

1. Einleitung

Der Begriff der Proteomanalyse ist im Laufe der letzten Jahre zu einem Schlagwort in der Post-Genomsequenzierungsära geworden. Allerdings blieb lange Zeit unklar, was genau der Terminus „Proteom“ umfasst und worauf eine Proteomanalyse letztlich abzielt. Grund genug, den Begriff zu Beginn der vorliegenden Arbeit näher zu beleuchten.

Während das Genom die DNA-Information der Chromosomen widerspiegelt, umfasst das Proteom den quantitativen Expressionszustand sämtlicher Proteine, die durch ein Genom kodiert werden. Diese noch ein wenig unscharfe Definition des Begriffs stammt aus dem Jahr 1994 und wurde von Marc Wilkins auf der 2-D-Elektrophorese-Konferenz in Siena formuliert. Publiziert wurde sie in dieser Form erstmals 1995 in einem speziellen Biotechnologie Journal ¹. Übergreifende Verbreitung fand sie im selben Jahr durch einen Artikel Kahns im Wissenschaftsmagazin Science ².

Folgt man dieser Definition, bleibt jedoch unklar, ob alle durch die genomische Sequenz kodierten oder lediglich die tatsächlich exprimierten Gene das Proteom ausmachen: Bilden alle kodierten, also alle theoretisch exprimierbaren Gene das Proteom, dann sollte per Definition jede Proteomanalyse als Subproteomanalyse bezeichnet werden. Denn letztlich analysiert sie nur den Teil der Proteine, der zu einem bestimmten Zeitpunkt und unter bestimmten Bedingungen exprimiert ist. Bilden hingegen lediglich die exprimierten Proteine das Proteom, dann sollte in der Konsequenz nicht mehr von „dem“ Proteom, sondern von den „Proteomen“ gesprochen werden, da der Expressionszustand der Proteine zu verschiedenen Entwicklungsstadien und unter verschiedenen Umweltbedingungen differiert.

Mit anderen Worten: Während die erste Interpretation der Definition zunächst die Gesamtheit aller möglichen Proteine eines Genoms subsumiert, fasst die zweite Auslegung den Proteinexpressionszustand eines Genoms, einer Zelle, eines Gewebes oder eines Organismus zu einem bestimmten Entwicklungsstadium und unter klar definierten Bedingungen zusammen. Diese zweite Auslegung führt also zu einer Präzisierung der Definition und ermöglicht damit eine Normierung der verschiedenen Proteomzustände. Dies stellt eine wichtige Grundlage für eine Laboratorien übergreifende Vergleichbarkeit der Resultate dar.

Wie man an der Datierung der ursprünglichen Definition erkennen kann, fällt diese nicht von ungefähr in den Zeitraum der Vervollständigung der Sequenzierung des ersten Genoms ³. Bereits zu diesem Zeitpunkt zeichnete sich ab, dass die DNA-Sequenz eines

Organismus zwar die gesamte Information möglicher Gene in sich trägt, aber keine genaue Information über den Einsatz derselben mitteilt.

Es folgten weitere groß angelegte Sequenzierprojekte, wie z. B. die Sequenzierung der Modellorganismen Hefe (*Saccharomyces cerevisiae*)⁴, des Fadenwurms (*Caenorabditis elegans*)⁵, der Fruchtfliege (*Drosophila melanogaster*)⁶, des Menschen (*Homo sapiens*)^{7,8} und natürlich die der ersten Landpflanzen Ackerschmalwand (*Arabidopsis thaliana*)⁹ und Reis (*Oriza sativa*)^{10,11}.

Aber auch diese umfangreicheren Projekte lieferten nicht die erhofften Resultate bezüglich der Funktionsweise von Lebewesen. Die generierten Daten glichen eher einer großen statischen Datenbank, die alle möglichen Bausteine des Lebens beinhaltet, aber über die dynamische, räumliche und zeitliche Verwendung nicht viel verrät.

Nichtsdestotrotz waren diese Sequenzierprojekte, vor allem die Sequenzierung des menschlichen Genoms, durchaus notwendige und wichtige Schritte, da sie ein weltweites Interesse an Biotechnologien weckten und somit die kostenintensive, biotechnologische Entwicklung von Hochdurchsatzmethoden in dem stattfindendem Maße überhaupt erst ermöglichten.

