzelnen Proteomzustände ist, um so klarer wird die Abfolge und Art der Veränderungen interpretiert werden können.

Die Proteomanalyse kann so eine neue Qualität von Antworten auf biologische Fragen geben, die zur Zeit mit keiner anderen Technik erhalten werden können. Nur das Zusammenspiel der Proteomanalyse mit den molekularbiologisch-gentechnischen Methoden der Genomanalyse wird uns einen besseren Einblick in die komplexen funktionellen Regulations- und Stoffwechselnetzwerke der Natur eröffnen.

### 3. Methoden der Proteomanalyse

Die Proteomanalyse weist gegenüber der klassischen Proteinanalytik den wesentlichen Vorteil auf, daß eine biologisch relevante Aussage unabhängig von der biologischen Aktivität des einzelnen Proteins erhalten wird. Es muß also bei der Trennung keine Rücksicht auf die Proteinaktivität genommen werden, was den Einsatz denaturierender Trennbedingungen erlaubt und damit wesentlich schnellere und effizientere Analysen ermöglicht. Allerdings ist die systematische und umfassende Analyse von Proteinen viel schwieriger als die Analyse von Genen, da – im Unterschied zu RNA oder DNA – bei Proteinen zum einen keine Amplifizierungsmöglichkeiten vorhanden sind (daraus folgt eine limitierte Sensitivität) und zum anderen die physikalisch-chemischen und biochemischen Eigenschaften von Proteinen äußerst unterschiedlich sind. Deshalb mußten neue Methoden entwickelt werden, um effizient mit komplexen Proteingemischen umgehen zu können, die mehr als 10000 verschiedene Proteinspezies enthalten können. Weltweit wurden die Methoden der Proteomanalyse vor allem in den universitären Forschungslabors entwickelt. In den einzelnen Labors wurden allerdings wegen limitierter Ressourcen jeweils nur einzelne Aspekte der Proteomanalytik bearbeitet. So entstand die heutige Situation, daß zwar alle Methoden zur Analyse eines Proteoms prinzipiell existieren, aber die Expertise für die Gesamtheit der dazu notwendigen Techniken nur an sehr wenigen Stellen der Welt in einer Gruppe vorhanden ist (Probenvorbereitung, 2D-Gelelektrophorese, enzymatische und chemische Spaltungen, Sequenzanalyse, Massenspektrometrie, Bioinformatik etc.).

Eine Proteomanalyse kann man in einzelne Schritte gliedern:

- Definition der Fragestellung und exakte Festlegung der Ausgangsbedingungen für die Proteomanalyse (oder für die einzelnen Zustände eines subtraktiven Ansatzes)
- Probenvorbereitung
- Trennung der Proteine
- Quantifizierung der Proteine
- Datenbankanalyse der Proteinmuster mit Methoden der Bioinformatik
- proteinchemische Charakterisierung der veränderten Proteine und Analyse der posttranslationalen Modifikationen

#### 3.1. Definition der Fragestellung und Festlegung der Ausgangsbedingungen für die Proteomanalyse

Die Methodik und damit auch der Aufwand und die Kosten einer Proteomanalyse werden wesentlich von der Fragestellung mitbestimmt: Eine Proteomanalyse im „puristischen“ Sinn soll über die Beschreibung der Unterschiede zwischen zwei oder mehreren Momentaufnahmen der Proteinmuster letztendlich Aussagen über funktionelle Netzwerke und/oder Beteiligungen von bestimmten Proteinen an einzelnen Reaktions- und Regelmechanismen treffen können. Die Einsatzmöglichkeiten sind beliebig vielfältig, so daß nur einige stellvertretend und stichpunktartig genannt werden sollen:

- Welche Proteine sind an Reaktionsnetzwerken oder biologischen Mechanismen beteiligt? Diese Information kann

als Grundlage für die Auswahl von Targetproteinen oder Markerproteinen dienen.<sup>[6b, 9, 10]</sup>

- Wie beeinflussen chemische Substanzen und Umweltbedingungen die Proteinexpression, und wie nehmen sie damit über pharmakologische und toxikologische Wirkungen Einfluß auf Differenzierung, Proliferation und Stoffwechsel? Was unterscheidet beispielsweise diesbezüglich Medikamente ohne und solche mit Nebenwirkungen? Welche Proteine oder welche Proteinkonstellationen sind für Nebenwirkungen charakteristisch (siehe auch Abbildung 6)?<sup>[6b, 10]</sup>
- Was sind die molekularen Grundlagen für effiziente Produktionsstämme in der Mikrobiologie? Kann die Expression von Proteinen auf rationaler Basis so beeinflusst werden, daß höhere Produktionsausbeuten oder Produktreinheiten erzielt werden können?<sup>[10b]</sup>

Ganz generell gilt, daß die Schwierigkeit, aus Veränderungen funktionelle Schlüsse ziehen zu können, mit der Zahl der veränderten Proteine stark zunimmt. Daher sollten die Unterschiede zwischen den einzelnen Zuständen einer Proteomanalyse nicht zu groß gewählt werden, um noch eine sinnvolle Aussage treffen zu können.

Vor einer Proteomanalyse sollte eine gründliche Überlegung dem Aufwand gelten. So ist aus technischen Gründen – sowohl bei der Trennung als auch bei der Detektion – eine vollständige quantitative Erfassung aller exprimierten Proteine (noch?) nicht möglich. Hydrophobe Proteine, sehr große, sehr kleine, sehr saure und sehr basische Proteine bereiten große Trennprobleme, die es verhindern, auch bei sonst optimalen Bedingungen ein vollständiges Proteom zu erstellen. Man geht heute davon aus, daß man von Organismen mit einem kleinem Genom, z. B. Hefe, deutlich mehr als 50 % der exprimierten Proteine quantitativ erfassen und analysieren kann. Da man heute einige Milligramm Proteinmaterial auf ein zweidimensionales Gel laden kann,<sup>[11]</sup> sind Proteine, die in > 100 000 Kopien pro Zelle vorkommen, ohne größere Probleme direkt nachzuweisen, zu quantifizieren und auch proteinchemisch zu charakterisieren. Die meisten Proteine sind allerdings weit weniger stark exprimiert und müssen über zusätzliche Schritte angereichert werden (siehe Abschnitt 3.2), was einen erheblichen Mehraufwand bedeutet. Daher sollten bei der Ausarbeitung der Fragestellung Überlegungen angestellt werden, wie weit eine (annähernd) vollständige Proteomanalyse überhaupt notwendig und sinnvoll (oder auch machbar) ist, oder ob die Untersuchungen nicht auf eine bestimmte Gruppe von Proteinen beschränkt werden können. Dies kann durch eine biologische Fraktionierung (z. B. von bestimmten Organellen) erreicht werden, kann aber auch eine physikalische Fraktionierung (z. B. eine Fällung) beinhalten. Wichtig ist aber in allen Fällen, daß das Ausgangsmaterial für die eigentliche Proteomanalyse reproduzierbar hergestellt werden muß.

Eine wichtiges Einsatzgebiet für die Proteomanalyse ist das Erkennen von Proteinen, die mit einem bestimmten Zustand (z. B. einer Krankheit) korreliert sind (z. B. diagnostische und therapeutische Marker). Dazu werden z. B. die Proteinmuster von gesunden und pathologischen Zellen, Geweben, oder Körperflüssigkeiten verglichen. Auch wenn die einzelnen Zustände der Proteomanalyse hier biologisch weit auseinander

liegen und nicht erwartet werden kann, aus den Veränderungen der Proteinmuster direkt komplexe Funktionszusammenhänge erklären zu können, wird das Ziel, wichtige, diagnostisch und therapeutisch interessante Proteine zu erkennen, häufig erreicht. Diese dienen dann im allgemeinen als Ausgangspunkt für weiterführende Analysen. Bei der Bewertung der Resultate sollten allerdings folgende Aspekte bedacht werden:

- Es ergeben sich fast immer Probleme bei der Probennahme von „Ex-vivo“-Material, die selten wirklich reproduzierbar durchgeführt werden kann. So sind unter Umständen mehrere Zelltypen in verschiedenen Proliferations- und Differenzierungszuständen vorhanden.
- Die „Normalwerte“ für einzelne Proteine können bei unterschiedlichen Probanden erheblich voneinander abweichen und sind statistisch abzusichern.
- Durch das relativ häufige Auftreten von „Polymorphismen“ wird sich das Proteinmuster auch unter identischen Randbedingungen signifikant unterscheiden. So schätzt man, daß sich die Genome von zwei Menschen in ca. einer Million Basenpaare unterscheiden, wobei sich viele dieser Unterschiede auf der Proteinebene als Aminosäureaustausche manifestieren. Die entstandenen isofunktionellen, aber physikalisch leicht unterschiedlichen Proteine werden bei der Proteomanalyse an unterschiedlichen Positionen gefunden und verändern somit das Proteinmuster.

Alle diese Schwierigkeiten mit der Probennahme und dem Probenmaterial können durch eine große statistische Basis gemildert werden, für die aber manchmal das Material fehlt.

Eine weitere Gruppe von Experimenten wird sehr oft auch der Proteomanalyse zugerechnet, da sie sich der methodisch gleichen Werkzeuge bedient. Es ist dies die Untersuchung von Proteingemischen, wie sie z. B. bei der Suche nach Interaktionspartnern eines Proteins über Immunpräzipitation vorliegen oder bei der Analyse der Zusammensetzung von Proteinkomplexen auftreten. Dabei ist das Ziel, alle Proteine zu charakterisieren, die an ein gegebenes Protein binden oder die Bestandteile eines Proteinkomplexes sind. Auch wenn hier alle „High-throughput“-Methoden der Proteomanalyse eingesetzt werden, ist der wesentliche Unterschied – und dabei eine enorme Erleichterung – zu einer echten Proteomanalyse, daß einerseits die Randbedingungen für die Analyse sehr viel weniger komplex sind und andererseits die genaue quantitative Bestimmung der Menge der vorhandenen Proteine nur eine untergeordnete Rolle spielt (das Fehlen einer absoluten Mengenbestimmung verhindert ohnehin eine einfache Aussage über die stöchiometrische Zusammensetzung von Komplexen).

### 3.2. Probenvorbereitung

Die Probenvorbereitung ist der erste wesentliche Schritt jeder Proteomanalyse. Für eine gute Beschreibung einzelner Proteomzustände müssen möglichst alle experimentellen Parameter erfaßt und konstant gehalten werden, da die Proteinexpression einer Zelle sehr sensitiv auf Veränderungen von äußeren Parametern reagiert. Um diese möglichst gut kontrollieren zu können, werden für Proteomanalysen meist

kultivierte Zellen Verwendung finden müssen, um sowohl die biologische Variabilität zu limitieren als auch die erforderliche Proteinmenge zur Verfügung zu haben. Alle, auch unbeabsichtigte Veränderungen bei der Probenahme werden quantitative und qualitative Abweichungen im Proteinhalt zur Folge haben, die dann wieder durch Mehrfachanalysen und aufwendige statistische Absicherungen kompensiert werden müssen.

Um die quantitativen Verhältnisse richtig beurteilen zu können, muß bei jeder Art von Probenvorbereitung darauf geachtet werden, daß die Proteine vollständig intakt bleiben, d. h., daß sie weder modifiziert noch abgebaut werden. Beim Zellaufschluß werden oft Proteasen freigesetzt, die Proteine sehr schnell abbauen können und damit artifiziell die Heterogenität des Proteingemisches vergrößern und gleichzeitig die natürlich vorhandenen quantitativen Verhältnisse der Proteine zueinander verändern. Das bedeutet, daß die Probenvorbereitung spezifisch für jedes Ausgangsmaterial optimiert werden muß.<sup>[12]</sup> So müssen eventuell Proteaseinhibitoren und/oder hohe Konzentrationen an chaotropen Agentien zur Denaturierung von proteolytischen Enzymen zugegeben werden, oder es muß ein niedriger pH-Wert beim Aufschluß eingestellt werden. In jedem Fall muß schnell und bei niedrigen Temperaturen gearbeitet werden. Relativ einfache Proben, beispielsweise Körperflüssigkeiten oder Zellen ohne extrem stabile Zellmembran, werden am besten direkt im Auftragspuffer für die 2D-Gelelektrophorese gelöst. Bei anspruchsvolleren Proben müssen für jede Probe spezifische Probenvorbereitungsvorschriften ausgearbeitet werden. Das oberste Ziel hierbei muß sein, die Probenvorbereitung möglichst einfach und vollständig, vor allem aber reproduzierbar durchzuführen, wozu standardisierte und weitgehend automatisierte Verfahren ausgearbeitet und eingesetzt werden sollten.