Um den statischen Charakter der generierten Information zu überwinden, gingen nachfolgende Konzepte dazu über, die Gesamtheit der Genomprodukte zu erforschen. Ein Ansatz, der bereits 1981 von NG Anderson und NL Anderson in ihrem „Human Protein Index“^{12,13} gefordert wurde, zielt auf die systematische Analyse aller menschlichen Proteine ab. Doch fand dieser Ansatz seinerzeit nur wenig finanzielle Unterstützung.

Größeren Zuspruch erfuhr hingegen die Analyse des ersten vom Genom erzeugten Genprodukts – der Boten-RNA (mRNA). Die Gründe hierfür waren zum einen die Tatsache, dass Nukleinsäuren per se unkomplizierter und homogener aufgebaut sind als Proteine. Nukleinsäuren bestehen aus nur 4 variablen Bausteinen, während Proteine mit 20 unterschiedlichen Aminosäuren 20 variable Bausteine besitzen. Aus dieser Tatsache resultiert der zweite Grund für die Auswahl der mRNA als Analyseobjekt für die Postgenomphase: Ein Großteil der entwickelten molekularbiologischen Analysemethoden – wie z. B. die Klonierung von DNA-Molekülen¹⁴, die Herstellung von cDNA-Molekülen mittels reverser Transkription¹⁵⁻¹⁹ die Klonierung dieser cDNA-Moleküle in Vektoren (Bibliotheken)²⁰⁻²³ sowie die Hybridisierung und Detektion von auf Membranen immobilisierten RNA-Molekülen (Northern)²⁴ – war für die Bearbeitung von Nukleinsäuren bereits entwickelt und etabliert. Darüber hinaus gab es im Bereich der Nukleinsäuren eine Methode, die den entscheidenden Vorteil in der Analysetechnik

ausmacht: die Vervielfältigung bestimmter klonierter oder genomischer DNA-Moleküle/Abschnitte mittels Polymerase-Kettenreaktion (PCR) ^{25,26}.

In Folge dieser Entwicklungen erschienen seit 1992 regelmäßig Publikationen zu Methoden, die ermöglichten, mRNA-Expression in hochparalleler und automatisierter Weise durch RNA/DNA-Hybridisierung auf Nylonmembranen ^{27,28} oder Glasobjektträgern ²⁹⁻³¹ zu messen (Microarrays). Diese hochparallelen Analysen von dynamischen Zuständen des Genoms ermöglichten einen ersten, komplexen Einblick in die Funktionsweise des Genoms auf Transkriptionsebene. Jedoch stellte sich hierbei bald heraus, dass die Expression der mRNA-Botenmoleküle in keiner linearen Korrelation zu den eigentlichen Effektoren, den Proteinen steht ³²⁻³⁵. Dementsprechend resultiert aus der Änderung in der mRNA-Konzentration nicht zwangsläufig eine entsprechende Änderungen in der Proteinkonzentration. In letzter Konsequenz bedeutet dies, dass kein Weg an der Proteomanalyse vorbeiführt.

1.1. Methoden der Proteomanalyse

Im Gegensatz zum jungen und ausgiebig diskutierten Konzept der Proteomforschung ^{1,36} ist die zentrale, hierbei zum Einsatz kommende Methode der 2-dimensionalen Gel-Elektrophorese (2-DE) schon relativ alt.

Die 2-DE basiert auf Methoden, die Mitte der 50-er Jahre des 20. Jahrhunderts entwickelt wurden. Eine der grundlegenden Arbeiten war die Verwendung des ersten mehrdimensionalen, elektrophoretischen Trennungssystems. Dies erfolgte bereits 1956, wobei Smithies und Poulik zum Auftrennen von Serumproteinen Filterpapier in der ersten Trenndimension und Stärkegelelektrophorese in der zweiten Trenndimension verwendeten ³⁷.

Weitere Meilensteine auf dem Weg zu der noch heute genutzten 2-DE waren die Entwicklung der ersten Polyacrylamidgelsysteme durch Raymond und Weintraub ³⁸, der Einsatz der Natriumdodecylsulfat-Gelelektrophorese (SDS-PAGE) zur Bestimmung von Molekulargewichten ³⁹, die Anwendung des diskontinuierlichen Tris-Glycin-Puffersystems von Laemmli ⁴⁰, die Entwicklung der theoretischen Grundlagen zur Erzeugung von pH-Gradienten mittels amphoterer Moleküle (Carrier Ampholine) ^{41,42} und schließlich die Synthese der ersten Carrier-Ampholine, die zur Erzeugung der theoretisch errechneten pH-Gradienten nach Svensson/Rilbe notwendig sind ^{43,44}.