Eine gute Probenvorbereitung kann auch eine Fraktionierung der Probe beinhalten und damit zu einer Reduktion der Komplexität der Probe führen, was für die nachfolgenden Schritte, die Trennung des Proteingemisches und die Quantifizierung der getrennten Proteine, äußerst vorteilhaft ist. Dabei können biologische Vortrennungen wie die Isolierung bestimmter Organellenfraktionen (z. B. Mitochondrien, Membranen, Zellkerne) oder physikalisch-chemische Methoden wie Fällungen, präparative elektrophoretische Verfahren (z. B. Free-flow-Elektrophorese), chromatographische Vorreinigungsschritte etc. zum Einsatz kommen.

Die Probenvorbereitung für eine Proteomanalyse ist vor allem wegen des breiten Spektrums der physikalisch-chemischen Eigenschaften der einzelnen Proteine äußerst schwierig. So werden sich zwangsläufig die Proteine in einem solch komplexen Proteingemisch, wie es das Proteom ist, in ihren Löslichkeiten ganz drastisch unterscheiden. Sie verhalten sich demzufolge ganz unterschiedlich und sollten im Idealfall auch unterschiedlich behandelt werden:

- In Wasser oder verdünnten Puffern leicht lösliche Proteine bereiten wenig Probleme. Hierbei handelt es sich meist um cytosolische Proteine oder Proteine aus Körperflüssigkeiten.
- In wäßrigen Lösungsmitteln schwer lösliche Proteine mit sehr stabilen Sekundär- und Tertiärstrukturen können erst

durch Zugabe chaotroper Verbindungen wie Harnstoff oder Guanidiniumhydrochlorid in Lösung gebracht werden. Dabei besteht die Gefahr von ungewollten, partiellen und nicht kontrollierbaren Modifikationen einzelner Proteine (z. B. durch Carbamoylierung), die dazu führen, daß ein ursprünglich einheitliches Protein in mehreren Formen vorliegt und das Proteingemisch insgesamt noch heterogener wird. Dies kann dann bei der Trennung und Quantifizierung erhebliche Probleme bereiten.

- Proteine, die große, unlösliche Komplexe bilden, wenn sie aus ihrer natürlichen Umgebung entfernt werden, können erst durch chemische Reaktionen, z. B. Reduktion von Disulfidbrücken, in Lösung gebracht werden.
- Besonders schwierig sind Membranproteine zu handhaben, deren natürliche Umgebung Lipidmembranen sind und die bei der Isolierung aus der Membran äußerst leicht aggregieren und damit unlöslich werden. Diese hydrophoben Proteine kann man nur durch Einsatz von Detergentien in Lösung halten, die aber die nachfolgenden Schritte zur effizienten Proteintrennung oft stören oder unmöglich machen.
- Hohe Proteinkonzentrationen, die für die meisten Trenn- und Analyseverfahren notwendig sind, bergen oft die Gefahr der Aggregation und damit der Präzipitation bestimmter Proteine. Andererseits bedingen geringe Proteinkonzentrationen, die für die Löslichkeit von Proteinen allgemein vorteilhaft sind, zusätzliche Schritte vor den Trenn- und Analyseverfahren, die wieder die Gefahr eines Proteinverlustes bergen.

### 3.3. Die Proteintrennung

Für die quantitative Analyse eines Proteoms müssen möglichst hochauflösende Trennmethode eingesetzt werden, die gleichzeitig alle Proteine mit den unterschiedlichsten Eigenschaften auftrennen sollen. Als hochauflösende Methoden im Molekülmassenbereich der Proteine stehen nur elektrophoretische und chromatographische Techniken zur Verfügung. Jedoch kann keine einzige dieser Trenntechniken in einer Analyse wesentlich mehr als ca. 100 Komponenten voneinander trennen. Da aber auch eine einfache Zelle wahrscheinlich wenigstens 10000 Proteinspezies enthält und zusätzlich die Mengen der in einer Probe vorhandenen Proteine sich um den Faktor  $10^6$  und mehr unterscheiden können, muß ein ausreichend großer Trennraum vorhanden sein, was nur durch gekoppelte – mehrdimensionale – Trennverfahren erreicht werden kann.

#### 3.3.1. Elektrophoretische Techniken

Proteine weisen zwitterionischen Charakter auf und sind je nach Pufferbedingungen positiv oder negativ geladen. Elektrophoretische Verfahren trennen Substanzen nach ihrer Mobilität in einem elektrischen Feld, wobei die elektrophoretische Mobilität eine für jedes Protein charakteristische Größe ist. Für die Proteomanalysen werden im wesentlichen zwei hochauflösende elektrophoretische Techniken eingesetzt, die isoelektrische Fokussierung (IEF) und die Natrium-

dodecylsulfat(SDS)-Polyacrylamidgel-Elektrophorese (SDS-PAGE).<sup>[3]</sup>

Bei der IEF wandern die einzelnen Proteine einer Mischung in einem pH-Gradienten zu einer Zone, die ihrem isoelektrischen Punkt entspricht, verlieren dort ihre Nettoladung und damit ihre elektrophoretische Mobilität. Falls sich ein Protein durch thermische Diffusion aus dem pH-Bereich seines isoelektrischen Punktes herausbewegt, erhält es eine Ladung und wird in diesem Moment wieder elektrophoretisch bis zu seiner „richtigen“ Position, die seinem isoelektrischen Punkt entspricht, bewegt. Die isoelektrische Fokussierung ist also eine konzentrierende Endpunktmethode, bei der sehr scharf begrenzte, hochkonzentrierte Proteinbanden entstehen. Sie ist auch mit (nichtionischen) Detergentien kompatibel, was eine wesentliche Voraussetzung für ihren Einsatz in der Proteomanalyse ist.

Bei der SDS-PAGE werden alle Proteinmoleküle mit SDS beladen und wandern als negative SDS-Protein-Komplexe in einem elektrischen Feld in Richtung der Anode und werden in der Polyacrylamidmatrix nach ihrer Größe (d.h. ihrer Molekülmasse) getrennt.

Bei der Kombination von IEF und SDS-PAGE, der von Klose und O'Farrell 1975 entwickelten zweidimensionalen (2D) Gelelektrophorese,<sup>[3]</sup> wird das IEF-Gel mit seinen hochkonzentrierten Proteinbanden auf ein SDS-Polyacrylamidgel gelegt und die Proteine in einem zweiten Trennschritt weiter nach ihrer Molekülmasse getrennt. Es entstehen hochaufgelöste zweidimensionale Proteinmuster, die nach Anfärbung sichtbar gemacht und quantitativ ausgewertet werden können (Abbildung 3). Die 2D-Elektrophorese ist zur Zeit die einzige Methode, die in der Lage ist, für Proteine einen Trennraum von mehreren Tausend Komponenten zur Verfügung zu stellen und komplexe Proteingemische in wenigen Stunden zu trennen. Sie hat allerdings erst in den letzten Jahren durch die Einführung von immobilisierten pH-Gradienten (IPG)<sup>[13]</sup> für die isoelektrische Fokussierung und durch verbesserte Probenauftragstechniken<sup>[11]</sup> einen Stand erreicht, der zu reproduzierbaren Proteinmustern führt.

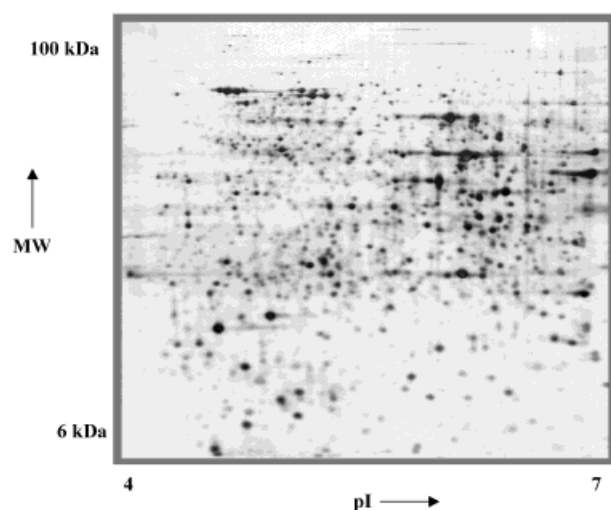


Abbildung 3. Typisches zweidimensionales Gel des Proteoms von *Saccharomyces cerevisiae* im pH-Wert-Bereich 4–7 (Immobilinteknik). MW = Molekülmasse.

Durch die Verwendung der IPG-Gele, die auch kommerziell als Fertiggele erhältlich sind, konnte ein großer Teil der Unzulänglichkeiten der isoelektrischen Fokussierung mit Ampholyten (Kathodendrift, nicht stabiles und in seiner Zusammensetzung variierendes Ampholytgemisch) behoben werden. Ein weiterer immenser Vorteil der Immobilinteknik ist, daß Gele mit sehr engen pH-Bereichen (z. B. ein Gradient von nur einer pH-Einheit auf einer Gelbreite von 18 cm) hergestellt werden können, die daher ein sehr hohes Auflösungsvermögen haben und auch noch Proteine trennen können, die sich in ihrem isoelektrischen Punkt um weniger als 0.01 pH-Einheiten unterscheiden (Abbildung 4).<sup>[14]</sup> Zudem haben diese Gele den Vorteil, mit einer großen Proteinmenge (bis ca. 15 mg) beladen werden zu können, so daß auch schwach exprimierte Proteine direkt sichtbar gemacht werden können.<sup>[11]</sup>

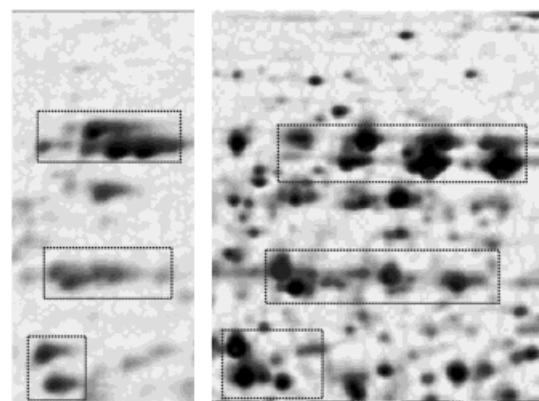


Abbildung 4. Gespreizte pH-Gradienten von IPG-Gele verbessern die Auflösung und die Beladbarkeit. Die Proteine in den markierten Bereichen auf dem links abgebildeten Gel (pH-Bereich 3–10) sind auf dem rechts abgebildeten Gel mit einem engeren pH-Bereich (5–7) deutlich besser getrennt.

Die besonderen Eigenschaften, welche die 2D-Gelelektrophorese zu einer so herausragenden Trenntechnik für die Proteomanalyse machen, sind:

- zwei komplementäre und effiziente Trennprinzipien (IEF trennt nach Ladung, SDS-PAGE trennt nach Molekulargewicht);
- prinzipiell für alle Proteine einsetzbar (Detergenskompatibilität);
- parallel und daher schnell;
- über enge pH-Gradienten können Teilbereiche des Proteoms mit sehr hoher Auflösung und bei hoher Proteinbeladung analysiert werden.

Den Vorteilen einer großen Universalität und Schnelligkeit stehen aber auch gravierende Nachteile gegenüber, die eine Weiterentwicklung dieser Technik dringend erforderlich machen:

- qualifizierte Probenvorbereitung notwendig;
- Probenauftrag ist mengenmäßig limitiert, so daß sehr schwach exprimierte Proteine nicht direkt erfaßt werden können;
- Proteintransfer von 1. auf 2. Dimension nicht quantitativ und schlecht reproduzierbar;

- keine einfache Quantifizierung, da die Proteine mit Farbstoffen angefärbt werden, die an verschiedene Proteine unterschiedlich binden;
- Proteine befinden sich in chemisch nicht inerten Polyacrylamidmatrix, die zu Modifikationen führen kann und die darüber hinaus eine direkte Analyse der getrennten Proteine verhindert;
- fehlende Automatisierung;
- die Durchführung der 2D-Gel-Analysen ist technisch anspruchsvoll, und es ist äußerst schwierig (zumindest zwischen einzelnen Laboratorien), dabei eine gute Vergleichbarkeit und Reproduzierbarkeit zu erreichen;
- nicht alle Proteine werden gleich gut erfaßt; außerdem werden keine brauchbaren Informationen für sehr kleine ( $M_r < 10000$ ) und sehr große Proteine ( $M_r > 100000$ ) erhalten.