Auf der Grundlage all dieser herausragenden Entwicklungen präsentierten im Jahr 1975 gleich 3 Gruppen das zukünftige Konzept der 2-DE zur Trennung komplexer Proteinmischungen ⁴⁵⁻⁴⁷. Bei der 2-DE, wie sie damals vorgestellt wurde und heute in

weiten Teilen noch durchgeführt wird, erfolgt die Trennung der Proteine jeweils in 2 unabhängigen Polyacrylamid-Gelelektrophoresesystemen, die unterschiedliche Eigenschaften der Proteine für ihre Separation ausnutzen. Im ersten Trennsystem (erste Dimension) werden die Proteine in einem pH-Gradienten, der in der Regel von pH 3 bis pH 10 reicht, nach ihrer Oberflächenladung getrennt. Das heißt, die Proteine werden im elektrischen Feld auf den pH-Wert fokussiert, bei dem ihre Oberflächen-Nettoladung ausgeglichen ist und sie daher keine Mobilität mehr im elektrischen Feld zeigen.

In dem zweiten Trennsystem (zweite Dimension) werden die fokussierten Proteine mittels einer SDS-PAGE nach ihrem Molekulargewicht getrennt. Das Auflösungsvermögen der 2-DE durch diese beiden unabhängigen Trennmethode ist ausgesprochen groß und kann mehr als 10.000 Proteinspezies betragen⁴⁸.

Die so getrennten Proteine können im Folgenden mittels verschiedener Färbetechniken wie z. B. Silber-Färbung⁴⁹⁻⁵¹, Coomassie-Färbung⁵²⁻⁵⁴, Zink/Imidazol-Färbung^{55,56} oder aber Fluoreszenz-Färbungen⁵⁷ visualisiert werden.

Ursprüngliches Ziel einer jeden Färbung war es, so viele Proteine wie möglich anzufärben. Später kamen jedoch noch andere Anforderungen wie akkurate Quantifizierbarkeit (Linearität der Färbungsintensität) der gefärbten Proteinspots oder aber Kompatibilität des Farbstoffs bzw. der Färbeprozedur zu anschließenden Proteinidentifizierungsmethoden hinzu.

Hiermit ist das nächste Schlagwort gefallen: die Proteinidentifikation. Sie ist heute der zentrale Aspekt und Motor der Proteomanalyse geworden.

Während zu Beginn der 2-DE-basierten Proteomforschung zumeist unterschiedliche Proteinmuster gegeneinander verglichen wurden, gewann die Frage nach der Identifizierung der getrennten Proteine zunehmend an Bedeutung. Anfang der 80-er Jahre bediente man sich zu diesem Zweck noch hauptsächlich zweier Methoden, zum einen der so genannten Komigration, zum anderen der Immunodetektion. Bei der Komigration lässt sich ein Protein dadurch identifizieren, dass es ein identisches elektrophoretisches Migrationsverhalten zu einem bekannten Referenzprotein zeigt⁵⁸. Diese Methode vermag jedoch ein Protein nicht eindeutig zu identifizieren, da unterschiedliche Proteine ein ähnliches Migrationsverhalten zeigen können.

Bei der Immunodetektion, die auch heute noch hohe Relevanz besitzt, werden markierte Antikörper zum Nachweis bestimmter Proteine auf Blots von 2-DE-Gelen⁵⁸ eingesetzt. Diese Methode ist bei Einsatz geeigneter Antikörper wesentlich eindeutiger als die Komigrationsmethode. Des Weiteren ist sie sehr sensitiv, d. h. man kann niedrig exprimierte Proteine mit einer Kopienzahl von 100–500 Kopien pro Zelle nachweisen⁵⁹.

Die Hauptlimitierung dieser Technik liegt – neben dem hohen Arbeitsaufwand – aber nach wie vor in der Abhängigkeit von spezifischen Antikörpern, die möglichst wenig unspezifische Kreuzreaktionen zu anderen Proteinen zeigen sollten.