Zusammenfassend stehen einer ausgezeichneten Auflösung und einer hervorragenden Kompatibilität mit fast allen Proteinklassen eine nur teilweise ausreichende Reproduzierbarkeit und Robustheit gegenüber. Dennoch ist die 2D-Gelelektrophorese zur Zeit die einzige Methode, die für Proteomprojekte eingesetzt wird.

### 3.3.2. Chromatographische Techniken

Bei den hochauflösenden chromatographischen Methoden (z.B. Ionenaustausch-, hydrophobe Interaktions-, Umkehrphasen-Chromatographie) wird die Selektivität durch die individuelle Wechselwirkung von Proteinmolekülen mit der chromatographischen Oberfläche, der stationären Phase, erreicht. Stark interagierende Proteine können irreversibel an die chromatographische Oberfläche adsorbieren. Leider sind diese Adsorptionsverluste schlecht vorhersehbar, da sie von der Gesamtproteinkonzentration, vom individuellen Protein und seiner Konzentration sowie von weiteren, schlecht kontrollierbaren Faktoren abhängen. Die adsorbierten Proteine ihrerseits verändern die Eigenschaften der stationären Phase, so daß eine reproduzierbare Trennung äußerst schwierig wird.

Ein schwerwiegendes Problem aller chromatographischen Methoden ist die Peakverbreiterung. Jedes Protein, auch wenn es in einem sehr kleinen Volumen (d.h. in hoher Konzentration) auf eine Säule aufgetragen wird, wird während der chromatographischen Trennung verdünnt und von der Säule in einer Konzentrationsverteilung eluiert, die einer Gauß-Kurve entspricht. Da man mit äußerst komplexen Proteingemischen zu tun hat, erfolgt der Fraktionswechsel zwar zu definierten Zeitpunkten, bezogen auf einzelne Proteine erfolgt er jedoch willkürlich. Dies führt dazu, daß viele Proteine in mehreren Fraktionen zu finden sind. Die letztendlich gewünschte Quantifizierung dieser Proteine wird dadurch erheblich erschwert.

Ein weiterer gravierender Nachteil bei chromatographischen Techniken ist, daß sie im allgemeinen seriell durchgeführt werden. Nach einer ersten Fraktionierung in z.B. 100 Fraktionen muß nun jede einzelne Fraktion der nächsten Trenndimension (z.B. einem weiteren Chromatographieschritt) zugeführt werden. Auch bei sehr schnellen Chromatographien (z.B. Gesamtdauer einer Analyse ½ h inklusive

Regenerierung des Säulenmaterials) wird die gesamte Aufarbeitung der Fraktionen aus der ersten Trennung mehrere Tage in Anspruch nehmen und darüber hinaus eine große Zahl von Fraktionen liefern, die weiter analysiert werden müssen. Proteine sind aber relativ labile Verbindungen, die in Lösung und vor allem im Gemisch mit anderen Proteinen schnell modifiziert, denaturiert oder abgebaut werden können. Inakzeptabel lange werden die Bearbeitungszeiträume, wenn etwa noch eine dritte chromatographische Trennung durchgeführt werden muß.

Aus den genannten Gründen haben multidimensionale chromatographische Trennungen zur Zeit noch keinen Eingang in die Proteomanalyse gefunden. Wegen der unbestreitbaren Vorteile der chromatographischen Techniken bei der Quantifizierung (UV-Detektion), der Automatisierung und der vielfältigen Möglichkeiten, Trennselektivitäten zu modulieren, besteht aber durchaus die Möglichkeit, daß multidimensionale Chromatographien eines Tages als zur Elektrophorese komplementäre Methoden für die Proteomanalyse eingesetzt werden.

## 3.4. Die quantitative Erfassung der getrennten Proteine

Da die meisten physiologischen und pathophysiologischen Prozesse mit quantitativen Veränderungen einzelner Proteinspezies einhergehen, ist der zentrale Punkt der Proteomanalyse die möglichst genaue Mengenbestimmung der getrennten Proteine. Dazu müssen die Proteine in einer auswertbaren Form sichtbar gemacht werden.

### 3.4.1. Detektion

Die am häufigsten eingesetzten Methoden sind verschiedene Färbungen. Die Problematik aller Färbemethoden ist, daß sie für jedes Protein (je nach seiner Aminosäurezusammensetzung und den vorhandenen Modifikationen) eine spezifische und nicht vorhersagbare Färbeintensität ergeben. Daraus folgt, daß keine absoluten Mengen einzelner Proteine angegeben werden können und Proteinmuster mit verschiedenen Arten von Färbungen nicht verglichen werden können. Darüber hinaus sind Färbungen aufgrund von Sättigungseffekten auch nur über einen sehr begrenzten Bereich linear (über maximal zwei Größenordnungen). Um einen großen dynamischen Bereich abzudecken, müssen mehrere Gele mit unterschiedlichen Proteinmengen hergestellt werden, damit eine grobe Eichkurve für jedes Protein erstellt werden kann. Für eine Quantifizierung eignen sich nur Proteine, die sich außerhalb des Sättigungsbereiches dieser Eichkurve befinden. Dazu ist die computerunterstützte Bildverarbeitung, gekoppelt mit einer leistungsfähigen Datenverarbeitung unerlässlich. Wenn verschiedene Gele miteinander verglichen werden müssen und dabei die Gesamtproteinmengen oder Färbeintensitäten unterschiedlich sind, stellt sich die Frage nach der Normierung. Dabei wird manchmal jeder Protein-spot auf die Intensität eines bekannten und konstitutiv exprimierten Proteins bezogen, was problematisch ist, da sich auch die Konzentrationen von konstitutiv exprimierten Proteinen signifikant ändern können. Der bessere Weg ist, die

Intensität jedes Proteins auf die Gesamtintensität aller Proteine zu beziehen.

Die bis jetzt am häufigsten für die Proteomanalyse eingesetzte Färbemethode ist die Silberfärbung, bei der die Proteine im Gel meist mit Trichloressigsäure fixiert werden und das Gel dann in eine Silbernitratlösung gelegt wird. Einige Silberionen werden von den Proteinen gebunden und durch Reduktion in Form elementaren Silbers präzipitiert. Durch die hohe Silberkonzentration an den Proteinen färben sich diese schnell dunkel. Man stoppt die Reaktion durch eine starke Änderung des pH-Wertes. Die Silberfärbung kann mit einer relativ hohen Empfindlichkeit Proteine sichtbar machen und weist einen linearen Bereich von ca. 0,5–20 ng Protein pro  $\text{mm}^{-2}$  auf. Allerdings ist diese Färbung sehr schwierig zu reproduzieren, da die Schwärzung stark von Entwicklungsdauer und Temperatur abhängt. Es existieren einige Durchführungsvorschriften, die sich in Dauer, Empfindlichkeit und Einfachheit unterscheiden und die fortlaufend verbessert werden.<sup>[15]</sup>

Etwas robuster als die Silberfärbungen sind Färbungen mit dem Triphenylmethanfarbstoff Coomassie Blue.<sup>[16]</sup> Sie sind aber auch deutlich weniger sensitiv (linearer Bereich ca. 50 ng bis 1  $\mu\text{g}$  Protein pro  $\text{mm}^{-2}$ ). Es existieren auch hier verschiedene Färbvorschriften, von denen die Methode nach Neuhoff,<sup>[16b]</sup> die eine Durchfärbung der Proteinspots erreicht, die besten Ergebnisse bezüglich einer quantitativen Auswertung liefert.

Proteindetektionsmethoden mit fluoreszierenden Farbstoffen wie SYPRO Orange oder SYPRO Red haben Vorteile gegenüber den konventionellen Färbungen mit Silber oder Coomassie Blue.<sup>[17]</sup> Sie sind etwa so empfindlich wie die Silberfärbung, sind allerdings viel schneller durchzuführen (ca. 30 min) und erfordern auch keine Fixierung der Proteine im Gel, so daß nachfolgende Spaltungen oder der Transfer der Proteine auf eine chemisch inerte Membran erleichtert sein sollten. Ein weiterer prinzipieller Vorteil der Fluoreszenzdetektion besteht darin, daß man, über unterschiedlich lange Meßzeiten, die Fluoreszenz häufig vorkommender Proteine neben der von seltenen Proteinen im gleichen Experiment detektieren und quantifizieren kann. Auch aus den fluoreszierenden Färbungen kann nicht auf die absolute Menge des vorhandenen Proteins geschlossen werden, da die Aminosäurezusammensetzung und die Art und das Ausmaß posttranslationaler Modifikationen die Farbeintensität beeinflussen. Da die Proteinspots nur unter UV-Licht sichtbar werden und ausgeschnitten werden müssen, ist bei Fluoreszenzfärbungen eine automatische Spoterkennung und das automatische Präparieren von Proteinspots für die nachfolgende Analytik besonders wichtig.

Prinzipiell wäre es auch möglich, alle Proteine eines Proteomzustandes vor ihrer Trennung mit einem fluoreszierenden Reagens kovalent zu verknüpfen. Dies brächte neben einer erhöhten Nachweisempfindlichkeit den Vorteil, daß man auch verschiedene Zustände mit verschiedenen Fluoreszenzfarbstoffen markieren könnte. Man könnte dann die Proteine beider Zustände unter identischen Bedingungen („Multiplexing“) im Gemisch trennen und über eine Detektion bei verschiedenen Emissions-Wellenlängen die entsprechenden Proteine den beiden Zuständen zuordnen. Differen-

zen wären so einfacher zu erkennen, da die technischen Probleme einer reproduzierbaren Trennung nicht berücksichtigt werden müssen.<sup>[18]</sup> Dabei müssen Fluoreszenzmarker mit unterschiedlichen optischen, aber elektrophoretisch identischen Eigenschaften eingesetzt werden. Die Problematik dieses Vorgehens liegt in der einheitlichen kovalenten Markierung der Proteine mit den Fluoreszenzfarbstoffen. Die Umsetzung ist eine chemische Reaktion an den funktionellen Gruppen eines Proteins. Diese Reaktion läuft in einem komplexen Gemisch sicherlich nicht vollständig ab. Selbst wenn eine Reaktionsausbeute von 99 % mit jeder einzelnen funktionellen Gruppe erreichbar wäre, würde aus einem einheitlichen Protein ein sehr heterogenes Gemisch, das durch hochselektive Trennmethode(n) (vor allem bei der isoelektrischen Fokussierung) auch aufgetrennt wird. Zwar mag jede der artifiziell generierten NebenkompONENTEN nur in einem kleinen Anteil von wenigen Promille entstehen, aber auch diese kleine Menge liegt praktisch immer in der Größenordnung von einigen natürlich vorhandenen Proteinen, da der Konzentrationsbereich unterschiedlicher Proteine in einer Zelle mindestens sechs Zehnerpotenzen umfaßt.

Man kann auch versuchen, die Proteine nicht vollständig, sondern mit sehr kleinen Mengen von Fluoreszenzreagentien umzusetzen. Dabei erreicht man, daß nur ein kleiner Teil des Proteins markiert wird und die Hauptmenge des Proteins unmodifiziert bleibt. Bei geeigneter Wahl des Reagens kann man erreichen, daß sich das Trennverhalten des modifizierten und des unmodifizierten Proteins nicht wesentlich unterscheiden.<sup>[18]</sup>

Teilweise könnte auch eine Fluoreszenzreaktion nach der Trennung in der ersten Dimension der 2D-Elektrophorese einen Ausweg bieten, da ja dann einerseits die hochauflösende Trennung durch die isoelektrische Fokussierung abgeschlossen ist und andererseits die durch wenige Fluoreszenzmoleküle verursachte Massenheterogenität unter dem Auflösungsvermögen der zweiten Dimension (der SDS-PAGE) liegt.<sup>[19]</sup>

Immunologische Färbungen mit Antikörpern sind sehr sensitiv und können auch sehr seltene Proteine direkt nachweisen. Leider gibt es keine für die Peptidbindung spezifischen Antikörper, so daß diese nicht als allgemeines Detektionswerkzeug eingesetzt werden können. Sehr wertvoll sind Antikörper aber, um ganz bestimmte Proteine oder Proteingruppen nachzuweisen (z.B. an Tyrosinresten phosphorylierte Proteine mit einem Phosphotyrosin-Antikörper). Ähnliches gilt für Lektin-Färbungen, mit denen glycosylierte Proteine nachgewiesen werden können.