Ein großer technischer Fortschritt war daher die Identifizierung von Proteinen direkt aus 2-DE-Gelen durch Edman-Sequenzierung^{60,61}. Bei der Edman-Sequenzierung wird die N-terminale Aminosäure eines Proteins markiert, abgespalten und identifiziert. Dieser Prozess kann einige Duzend Male wiederholt werden, wodurch sich eine Aminosäuresequenz des jeweiligen Proteins generieren lässt. Der Hauptnachteil dieser Methode besteht im hohen Materialaufwand, der zur Identifikation eines Proteins notwendig ist. Häufig mussten jeweils Proteinspots mehrerer 2-DE-Gele gepoolt werden, um die für die Sequenzierung erforderlichen picomolaren Mengen zu erhalten. Ferner bedeutete dies, dass niedrig exprimierte Proteine, selbst durch Poolen mehrerer 2-DE-Gelspots nicht im Bereich der Nachweisgrenze dieser Methode liegen. Darüber hinaus ist der Zeit- und Arbeitsaufwand dieser Methode mit ca. 1 Aminosäure pro Stunde sehr hoch⁶².

An diesem Punkt kommen wir zur eigentliche Triebfeder der modernen Proteomforschung, dem Einsatz der Massenspektrometrie. Diese Technik ermöglichte erstmals eine schnelle und sensitive Identifikation von Proteinen aus 2-DE-Gelen. Die Massenspektrometrie, wie sie heute in der Proteomforschung Verwendung findet, hat ihre Wurzeln ebenfalls im Beginn der 50-er Jahre des 20. Jahrhunderts. Zu dieser Zeit wurden die ersten wissenschaftlichen Ergebnisse der Flugzeit-^{63,64} und der Quadrupol-Massenspektrometrie^{65,66} publiziert.

Lange Zeit konnten diese sensitiven Messmethoden jedoch nicht für die Messung von Biopolymeren wie Proteine und Nukleinsäuren eingesetzt werden, da eine wichtige Grundvoraussetzung nicht erfüllt wurde: Als wichtigstes Kriterium für die Massenspektrometrie müssen die Analytmoleküle als Ionen in der Gasphase vorliegen. Das generelle Problem bei dieser Ionisation ist es, thermisch labile hochmolekulare Biomoleküle aus Festphase oder Flüssigkeit in die Gasphase zu transferieren – und dies möglichst unbeschadet.

Die technischen Voraussetzungen hierfür wurden letztlich erst zwischen 1987–1990 in grundlegenden Publikationen veröffentlicht⁶⁷⁻⁷³.

Bei der Etablierung dieser Techniken ragte die Arbeitsgruppe um Franz Hillenkamp, bei der Entwicklung der Matrix-Assistierten-Laser-Desorption-Ionisation (MALDI) für die Flugzeit-Massenspektrometrie heraus. Bei dieser Methode, die als *energy sudden*-Methode bezeichnet wird, werden die hochmolekularen Analytmoleküle in einer

niedermolekularen kristallinen Matrix eingebettet. Diese Matrix, die häufig aus einer organischen Säure besteht, ist in der Lage, Energie aus Laserlicht zu absorbieren. Wenn man die Matrix nun mit einem gepulsten Laserstrahl beschießt, entsteht durch die lokale Energiezufuhr ein Zustand, der als *pressure pulse* bezeichnet wird und zu einem Druckgradienten gegenüber der Probenoberfläche führt. Als Folge dieses Druckgradienten expandieren die Matrixmoleküle samt der Analytmoleküle explosionsartig in die Gasphase. Hierbei können sie aufgrund der hohen Energiedichte für kurze Zeit ($t < 20$ ns) Temperaturen von über 1000 °C erreichen^{74,75}. Es wird nun vermutet, dass in dieser expandierenden Matrixwolke, Cluster aus hochgeladenen Matrix- und Analytmolekülen bestehen. Diese werden nach und nach durch eine sukzessive Neutralisation der Ladung mittels Protonierung bzw. Übertragung von Elektronen und durch Sublimationsprozesse in einfach geladenen Analytmolekülen überführt (*lucky survivor Theory*)⁷⁶. Die meisten Details und der präzise Mechanismus der Ionisation im MALDI-Prozess sind jedoch bis heute nicht ganz aufgeklärt⁷⁷.

Bei der Entwicklung der zweiten neuen Ionisationsmethode, die fast zeitgleich und in Verbindung mit dem bereits erwähnten zweiten Massenanalysator, dem Quadrupol, stattfand, ragte die Arbeitsgruppe um John Fenn heraus. Es handelt es sich hierbei um die Elektro-Spray-Ionisation (ESI)⁷³. Bei dieser Methode werden die Analytmoleküle, die in gelöster Form in einem sehr sauren Lösungsmittel vorliegen, über eine feine, elektrisch geladene Kapillare zu einem Flüssigkeitsnebel versprüht. Durch Verdunstung wird die Analyt- und Ionenkonzentration in den versprühten Tröpfchen immer größer, sodass nach mehreren *kolumbschen Explosionen*^{77,78} immer kleinere Tröpfchen entstehen. Letzten Endes enthalten diese Tropfen nur noch ein Analytmolekül, das durch die stetig anwachsende Ionenkonzentration selbst eine oder aber mehrere Ladungen angenommen hat. Ganz zum Schluss verdunstet die letzte Flüssigkeit aus dem Tropfen, sodass sich das ionisierte Analytmolekül in der Gasphase befindet.

Auch bei dieser Ionisationstechnik, die als homogenes Ionisationsverfahren bezeichnet wird, gibt es verschiedene Modelle, welche die Entstehung des einzelnen Ions in all seinen Facetten mehr oder weniger hinreichend erklären^{78,79}.

Das Ziel beider Ionisationstechniken ist letztlich, die in der Gasphase befindlichen Ionen über elektrische Felder zu beschleunigen, um dann mittels Quadrupol-Massenfilter oder einem TOF-Analysator ihr Masse-zu-Ladungsverhältnis (m/z) zu bestimmen. Da die Ladung der gebildeten Ionen ganzzahlige positive oder negative Vielfache der Ladung eines Elektrons entsprechen, können aus diesen Werten unter Berücksichtigung der Masse des Elektrons die exakten Molekülmassen der zugehörigen Analyten abgeleitet werden.

Die Messgenauigkeit beider Methoden liegt im unteren Millionstel Bereich (ppm) und bildet in Verbindung mit den geringen benötigten molaren Analytmengen die Voraussetzung für höchstgenaue Messungen und sensitive Proteinidentifikation. Beide Methoden und ihre Erfinder (Tanaka und Fenn) wurden im Jahr 2002 mit dem Nobelpreis für Chemie ausgezeichnet.

Diese hochgenauen Messmöglichkeiten von Molekülmassen lieferten schließlich die Grundlagen für eine neue Strategie zur Identifizierung von Proteinen. Die Methode wurde im Jahr 1993 von 5 Arbeitsgruppen nahezu zeitgleich veröffentlicht⁸⁰⁻⁸⁴. Faktisch handelt es sich bei dieser Methode um den so genannten Protein-Massen-Fingerabdruck (PMF). Das Prinzip ist, dass die Molekülmassen der Peptide eines Proteins, das unter Verwendung einer spezifischen Protease in distinkte Peptidfragmente (Peptid-Massen-Fingerabdruck) geschnitten wird, mit einem Massenspektrometer akkurat bestimmt werden können. Diese Massen vergleicht man dann mit den erwarteten Peptid-Massen-Fingerabdrücken eines jeden Proteins in einer Sequenzdatenbank. Ein Protein gilt hierbei als identifiziert, wenn es eine ausreichende Übereinstimmung zwischen den berechneten und den gemessenen Peptidmassen gibt. Bei genügend hoher Messgenauigkeit reicht häufig eine Übereinstimmung von nur 3–5 Peptidmassen, um ein Protein zu identifizieren⁸⁵. Diese neue Methode ermöglicht erstmals die Identifizierung von Proteinen aus 2-DE-Gelen, die in Nanogrammmengen vorliegen. Sie hat jedoch eine entscheidende Limitierung: Es lassen sich lediglich Proteine identifizieren, welche bereits in sequenzierter Form in Datenbanken vorliegen. Diese Limitierung verliert jedoch durch die Anzahl bekannter Genom- und Proteinsequenzdaten, die mit den immer weiter fortschreitenden Genomsequenzierungsprojekten (eine Übersicht kann man unter www.tigr.org finden) produziert werden, von Tag zu Tag an Bedeutung.

Zusätzlich zu den hier beschriebenen Massenspektrometern haben sich im Laufe der letzten Jahre sowohl die Analysatoren als auch die Ionisationstechniken immer weiter entwickelt. Mittlerweile lassen sich Kombinationen verschiedenster Ionisationsquellen mit unterschiedlichsten Analysatoren wählen⁸⁶⁻⁹¹, sodass man Instrumente mit unterschiedlichen Eigenschaften erwerben kann.