Für spezielle Fragestellungen, bei denen eine Inkorporierung von Radioaktivität während der Kultivierung von Zellen möglich ist, können radioaktive Markierungen sehr sinnvoll eingesetzt werden. Dies bedeutet normalerweise den Umgang mit hohen Radioaktivitätsmengen, wofür speziell ausgerüstete Labors vorhanden sein müssen. So können z.B. über metabolische Markierung mit [ $^{35}\text{S}$ ]Met oder [ $^{35}\text{S}$ ]Cys in Pulse-chase-Experimenten spezifisch Proteine nachgewiesen werden, die zu bestimmten Zeitpunkten synthetisiert werden. Oder es kann eine selektive Analyse phosphorylierter Proteine (Phosphoprotein-Partialproteom) durch Markierung mit [ $\gamma$ - $^{32}\text{P}$ ]ATP erreicht werden.

Auch wenn die radioaktive Markierung eine sehr empfindliche Detektionsmethode ist und die Quantifizierung durch die Entwicklung der Phosphoimaging-Techniken in ihrer Linearität stark verbessert wurde, so reflektiert auch sie nicht die absolute Menge eines Proteins, da der Einbau der Markierung von der Aminosäurezusammensetzung der einzelnen Proteine abhängig ist.

### 3.4.2. Quantifizierung

Der nächste Schritt nach der Detektion der getrennten Proteine ist deren Quantifizierung. Außer der Aminosäureanalyse existiert keine Möglichkeit, die Absolutmengen der einzelnen in einem 2D-Gel getrennten Proteine zu bestimmen. Die Aminosäureanalyse ist aber sehr kontaminationsempfindlich und arbeitsaufwendig (wenn auch gut automatisierbar) und liefert nur zuverlässige Resultate bei Proteinmengen von mehr als ca. 0.4 µg. Damit ist sie für den Nachweis der meisten Proteine in einem 2D-Gel zu unempfindlich.

Normalerweise werden die gefärbten Gele computerunterstützt laserdensitometrisch vermessen. Es hat sich gezeigt, daß nur Laserscanner den erforderlichen hohen dynamischen Bereich aufweisen, um gute Resultate zu liefern. Zur Datenaufnahme und weiteren Auswertung werden spezielle Softwarepakete zur automatisierten Spoterkennung, Quantifizierung und Darstellung der Ergebnisse verwendet. Die Rohdaten und Resultate werden in strukturierter Form in großen Datenbanken abgelegt, auf die dann spezialisierte Datenbankprogramme zum Zweck des „Data-mining“ zugreifen können.

### 3.5. Datenauswertung mit Methoden der Bioinformatik

Die große Zahl von Daten, die bei einer Proteomanalyse anfallen, können nicht mehr ohne massive Unterstützung von spezialisierter Datenbanksoftware ausgewertet werden. Dabei ist nicht nur ein schneller Zugriff auf die Daten von Gelposition, Menge und Identität eines Proteins wichtig, sondern auch auf die zum Teil recht unstrukturierten Daten über Herkunft, Herstellung und Probenaufarbeitungshistorie der Probe. Weiterhin müssen auch klinische Datenbanken und Literaturdaten eingebunden werden können. In der Regel sind die gewünschten Informationen weltweit in unterschiedlichen Datenbanken abgelegt, die über On-line-Verbindungen verknüpft werden müssen. Teilweise sind die Informationen aus öffentlich zugänglichen Datenbanken über das World Wide Web abrufbar,<sup>[1b]</sup> andere Informationen sind aber so sensibel, daß sie nicht öffentlich zugänglich sein können. Die Datenbanksoftware muß auch komplexe Abfragen zulassen, z.B.: Welches Medikament beeinflusst die Expression welcher Proteine in welcher Patientengruppe positiv, oder welche Substanzen beeinflussen die Expression bestimmter Proteine in ähnlicher Weise? Für solche Fragen sind die zur quantitativen Auswertung der 2D-Gele verwendeten Programme hoffnungslos überfordert. Spezielle Datenbankprogramme für ein extensives Data-mining von RNA-Expressionsdaten und Proteomdaten sind zur Zeit ein

Schwerpunkt der Entwicklungen in der Bioinformatik. Auch der anschaulichen Präsentation der Daten aus Proteomanalysen kommt ein hoher Stellenwert zu. Dies können einfache Balkendiagramme sein, in denen die Proteinmengen für einzelne Proteine/Proteingruppen aus verschiedenen Experimenten dargestellt werden kann (Abbildung 5). Man

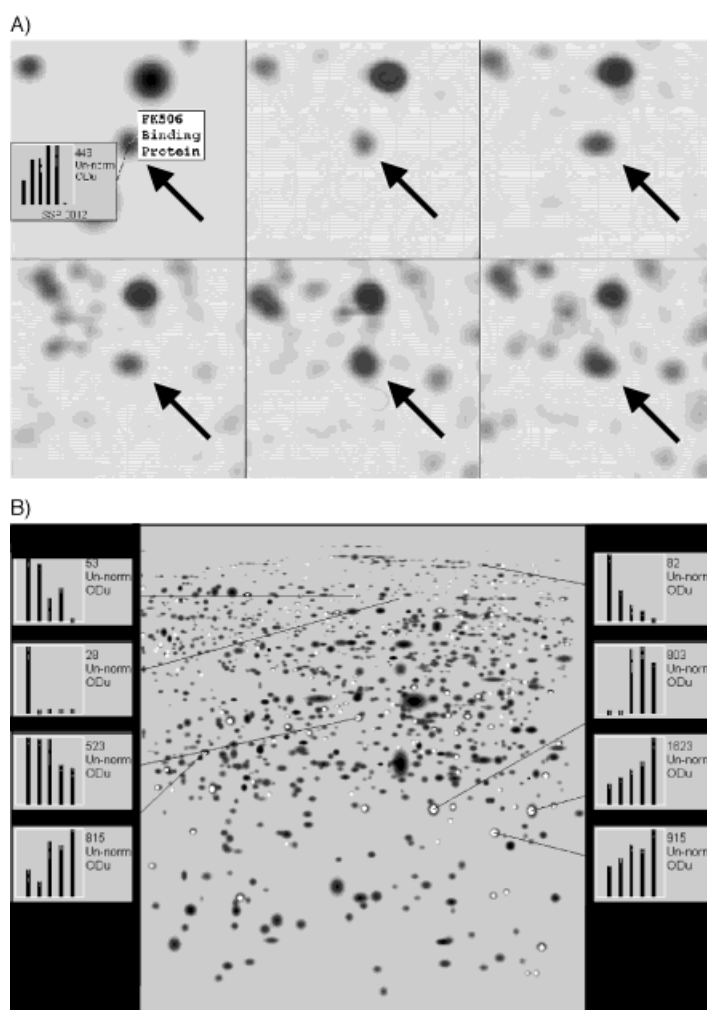


Abbildung 5. Darstellung von Proteomdaten. a) Veränderung der Menge eines Proteins bei verschiedenen Experimenten. b) Darstellung der Veränderungen mehrerer Proteine bei verschiedenen Experimenten.

erkennt gut den Verlauf der zu jedem Zeitpunkt vorhandenen Proteinmenge, wobei Proteine mit ähnlichem quantitativen Verlauf über Cluster- und Korrelationsanalysen weiter zusammengefaßt und in Beziehung gesetzt werden können.<sup>[20]</sup> Anschaulich sind auch Darstellungen, die komplexere Zusammenhänge verdeutlichen können und die auch Signifikanzparameter anschaulich berücksichtigen können (Abbildung 6).

### 3.6. Proteinchemische Analytik der gelelektrophoretisch getrennten Proteine

Die Methoden der Analytik gelelektrophoretisch getrennter Proteine war in den letzten 15 Jahren Gegenstand intensiver Bemühungen. Da die Proteine in einer Gelmatrix

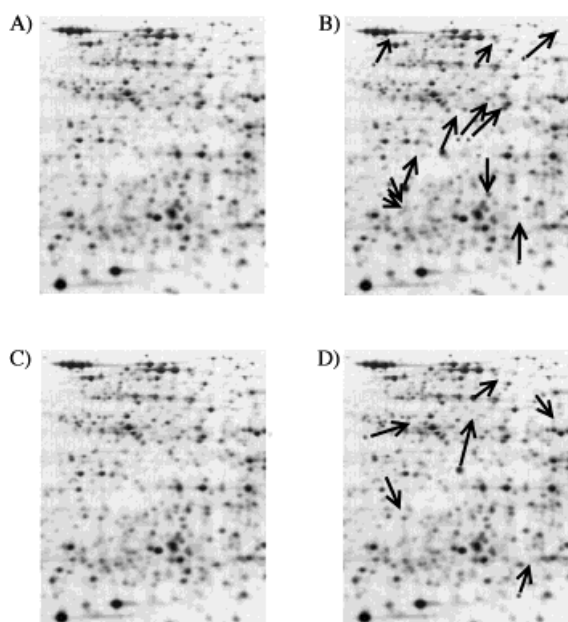


Abbildung 6. Darstellung einer Proteomanalyse unter Berücksichtigung von statistischen Parametern. Gezeigt ist ein schematisches Beispiel zur Abschätzung der Wirkung eines Medikaments. A) Proteom einer normalen Zelle. B) Proteom einer pathologisch veränderten Zelle. Die Pfeile zeigen die für den Krankheitszustand charakteristisch veränderten Proteine an. C) Pathologisch veränderte Zelle unter Einfluss eines idealen Medikaments. Die pathologischen Veränderungen werden revertiert; die Zelle erscheint normal. D) Pathologisch veränderte Zelle unter Einfluss eines realen Medikaments. Einige pathologische Veränderungen werden korrigiert, aber neue Veränderungen des Proteoms werden möglicherweise induziert – ein Indiz für Nebenwirkungen. Die Steigung des Pfeiles nach oben oder unten ist hier ein Maß für Zu- bzw. Abnahme der jeweiligen Proteinmenge. Die Länge des Pfeiles zeigt die statistische Signifikanz dieser Veränderung an (alle unveränderten Proteine hätten waagerechte Pfeile, die der Übersichtlichkeit halber nicht abgebildet sind).

für die normalen proteinchemischen Methoden wie Sequenzanalyse, Aminosäureanalyse und Massenspektrometrie praktisch nicht zugänglich sind, war der erste Erfolg, die Proteine aus dem Gel auf chemisch inerte Membranen zu transferieren und sie dort zu immobilisieren (Blotting).<sup>[4a-d]</sup> Die intakten immobilisierten Proteine können einerseits durch Aminosäuresequenzanalyse, durch Aminosäureanalyse oder seit neuestem auch mit massenspektrometrischen Methoden<sup>[21, 22]</sup> untersucht werden. Da jedoch die Analyse der intakten Proteine aufwendig ist und selten ausreichende Informationen für eine eindeutige Proteinidentifizierung liefert, hat sich die Analyse von internen Fragmenten nach einer enzymatischen oder chemischen Spaltung der Proteine als die effizientere Strategie erwiesen. Das heute generell verwendete Vorgehen im Rahmen der Proteomanalytik ist in Abbildung 7 dargestellt und wird in den folgenden Abschnitten diskutiert.

### 3.6.1. Analyse intakter Proteine

#### Aminosäuresequenzanalyse

Die Aminosäuresequenzanalyse geblotteter Proteine ist ein automatisiertes, relativ aufwendiges, teures und langsames Standardverfahren und kann ohne größere Probleme für Proteinmengen im Picomolbereich eingesetzt werden.<sup>[4a-e]</sup>

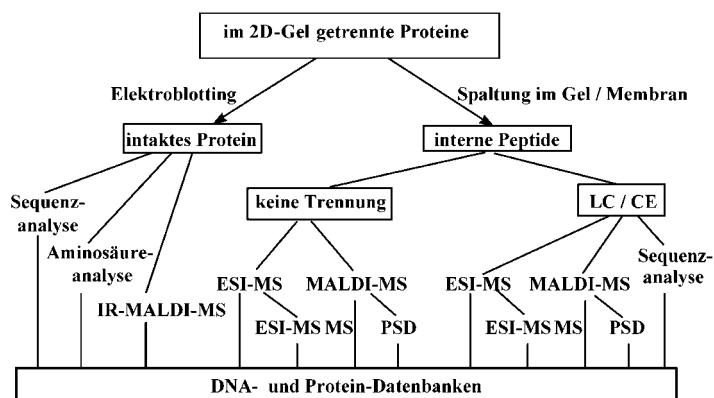


Abbildung 7. „High-throughput“-Proteinidentifizierung. Strategien zur Charakterisierung gelelektrophoretisch getrennter Proteine sind schematisch dargestellt. CE = Kapillarelektrophorese.

Dabei wird ein Protein oft anhand seiner N-terminale Sequenz identifiziert oder zumindest eindeutig charakterisiert. Die erreichbare Sequenzlänge (maximal 40 Aminosäuren) genügt zwar für die Identifizierung oder für Homologiesuchen in Datenbanken, sie deckt aber immer nur einen kleinen Teil der Gesamtsequenz des Proteins ab. Daher können nahe verwandte Isoenzyme, Spleißvarianten, Modifikationen oder Punktmutationen außerhalb des untersuchten aminoterminalen Bereichs des Proteins nicht erkannt werden. Das Hauptproblem ist aber, daß mehr als die Hälfte aller Proteine N-terminal modifiziert und damit einer Sequenzanalyse nicht zugänglich sind.

#### Aminosäureanalyse

Jedes Protein weist eine charakteristische Aminosäurezusammensetzung auf. Daher kann man fast jedes Protein allein über die Verhältnisse seiner Aminosäuren zueinander durch Suche in einer Proteindatenbank mit hoher Wahrscheinlichkeit identifizieren.<sup>[4e-h]</sup> In der Praxis sind aber der Genauigkeit der Bestimmung der einzelnen Aminosäuren prinzipielle Grenzen gesetzt. Jede Aminosäureanalyse ist, wegen der bei der Hydrolyse der Peptidbindungen anzuwendenden drastischen Reaktionsbedingungen, ein Kompromiß zwischen vollständiger Spaltung der Peptidbindungen und möglichst geringfügiger Zerstörung von empfindlichen Aminosäuren. So werden nur die stabilsten Aminosäuren, mit relativ großen Fehlergrenzen, für eine Datenbanksuche verwendet. Die Methode läßt sich gut automatisieren und ist annähernd ähnlich empfindlich wie die Sequenzanalyse, hat ihr gegenüber aber den Vorteil, daß sie auch bei N-terminal blockierten Proteinen Resultate liefern kann. Gleichzeitig ist die Aminosäureanalyse die einzige Methode, die für ein Protein eine absolute Mengenangabe liefern kann. Leider stehen in der Praxis diesen Vorzügen einige schwerwiegende Limitationen gegenüber, die den generellen Wert der Aminosäureanalyse in der Proteomanalytik stark einschränken. Die Technik kann nur Proteine identifizieren, deren Sequenz in einer Datenbank abgelegt sind. Bei den meisten Proteinen liegt die Proteinsequenz aber nur als Übersetzung der DNA-Sequenz vor, die sich oft signifikant von der Sequenz der natürlich vorkommenden Form des Proteins unterscheidet. Nach der

Translation werden Signalsequenzen abgespalten und Proteine weiter prozessiert und modifiziert. Damit unterscheidet sich die Aminosäurezusammensetzung des experimentell zugänglichen Proteins von seiner theoretischen in der Datenbank abgelegten. Dies ist in den Datenbanken normalerweise gar nicht oder nur marginal berücksichtigt. Wegen der genannten technischen und prinzipiellen Probleme und wegen ubiquitär vorhandener und nicht zu vermeidender Kontaminationen (freie Aminosäuren aus Puffern, Keratine etc.) ist der Fehlerbereich bei der Bestimmung der einzelnen Aminosäuren so groß, daß eine Identifizierung vor allem schwach exprimierter Proteine oft nicht sicher ist. Einzelne Aminosäureaustausche oder Modifikationen eines identifizierten Proteins können in der Praxis nicht erkannt werden.

### IR-MALDI-Massenspektrometrie

Ein signifikantes Charakteristikum eines Proteins ist auch seine Molekülmasse. Auch wenn die heutigen Methoden die Massen von Proteinen nicht genau genug bestimmen können (Fehler der Massenbestimmung > 100 ppm), um ein Protein allein anhand seiner Masse in einer Datenbank zu identifizieren, so ist die Information über die Masse des Gesamtproteins in Verbindung mit anderen Daten äußerst wertvoll. Die Qualität der Massenbestimmung reicht nämlich – zumindest für kleinere Proteine – aus, um posttranslationale Modifikationen über den Vergleich der theoretisch aus der DNA-Sequenz berechneten mit der experimentellen Masse zu erkennen. Dazu muß allerdings die Identität eines Proteins über Sequenzanalyse, Aminosäureanalyse oder die in Abschnitt 3.6.2 beschriebenen Methoden ermittelt sein.

Die MALDI-MS bietet sich für die Analyse von immobilisierten Proteinen an (Abbildung 8).<sup>[8, 21, 22]</sup> Die Proteine werden bei einer Standardpräparation für die MALDI-MS in eine kristalline Matrix kleiner organischer Moleküle eingebettet, in der sie dann durch Laserbeschuß verdampft und ionisiert werden. Dabei hat die Matrix die Aufgaben, die einzelnen Proteinmoleküle zu separieren, das Laserlicht zu absorbieren und die Energie durch Relaxation in kurzer Zeit in das Festkörpergitter zu übertragen. Dadurch wird eine explosionsartige Auflösung eines kleinen Bereiches der Festkörperoberfläche und ein Übergang von Matrix und Proteinmolekülen in die Gasphase erreicht. Dieser Vorgang ist bei richtiger Wahl der Laserenergie aber so schonend, daß auch die großen, thermisch labilen Proteinmoleküle intakt bleiben. Wahrscheinlich spielt die Matrix auch eine Rolle bei der Ionisierung der Proteinmoleküle. Es können verschiedene Wellenlängen des Laserlichtes eingesetzt werden, wobei das Zusammenspiel von Laserart (UV- oder IR-Laser) und Matrix (z. B. 2,5-Dihydroxybenzoesäure für UV-MALDI oder Bernsteinsäure für IR-MALDI) wichtig ist. Die entstandenen Protein-Ionen werden dann in einen Flugzeitanalysator beschleunigt und die genaue Flugzeit der Ionen von der Ionisierung bis zur Detektion gemessen (Abbildung 8a).<sup>[7a]</sup> Da die Flugzeit bei einer gegebenen Beschleunigungsspannung und Flugstrecke nur von der Wurzel aus dem Masse/Ladungs-Verhältnis ( $m/z$ ) abhängig ist, kann über eine Eichung eine genaue Massenbestimmung der Analytmoleküle erfolgen.

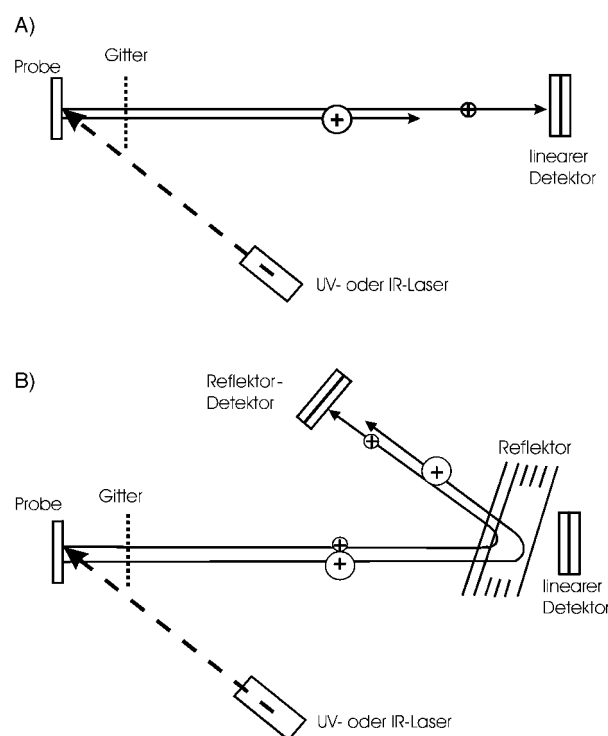


Abbildung 8. a) Matrix-assistierte-Laser-Desorption/Ionisation (MALDI)-Massenspektrometrie mit einem linearen (A) und einem Reflektor-Time-of-flight (TOF)-Detektor (B). A) Ein Laserpuls verdampft und ionisiert die Moleküle der Probe. Diese werden durch die am Gitter anliegende Spannung in Richtung des TOF-Analysators beschleunigt. Die Fluggeschwindigkeit der Ionen ist proportional zu  $1/\sqrt{m}$  ( $m$  = Ionenmasse). B) Während des Fluges durch die feldfreie Driftstrecke können Ionen zerfallen (PSD). Diese Bruchstücke, die Strukturinformation enthalten, fliegen mit derselben Geschwindigkeit und würden den linearen Detektor zur selben Zeit erreichen. Zur Trennung werden sie vom Potentialberg des Reflektors gebremst, zur Umkehr gezwungen und in Richtung des Reflektor-Detektors beschleunigt. Da Ionen mit größerer Masse tiefer in das Reflektorfeld eindringen, durchlaufen sie eine längere Strecke bis zum Detektor als leichtere Ionen.

Um elektrophoretisch getrennte Proteine nach einem Transfer auf eine chemisch inerte Membran direkt mit MALDI-MS analysieren zu können, muß sofort nach dem Elektrophoretikum die Membran mit der Matrix inkubiert werden. Die UV-MALDI-MS benötigt als Matrix meist sehr hydrophobe Moleküle, die nur in organischen Lösungsmitteln löslich sind. Bei der Inkubation einer Membran mit einer solchen hydrophoben Matrix werden die Proteinmoleküle zum Teil gelöst, und die geometrische Anordnung der Proteinspots geht durch Diffusion verloren.<sup>[22a]</sup>

Im Unterschied dazu wurde gezeigt, daß bei der Inkubation einer Membran mit einer hydrophilen Matrix die lokale Anordnung der geblotteten Proteinspots auf der Membran und somit die Auflösung der Gelelektrophorese und sogar die Intensitätsverteilung innerhalb eines Proteinspots vollständig erhalten bleiben.<sup>[21]</sup> Die IR-MALDI-MS verwendet solche hydrophilen Matrices und ist so in der Lage, durch Elektrophoretikum immobilisierte Proteine sehr empfindlich (bis in den Attomolbereich) zu analysieren.<sup>[21b]</sup> Dabei können die Grenzen und Intensitäten von Proteinspots gut erkannt werden, und es können auch Proteingemische innerhalb eines 2D-Gel-Spots erkannt werden. Leider ermöglichen die für die einzel-

nen Proteine sehr unterschiedlichen Signalintensitäten keine Abschätzung der jeweiligen Proteinmengen. Die Entwicklungen bei der Automatisierung sowie methodische Verbesserungen bei der MALDI-MS bezüglich Ionenausbeute und Massengenauigkeit für größere Proteine lassen erwarten, daß in naher Zukunft eine Detektion von durch Elektroblothing immobilisierten Proteinen durch ein automatisches Abscannen von Blottingmembranen mit MALDI-MS erfolgen kann, wobei dann ein Spot außer durch seine elektrophoretische Position auch durch die Masse charakterisiert wird.

### 3.6.2. Analyse interner Fragmente

#### *Enzymatische Spaltung und Elution der Peptide*

Da die Analyse der intakten Proteine aufwendig ist und selten ausreichende Informationen liefert, hat sich als die effizienteste Strategie für die Charakterisierung und Identifizierung elektrophoretisch getrennter Proteine die Analyse von internen Fragmenten dieser Proteine erwiesen. Nach der enzymatischen Spaltung der getrennten Proteine, die direkt in der Polyacrylamidgel-Matrix durchgeführt werden kann,<sup>[4i-o]</sup> werden die erhaltenen Peptide eluiert und über massenspektrometrische und/oder andere proteinchemische Methoden analysiert. Eine Unzahl von Vorschriften existiert, die jeweils leichte Abwandlungen aufweisen, aber prinzipiell auf eine einzige Arbeit<sup>[4i]</sup> zurückgehen. Der Spot des gelelektrophoretisch getrennten Proteins wird möglichst eng ausgestochen, gewaschen und entweder durch Trocknen oder durch Zugabe von Acetonitril dehydratisiert. Auf das geschrumpfte Gelstück wird ein kleines Volumen gepufferter Enzymlösung pipettiert. Als Enzyme werden hauptsächlich Trypsin, Endoprotease Lys-C, Endoprotease Glu-C oder Endoprotease Asp-N verwendet, die alle Proteine sehr spezifisch und vollständig schneiden. Vor allem wenn massenspektrometrische Analysen folgen, wird Trypsin sehr häufig eingesetzt, da es sich relativ schnell selbst spaltet und die Autoproteolysefragmente als interne Referenzpeptide für die Massenkalisierung dienen. Nach Inkubation über einige Stunden bei erhöhter Temperatur wird die Reaktionsmischung, die jetzt die gespaltenen Peptide teils im Überstand, teils noch im Gel enthält, für die nachfolgenden Analysenmethoden unterschiedlich aufgearbeitet. Der Überstand der Spaltungslösung kann direkt analysiert werden, wobei die Ausbeuten an großen oder sehr hydrophoben Peptiden oft schlecht sind. Meist erfolgt zusätzlich die Elution dieser Peptide aus den Gelstücken mit einer Lösung einer flüchtigen Säure in einem organischen Solvens, z.B. 0.1–1proz. Trifluoressigsäure in Acetonitril. Die Weiterverarbeitung der eluierten Peptide richtet sich vor allem danach, ob das zu untersuchende Proteom von einem Organismus stammt, dessen Genom bereits vollständig sequenziert ist. Der hohe Probenanfall bei der Proteomanalytik führte schon zur Automatisierung der Spaltung und der nachfolgenden Elution von gelelektrophoretisch getrennten Proteinen.<sup>[23]</sup> Es stehen heute käufliche Geräte zur Verfügung, die 50 Proben pro Tag (bei Probenmengen von >1 pmol) automatisch bearbeiten können.

#### *Interne Sequenzen von Proteinen aus einem Organismus mit vollständig sequenziertem Genom*

Eine sehr einfache und schnelle Identifizierung der Proteine eines Organismus mit vollständig sequenziertem Genom kann mit ausschließlich massenspektrometrischen Methoden erreicht werden. Dabei werden die eluierten Peptide ohne weitere Trennung entweder mit MALDI-MS oder mit Nano-Elektrospray-MS analysiert.<sup>[7d, 24]</sup>

Im allgemeinen wird man das durch Spaltung erhaltene Peptidgemisch im Gel direkt der MALDI-MS zuführen, wobei auch hier automatisierte Probenauftragssysteme zum Einsatz kommen. Die MALDI-MS ist heute schon so weit entwickelt, daß die Proben automatisch mit sehr hoher Massengenauigkeit (Fehler <20 ppm) analysiert und on-line so ausgewertet werden können, daß aus dem erhaltenen Peptidmassenmuster das Protein in Sequenzdatenbanken mit sehr hoher Wahrscheinlichkeit identifiziert werden kann. Über diesen Weg können mehr als 90% der analysierten Proteine unzweifelhaft und mit hoher Sequenzabdeckung direkt identifiziert werden (Abbildung 9). Bei einigen Proteinen ist die Identifizierung aber nicht eindeutig, so daß weitere Analysen durchgeführt werden müssen. Dies kann mit der MALDI-MS über Post-source-decay(PSD)-Spektren erreicht werden.<sup>[25]</sup> Dabei wird die Tatsache genutzt, daß Proteine und Peptide während des Fluges in der feldfreien Driftstrecke des TOF-Analysators spontan zerfallen (metastabiler Zerfall). Die Fragmente eines Ions fliegen alle mit derselben Geschwindigkeit weiter und erreichen den Detektor eines linearen TOF-Analysators zur selben Zeit. Wenn man die Fragment-Ionen aber in einem TOF-Analysator mit Reflektor gegen ein gleichgeladenes Feld (den Reflektor) fliegen läßt, kommen die Fragment-Ionen zum Stillstand und kehren um (Abbildung 8b). Dabei dringen die verschiedenen Massen unterschiedlich tief in das Reflektorfeld ein, haben daher unterschiedliche Wege zurückzulegen und erreichen den Detektor des Reflektor-Instruments zu unterschiedlichen Zeiten. So werden die PSD-Fragmente getrennt und können über Eichungen mit Referenzverbindungen auch in ihrer Masse bestimmt werden. Die Interpretation der Spektren ist relativ kompliziert und langwierig, so daß bei einem Proteomansatz mit den vielen zu analysierenden Proben eine exakte Auswertung dieser Spektren nicht durchgeführt werden kann. Daher werden die so erhaltenen, nicht interpretierten Spektren automatisch und on-line mit einer Datenbank von berechneten Spektren verglichen, die aus allen theoretisch möglichen Peptiden eines Organismus in silico erzeugt wurden. Die Software liefert dann über die Ähnlichkeit einen Identifikationsvorschlag.<sup>[26]</sup>

Eine Alternative zur MALDI-MS-Analyse des aus der Spaltung im Gel resultierenden Peptidgemisches ist die Elektrospray-Ionisations(ESI)-MS (Abbildung 10).<sup>[7b-e]</sup> Bei der ESI-MS wird die Probe in flüssiger Form kontinuierlich mit Hilfe eines elektrostatischen Feldes in Form kleiner Tröpfchen versprüht, die dann schnell Lösungsmittel verlieren, wobei die Ladungsdichte an der Oberfläche der Tröpfchen immer größer wird. Nach wiederholtem spontanen Zerfall der Tröpfchen (Coulomb-Explosionen) treten schließlich desolvatisierte, mehrfach geladene Moleküle in das

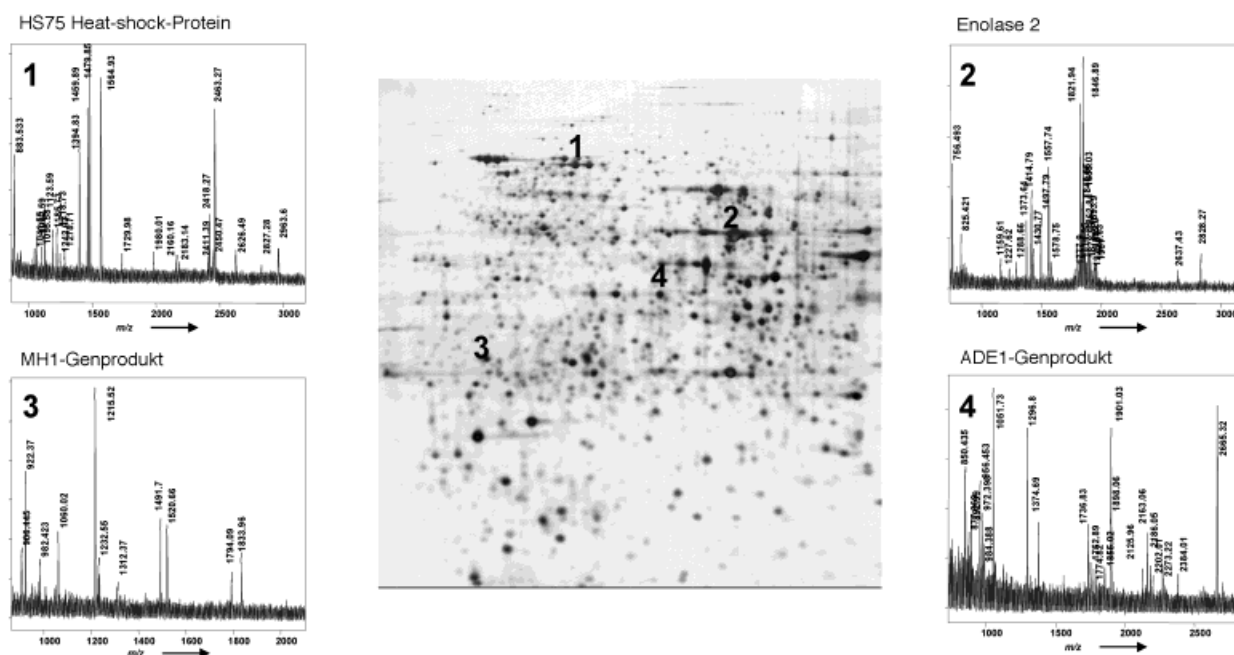


Abbildung 9. Identifizierung von 2D-elektrophoretisch getrennten Proteinen anhand von MALDI-Massenspektren. Die enzymatische Spaltung eines Proteins liefert ein Peptidgemisch, dessen Massen computerunterstützt durch Vergleich mit Datenbanken analysiert werden.

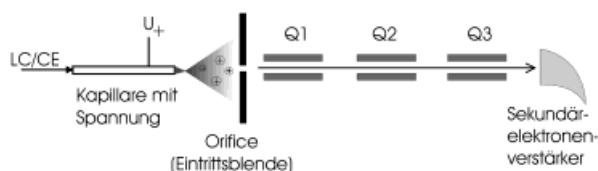


Abbildung 10. Schematische Darstellung des Aufbaus eines Elektrospray-Ionisations-Massenspektrometers. Die Probe wird in flüssiger Form (HPLC, CE, Spritze) kontinuierlich durch die Kapillare, an der eine hohe Spannung anliegt, gefördert. Der entstehende Sprühnebel (Spray) enthält die (mehrfach) ionisierten Moleküle, die über die Eintrittsblende (Orifice) in das Hochvakuumfeld des Quadrupol-Massenspektrometers eingebracht werden. Für die Massenbestimmung werden die Quadrupole Q1 und Q2 so geschaltet, daß alle Ionen passieren können. Die eigentliche Massenbestimmung findet im Quadrupol Q3 statt. Für Strukturuntersuchungen läßt der Massenfilter Q1 nur Ionen einer einzigen Masse nach Q2 passieren. In Q2, der mit Argongas gefüllt ist, findet die Fragmentierung der Ionen statt, und in Q3 werden die Massen der entstandenen Fragmente analysiert.

Massenspektrometer ein, die normalerweise mit einem Quadrupol-Massenanalysator detektiert werden. Dieser Typ von Massenanalysator ist ein Massenfilter, der unter vorgegebenen physikalischen Bedingungen nur Ionen mit einem definierten Masse/Ladungs-Verhältnis durchläßt. Alle anderen Ionen können den Analysator nicht passieren und werden nicht erfaßt. Durch kontinuierliche Veränderung der am Quadrupol anliegenden Spannungen werden nacheinander die Ionen verschiedener Massen durchgelassen (Scannen) und die Intensität des Ionenstroms am Detektor in Abhängigkeit vom Verhältnis  $m/z$  aufgezeichnet. Die Genauigkeit der Massenbestimmung ermöglicht es, aus der Isotopenverteilung den zu jedem Massensignal gehörigen Ladungszustand zu ermitteln, somit mehrfach geladene Ionen zu erkennen und computerunterstützt die Masse des einfach geladenen Ions zu berechnen.

Da der gemessene Ionenstrom mit der Konzentration der versprühten Probe korreliert, versucht man, mit sehr kleinen Flußgeschwindigkeiten ( $\text{nL min}^{-1}$ ) hochkonzentrierte Lösungen zu versprühen (Nanospray-ESI).<sup>[7d]</sup> Dies hat neben einer gesteigerten Empfindlichkeit den Vorteil, daß mit einer Probe von wenigen  $\mu\text{L}$  sehr lange Meßzeiten zur Verfügung stehen.

Die ESI-MS bietet auch die Möglichkeit, über eine Fragmentierung einzelner Peptide eine zumindest partielle Struktur- oder Sequenz-Information zu gewinnen.<sup>[27]</sup> Dazu verwendet man ein Triple-Quadrupol-Gerät, bei dem der erste Quadrupol für die Selektion eines Peptidions verwendet wird. Diese selektierten Ionen werden in einen zweiten Quadrupol geführt, in dem sie mit Argongas kollidieren und dabei fragmentiert werden. Die entstandenen Fragmente werden dann mit einem dritten Quadrupol analysiert (Abbildung 10). Die ESI-Tandem-Fragmentspektren sind etwas klarer und besser zu interpretieren als MALDI-PSD-Spektren, aber auch hier werden bei der Proteomanalyse zunehmend automatisierte, softwaregesteuerte Interpretationsprogramme eingesetzt. Die C-terminale Markierung der Peptide mit  $^{18}\text{O}$ , was durch enzymatische Spaltung des Proteins in Gegenwart von  $\text{H}_2^{18}\text{O}$  erreicht wird, erleichtert die Interpretation der Fragmentspektren erheblich.<sup>[28]</sup> Eine Alternative zu den Quadrupolgeräten sind die Ionenfallen, bei denen Ionen in einem geeigneten elektrischen Feld eingefangen werden und dort auf stabilen Bahnen gehalten werden können.<sup>[29]</sup> Über Veränderungen der elektrischen Bedingungen werden dann die einzelnen Ionen aus der Falle katapultiert und in einem Detektor nachgewiesen. Die Scangeschwindigkeiten sind hier bis zu zehnmal höher als bei einem Quadrupoldetektor, daher eignet sich die Ionenfalle besonders als schneller Detektor für die Kopplung mit der HPLC, bei der jeder Substanzpeak nur für eine limitierte Zeit zur Verfügung steht. Auch die Ionenfalle eignet sich wie die Triple-Quadrupolgeräte für

die Strukturbestimmung von Peptiden, da einzelne Ionen in der Falle selektiert und durch Kollision mit Inertgasatomen fragmentiert werden können. Aus den Ergebnissen der massenspektrometrischen Sequenzanalyse, die meist keine vollständigen Peptidsequenzen, sondern nur „Sequence tags“ liefert, und durch Kombination mit den Peptidmassen-Fingerprints kann in der Regel jedes Protein eindeutig identifiziert werden. Neuere Entwicklungen ermöglichen auch die Kopplung von Elektrospray-Ionenquellen mit TOF-Analysatoren (orthogonale TOF-Geräte (Q-TOF), Abbildung 11). Dabei

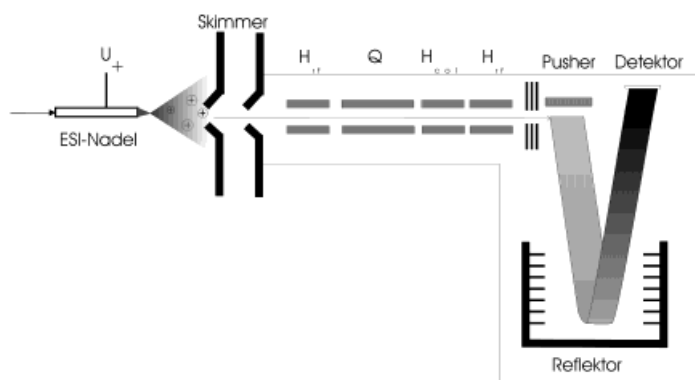


Abbildung 11. Schematische Darstellung des Aufbaus eines ESI-Q-TOF-Massenspektrometers. Der durch ESI erzeugte Ionenstrom wird über Quadrupole (Q) und Hexapole ( $H_{rf}$ ) zum Pusher geführt, der die Ionen einem Reflektor-TOF-Analysator zuführt. Eine Strukturinformation wird gewonnen, indem man Q zur Selektion eines Ions benutzt, das in der Kollisionszelle ( $H_{coll}$ ) fragmentiert wird. Die Fragment-Ionen werden in der TOF-Einheit analysiert.

wird der gesamte Ionenstrom über Quadrupole und Hexapole zu einem orthogonalen Beschleunigungssystem geführt (Pusher), das die Ionen einem Reflektor-TOF-Analysator zuführt. Auch diese Geräte ermöglichen eine Strukturaufklärung, indem man den vorgeschalteten Quadrupol zur Selektion nur eines Ions benutzt, in einer nachgeschalteten Kollisionszelle wird das selektierte Ion fragmentiert und dann durch den Pusher in die TOF-Einheit geführt und analysiert.

#### *Interne Sequenzen von Proteinen aus einem Organismus mit nicht sequenziertem Genom*

Die Strategie, Proteine mit der Massenspektrometrie allein über Peptidmassenmuster zu identifizieren, hat heute noch zwei wesentliche Limitationen.

- Es muß die Genomsequenz des Organismus weitgehend bekannt sein. Dies hat zwar heute noch eine praktische Bedeutung, wenn man z.B. bedenkt, daß das Humangenomprojekt noch nicht abgeschlossen ist, aber es ist in naher Zukunft abzusehen, daß die DNA-Sequenz der wichtigsten Organismen bekannt sein wird.
- Posttranslationale Modifikationen werden aus verschiedenen Gründen nur sehr begrenzt erkannt.

Die Methoden der schnellen Proteinidentifizierung über Peptidmassenmuster versagen, wenn die Sequenz des zu untersuchenden Proteins nicht in einer Datenbank vorhanden ist, wenn das Protein stark modifiziert ist oder wenn in einem Proteinspot mehrere Proteine vorhanden sind. Auch wenn

große Teile eines Genoms nur in Teilsequenzen, den EST-Datenbanken (EST=expressed sequence tag), zugänglich sind, wird die Analyse mit Peptidmassenmustern zu keinem Erfolg führen. Weiterhin führen Schwierigkeiten bei der Spaltung und bei der Elution von größeren Peptiden aus dem Gel zu einer schlechten Sequenzabdeckung, und auch bei der Massenspektrometrie selbst sind inhärente Probleme vorhanden. So verhindern Suppressioneffekte eine Quantifizierung einzelner Peptide,<sup>[30]</sup> Signale von ubiquitär vorhandenen Kontaminationen (z.B. Keratinen) komplizieren die Spektren, und artifizielle Modifikationen während der Probenaufarbeitung (z.B. Oxidationen oder Modifikationen von Cysteinsten) oder der Messung selbst (Oxidationen, Fragmentierungen) verhindern eine einfache Zuordnung der gefundenen Signale.

Für detaillierte Analysen wie der Bestimmung von posttranslationalen Modifikationen eines Proteins oder für die Charakterisierung von Proteinen aus Organismen mit noch nicht sequenziertem Genom müssen zusätzliche, aufwendige und langsame proteinchemische Mikrotechniken eingesetzt werden (Abbildung 7). Zur Trennung der Peptide nach der enzymatischen Spaltung im Polyacrylamidgel wird die Kapillar-HPLC oder Kapillarelektrophorese (CE) mit on-line gekoppelten massenspektrometrischen Verfahren eingesetzt. Solche Analysen, bei denen das Protein nicht nur identifiziert, sondern wirklich genau untersucht wird, sollten ausgehend von mindestens zwei unterschiedlichen enzymatischen Spaltungen durchgeführt werden, um mit den Sequenzdaten möglichst das gesamte Protein abzudecken. Im Endeffekt müssen fast immer Methoden angewendet werden, die eine Sequenzinformation „de novo“ liefern, also entweder massenspektrometrische Sequenzierungen, die aufwendig sind und leider oft auch mit hoher Unsicherheit behaftet sind, oder der klassische Edman-Abbau, der zwar eindeutige Ergebnisse liefert, aber bei ca. 1 pmol Peptid als Ausgangsmaterial an seine Empfindlichkeitsgrenze stößt. Beide Sequenzierungstechniken sind langsam und können heute noch nicht dem hohen Probenanfall einer Proteomanalyse gerecht werden.

#### **3.6.3. Analyse posttranslationaler Modifikationen**

Ein wichtiger Bereich einer Proteomanalyse ist die Analyse der posttranslationalen Modifikationen, die einen erheblichen Einfluß auf die Funktionen und Eigenschaften eines Proteins haben. Da diese detaillierte Proteinanalyse sehr aufwendig und langwierig ist, sollten nur die Proteine genauer untersucht werden, bei denen Hinweise auf posttranslationale Modifikationen vorliegen. Diese Information kann aus der beschriebenen massenspektrometrischen Analyse (siehe Abschnitt 3.6.1) des Gesamtproteins über IR-MALDI-MS erhalten werden. Auch eine Abweichung des gemessenen vom aus der DNA-Sequenz berechneten isoelektrischen Punkt eines Proteins ist ein guter Hinweis auf eine Modifikation.<sup>[31]</sup> Zur genauen Ermittlung der Art und Position der posttranslationalen Modifikation werden vor allem spezielle massenspektrometrische Techniken, wie Vorläufer-Ionen-Analyse („precursor scan“) oder Neutralverlust-Analyse („neutral loss scan“) und die massenspektrometrischen Sequenzierungsmethoden MALDI-PSD-MS und Nanospray-ESI-MS-

MS eingesetzt.<sup>[32]</sup> Ergänzt werden sie durch die klassischen Strukturermittlungsverfahren, hauptsächlich durch die Sequenzanalyse nach Edman. Besonders schwierig zu analysieren sind partielle Modifikationen an mehreren Stellen eines Proteins, was häufig bei Phosphorylierungen und Glycosylierungen auftritt. Hier müssen fast immer Trennungen des Peptidgemisches über Nano-HPLC mit online-massenspektrometrischer Analyse stattfinden. Zusammenfassend ist die Charakterisierung von posttranslationalen Modifikationen auch heute noch eine Herausforderung an den Proteinchemiker und kann – trotz enormer Fortschritte in den letzten Jahren – sicher noch nicht mit „High-throughput“-Methoden gelöst werden.<sup>[33]</sup>

#### 4. Zusammenfassung und Ausblick

Ein Proteom als das quantitative Proteinmuster eines Organismus, einer Zelle, oder einer Körperflüssigkeit unter exakt definierten Randbedingungen ist – im Gegensatz zum statischen Genom – hoch dynamisch. Zu einem einzigen Genom existieren viele Proteome, die jedes einen aktuellen Entwicklungs- und Stoffwechselzustand zu einem bestimmten Zeitpunkt widerspiegeln. Bei geeigneter Wahl unterschiedlicher Zustände können aus den unterschiedlichen Proteinmustern der entsprechenden Proteome mit Hilfe der Bioinformatik funktionelle Schlüsse gezogen werden.<sup>[13]</sup> Ihr ganzes Potential wird die Proteomanalyse sicher erst im Zusammenspiel mit anderen Teilbereichen der Biowissenschaften, wie Molekularbiologie, Genetik, Immunologie und Medizin, ausspielen können (Abbildung 12). Zur Zeit befindet sich die

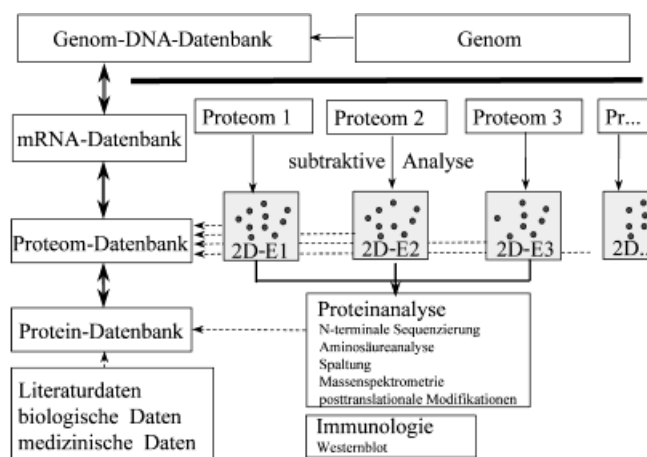


Abbildung 12. Die Proteomanalyse im Kontext der Teilbereiche der Biowissenschaften. 2D-E = zweidimensionale Elektrophorese.

Proteomanalyse noch „in statu nascendi“, die Methoden sind noch unausgereift, werden aber in erstaunlichem Tempo auf allen Ebenen weiterentwickelt. Bei den Trennungen wird versucht, die einzige bisher erfolgreiche Methode, die 2D-Gelelektrophorese, zu verbessern oder zu ersetzen, und die Quantifizierung erhält durch Automatisierung, neue Detektionsverfahren sowie den Einzug der Bioinformatik immer größere Bedeutung. Am weitesten fortgeschritten ist die

Identifizierung und Charakterisierung der getrennten Proteine. Sie ist für Proteine aus Organismen mit vollständig sequenziertem Genom schon fast zur Routine geworden, obwohl auch hier noch weiter Raum für Verbesserungen bezüglich Durchsatz und Empfindlichkeit gegeben ist. Das Erkennen und Lokalisieren posttranslationaler Modifikationen im „High-throughput“-Verfahren wird der nächste Schritt der Entwicklung sein. Die Proteomanalyse von Organismen mit nicht vollständig aufgeklärtem Genom ist heute noch so aufwendig, daß sie in naher Zukunft nur auf spezielle Fragestellungen beschränkt bleiben wird.

Da die Proteomanalytik auf eine Vielzahl von komplexen Fragestellungen angewendet werden kann, ist es sicher nur eine Frage der Zeit und des finanziellen Einsatzes, wann die ersten größeren praxisorientierten Proteomprojekte starten werden. Die erste Realisierung eines umfassenden Proteomprojektes wird sicher nur über eine Vernetzung der unterschiedlichen und nur teilweise vorhandenen Techniken mit gleichzeitiger massiver Automatisierung und Weiterentwicklung auf allen Ebenen der Proteomanalyse möglich sein, wobei auch die Entwicklung neuer Technologien abzusehen ist. Der große instrumentelle Aufwand und die große Expertise, die für eine Proteomanalyse gebraucht werden, lassen es aber wahrscheinlich erscheinen, daß größere Proteomprojekte nur in einigen dafür spezialisierten Zentren durchgeführt werden, von denen heute die ersten im Entstehen sind.

Besonders interessant wird die Korrelation des Proteoms mit der mRNA-Expression (dem Transkriptom) sein, die ja beide auf ganz unterschiedlichen Ebenen reguliert werden. Auch der Fluß und die Veränderung von Stoffwechselprodukten und kleinen Molekülen (dem Metabolom/Fluxosom/Physiom – die Namensgebung ist hier noch nicht endgültig erfolgt), die ihrerseits wieder vielfältige Rückkopplungen auf das Transkriptom und Proteom haben, werden in Zukunft eingehender untersucht werden. Letztendlich werden die Einblicke in das Zusammenspiel von Genom, Transkriptom, Proteom und Metabolom sicher ganz neue Erkenntnisse über die biologisch relevanten und sehr komplex vernetzten Mechanismen lebender Organismen bringen.

Eingegangen am 16. Februar 1999 [A 326]

International Edition: *Angew. Chem. Int. Ed.* **1999**, 38, 2476–2492

- [1] a) <http://pedant.mips.biochem.mpg.de/>; b) <http://www.expasy.ch/>.
- [2] a) P. Edman, *Acta Chem. Scand.* **1950**, 4, 283–290; b) F. Sanger, H. Tuppy, *Biochem. J.* **1951**, 49, 481–490.
- [3] a) P. H. O'Farrel, *J. Biol. Chem.* **1975**, 250, 4007–4021; b) J. Klose, *Humangenetik* **1975**, 26, 231–243.
- [4] a) J. Vandekerckhove, G. Bauw, M. Puype, J. Van Damme, M. Van Montagu, *Eur. J. Biochem.* **1985**, 152, 9–19; b) R. Aebersold, D. B. Teplow, L. E. Hood, S. B. Kent, *J. Biol. Chem.* **1986**, 261, 4229–4239; c) P. Madsudaira, *J. Biol. Chem.* **1987**, 262, 10035–10038; d) C. Eckerskorn, W. Mewes, H. W. Goretzki, F. Lottspeich, *Eur. J. Biochem.* **1988**, 176, 509–519; e) C. Eckerskorn, P. Jungblut, W. Mewes, J. Klose, F. Lottspeich, *Electrophoresis* **1988**, 9, 830–838; f) M. Ploug, A. L. Jensen, V. Berkholt, *Anal. Biochem.* **1989**, 181, 33–39; g) G. I. Tous, J. L. Fausnaught, O. Akinyosoye, H. Lachland, P. Winter-Cash, F. J. Vitoria, S. Stein, *Anal. Biochem.* **1989**, 179, 50–55; h) S. Nakagawa, T. Fukuda, *Anal. Biochem.* **1989**, 181, 75–78; i) C. Eckerskorn, F. Lottspeich, *Chromatographia* **1989**, 28, 92–94; j) R.

- Aebersold, J. Leavitt, L. E. Hood, S. H. Kent, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 6970–6974; k) G. Bauw, M. Van Den Bulcke, J. Van Damme, M. Puype, M. Van Montagu, J. Vandeckerkhove, *Electrophoresis* **1990**, *11*, 528–536; l) M. J. Walsh, J. McDougall, B. Wittmann-Liebold, *Biochemistry* **1988**, *27*, 6867–6876; m) P. Tempst, A. J. Link, L. R. Riviere, M. Fleming, C. Elicone, *Electrophoresis* **1990**, *11*, 537–543; n) S. D. Patterson, D. Hess, T. Youngwirth, R. Aebersold, *Anal. Biochem.* **1992**, *202*, 193–203; o) J. Fernandez, M. DeMott, D. Atherton, S. M. Mische, *Anal. Biochem.* **1992**, *201*, 255–264; p) F. Lottspeich, C. Eckerskorn, R. Grimm in *Cell Biology: A Laboratory Handbook, Vol. 3* (Hrsg.: J. E. Celis), Academic Press, Orlando, **1994**, S. 417–421.
- [5] R. M. Hewick, M. W. Hunkapiller, L. E. Hood, J. Dreyer, *J. Biol. Chem.* **1981**, *256*, 7990–7997.
- [6] a) L. Anderson, J. Seilhamer, *Electrophoresis* **1997**, *18*, 533–537; b) N. L. Anderson, N. G. Anderson, *Electrophoresis* **1998**, *19*, 1853–1861.
- [7] a) K. Biemann, S. A. Martin, *Mass Spectrom. Rev.* **1987**, *6*, 1–76; b) J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science* **1989**, *246*, 64–67; c) K. Biemann, *Annu. Rev. Biochem.* **1992**, *61*, 977–1010; d) J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Mass Spectrom. Rev.* **1990**, *9*, 37–70; e) M. Wilm, A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis, M. Mann, *Nature* **1996**, *379*, 466–469.
- [8] M. Karas, F. Hillenkamp, *Anal. Chem.* **1988**, *60*, 2299–2301.
- [9] S. Müller, T. Neumann, F. Lottspeich, *Arzneim. Forsch./Drug Res.* **1998**, *48*, 93–95.
- [10] a) D. F. Hochstrasser in *Proteome Research* (Hrsg.: M. R. Wilkins, K. L. Williams, R. D. Appel, D. F. Hochstrasser), Springer, Berlin, **1997**, S. 187–219; b) K. L. Wilkins, V. Pallini in *Proteome Research* (Hrsg.: M. R. Wilkins, K. L. Williams, R. D. Appel, D. F. Hochstrasser), Springer, Berlin, **1997**, S. 221–232.
- [11] a) B. Bjellqvist, J. C. Sanches, C. Pasquali, F. Ravier, N. Paquet, S. Frutiger, G. J. Hughes, D. Hochstrasser, *Electrophoresis* **1993**, *14*, 1375–1378; b) T. Rabilloud, C. Valette, J. J. Lawrence, *Electrophoresis* **1994**, *15*, 1552–1558; c) J. C. Sanchez, V. Rouge, M. Pisteur, F. Ravier, L. Tonella, M. Moosmayer, M. R. Wilkins, D. F. Hochstrasser, *Electrophoresis* **1997**, *18*, 324–327.
- [12] T. Rabilloud, *Electrophoresis* **1996**, *17*, 813–829.
- [13] a) B. Bjellqvist, P.-G. Righetti, E. Gianazza, A. Görg, R. Westermeyer, W. Postel, *J. Biochem. Biophys. Methods* **1982**, *6*, 317–339; b) A. Görg, W. Postel, S. Gunther, *Electrophoresis* **1988**, *9*, 351–346.
- [14] A. Görg, W. Postel, J. Weser, W. Patutchnig, H. Cleve, *Am. J. Hum. Genet.* **1985**, *37*, 922–930.
- [15] a) C. R. Merril, J. E. Joy, G. J. Creed in *Cell Biology: A Laboratory Handbook, Vol. 3* (Hrsg.: J. E. Celis), Academic Press, Orlando, **1994**, S. 281–287; b) A. Wallace, H. P. Saluz in *Cell Biology: A Laboratory Handbook, Vol. 3* (Hrsg.: J. E. Celis), Academic Press, Orlando, **1994**, S. 289–298; c) H. Blum, H. Beier, H. J. Gross, *Electrophoresis* **1987**, *8*, 93–99; d) H. M. Poehling, V. Neuhoff, *Electrophoresis* **1981**, *2*, 141–147.
- [16] a) W. Dietzel, G. Kopperschlager, E. Hofmann, *Anal. Biochem.* **1973**, *48*, 617–620; b) V. Neuhoff, R. Stamm, H. Eibl, *Electrophoresis* **1985**, *6*, 427–448.
- [17] T. H. Steinberg, R. P. Haugland, V. L. Singer, *Anal. Biochem.* **1996**, *239*, 238–245.
- [18] M. Ünlü, M. Morgan, J. S. Minden, *Electrophoresis* **1997**, *18*, 2071–1077.
- [19] P. Jackson, V. E. Urwin, C. D. Mackay, *Electrophoresis* **1988**, *9*, 330–339.
- [20] J. N. Weinstein, T. G. Myers, P. M. O'Connor, S. H. Friend, A. J. Fornace, Jr., K. W. Kohn, T. Fojo, S. E. Bates, L. V. Rubinstein, N. L. Anderson, J. K. Buolamwini, W. W. van Osdol, A. P. Monks, D. A. Scudiero, V. N. Viswanadhan, G. S. Johnson, R. E. Wittes, K. D. Paull, *Science* **1997**, *275*, 343–349.
- [21] a) C. Eckerskorn, K. Strupat, F. Hillenkamp, F. Lottspeich, *Electrophoresis* **1992**, *13*, 664–665; b) C. Eckerskorn, K. Strupat, D. Schleuder, D. F. Hochstrasser, J. C. Sanchez, F. Lottspeich, F. Hillenkamp, *Anal. Chem.* **1997**, *69*, 2888–2892; c) K. Strupat, M. Karas, F. Hillenkamp, C. Eckerskorn, F. Lottspeich, *Anal. Chem.* **1994**, *66*, 464–470; c) W. Sutton, C. H. Wheeler, J. M. Corbett, J. S. Cottrell, M. J. Dunn, *Electrophoresis* **1997**, *18*, 424–431.
- [22] a) M. Schreiner, K. Strupat, F. Lottspeich, C. Eckerskorn, *Electrophoresis* **1996**, *17*, 954–961; b) M. M. Vestling, C. Fenselau, *Anal. Chem.* **1994**, *66*, 47–477; c) J. C. Blais, P. Nagnan-Le-Meillour, G. Bolbach, J. C. Tablet, *Rapid Commun. Mass Spectrom.* **1994**, *5*, 230–237; d) S. D. Patterson, *Electrophoresis* **1993**, *16*, 1104–1114.
- [23] a) F. Hsieh, H. Wang, C. Elicone, J. Mark, S. Martin, F. Regnier, *Anal. Chem.* **1996**, *68*, 455–462; b) T. Houthaeve, H. Gausepohl, M. Mann, K. Ashman, *FEBS Lett.* **1995**, *376*, 91–94.
- [24] S. D. Patterson, R. H. Aebersold, *Electrophoresis* **1995**, *16*, 1791–1814.
- [25] B. Spengler, D. Kirsch, R. Kaufmann, E. Jaeger, *Rapid Commun. Mass Spectrom.* **1992**, *6*, 105–108.
- [26] a) K. J. Eng, A. L. McCormac, J. R. Yates III, *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976–989; b) J. R. Yates III, *Electrophoresis* **1998**, *19*, 893–900.
- [27] M. Mann, M. Wilm, *Anal. Chem.* **1994**, *66*, 4390–4399.
- [28] M. Schnölzer, P. Jedrzejewski, W. D. Lehmann, *Electrophoresis* **1996**, *17*, 945–953.
- [29] a) K. R. Jonscher, J. R. Yates, *Anal. Biochem.* **1997**, *244*, 1–15; b) R. E. March, *J. Mass Spectrom.* **1997**, *32*, 351–369.
- [30] R. Kratzer, C. Eckerskorn, M. Karas, F. Lottspeich, *Electrophoresis* **1998**, *19*, 1910–1919.
- [31] a) B. Bjellqvist, G. Hughes, C. Pasquali, N. Paquet, F. Ravier, J. C. Sanchez, S. Frutiger, D. Hochstrasser, *Electrophoresis* **1993**, *14*, 1023–1031; b) B. Bjellqvist, B. Basse, E. Olsen, J. E. Celis, *Electrophoresis* **1994**, *14*, 1023–1031.
- [32] a) C. Eckerskorn in *Bioanalytik* (Hrsg.: F. Lottspeich, H. Zorbas), Spektrum, Heidelberg, **1998**, S. 323–368; b) J. W. Metzger, C. Eckerskorn, C. Kemper, B. Behnke in *Microcharacterization of Proteins*, 2. Aufl. (Hrsg.: R. Kellner, F. Lottspeich, H. E. Meyer), WILEY-VCH, Weinheim, **1999**, S. 213–234.
- [33] a) H. E. Meyer in *Microcharacterization of Proteins*, 2. Aufl. (Hrsg.: R. Kellner, F. Lottspeich, H. E. Meyer), WILEY-VCH, Weinheim, **1999**, S. 159–175; b) A. A. Gooley, N. H. Packer in *Proteome Research* (Hrsg.: M. R. Wilkins, K. L. Williams, R. D. Appel, D. F. Hochstrasser), Springer, Berlin, **1997**, S. 65–91.