Nach wie vor lässt die Kombination von MALDI-Ionisation und der Flugzeit-Massenanalyse den höchsten Probendurchsatz zu.

1.2. Stand der Proteomanalyse bei Pflanzen

Pflanzenproteomanalyse im eigentlichen Sinne, das heißt Projekte, die Proteintrennungsmethoden mittels 2-DE mit anschließender Proteinidentifikation

verbinden, sind relativ selten und liegen einige Jahre hinter der Proteomforschung in Prokaryonten und einzelligen- bzw. mehrzelligen Eukaryonten⁹² zurück.

Ein Großteil der bisher durchgeführten Projekte war analytischer Art; Proteine verschiedener Gewebe oder Organismen wurden mittels 2-DE getrennt und die unterschiedlichen Proteinstreifen verglichen bzw. analysiert. Es gab und gibt nur wenige Projekte, die Identifikationstechniken wie die Edman-Sequenzierung oder die Massenspektrometrie im großen Umfang eingesetzt haben.

Zu Beginn der „Pflanzen-Proteomforschung“ lag der Hauptfokus in der Herstellung von 2-DE-basierten Proteinkarten verschiedener Pflanzenarten, -linien und -spezies⁹³⁻⁹⁵. Wie man an der Datierung der Publikationen erkennt, begann im Pflanzenbereich der Einsatz der 2-DE etwas verzögert zur Publikation der Methode im Jahre 1975.

Im folgenden Schritt versuchte man diese 2-DE-Proteinstreifen oder -Proteinkarten zu verwenden, um Protein-Marker zu identifizieren. Diese Marker sollten dazu dienen, genetische oder phylogenetische Abstände bzw. Unterschiede zwischen verschiedenen Pflanzenlinien, -arten oder aber -familien zu bestimmen⁹⁶⁻⁹⁸. Hierbei ist noch einmal darauf hinzuweisen, dass diese Marker in den meisten Fällen anonyme schwarze Punkte auf 2-DE-Karten waren, deren Proteinidentität sich nicht aufklären ließ. Bei der Feststellung dieser Marker ging es weitestgehend um die Frage, ob ein Protein in der einen Spezies vorhanden ist, während es in der anderen fehlt.

In den folgenden Jahren benutzte man 2-DE-Gele auch dazu, verschiedene Organe oder Gewebe von Pflanzen miteinander zu vergleichen⁹⁹. Dies sollte dazu dienen, Organ- oder Gewebespezifität einer Pflanze auf Proteinebene zu analysieren. Interessant war, dass die Variabilität zwischen verschiedenen Organen oder Geweben unterschiedlicher Linien starke Unterschiede aufweisen können. So zeigen z. B. die 2-DE-Muster von Mesophyllzellen (Blattzellen) zweier verschiedener Pflanzenlinien 12,6 % Unterschiede, während Bündelscheidenzellen der gleichen Linien im selben Vergleich nur um 7,6 % variieren^{93,99}.

Weitere Untersuchungen beschäftigten sich mit den Auswirkungen, die verschiedene geographische und klimatische Umwelteinflüsse auf den morphologischen und molekularen Phänotyp haben. Hierzu wurden Weizenpflanzen, abstammend von einer bestimmten Ausgangslinie, unter verschiedenen Umweltbedingungen aufgezogen und morphologisch und mittels 2-DE analysiert. Das erstaunliche Resultat war, dass die unterschiedlich gewachsenen Pflanzen neben unterschiedlichem Erscheinungsbild auch variierende 2-DE-Muster aufwiesen im Vergleich zur Ausgangslinie¹⁰⁰.

In der Folgezeit, die durch die Identifikation und Klonierung zahlreicher Mutanten geprägt war, ging man in der Pflanzen-Proteomanalytik dazu über, Mutante- und Wildtyppflanzen in differentiellen Ansätzen zu vergleichen¹⁰¹⁻¹⁰⁶. Hauptanliegen dieser Projekte war es, Zielproteine bestimmter Mutanten aufzuspüren und diese dann in einen physiologischen Kontext der Mutation zu stellen. Dieser differentielle Ansatz, kranke Pflanze gegen gesunde Pflanze, Mutante gegen Wildtyp, gestresste gegen ungestresste Pflanze zu vergleichen, und die daraus resultierenden Analysen der unterschiedlichen Proteine haben sich bis heute am stärksten etabliert. Etwa zur selben Zeit wurde auch das erste umfangreichere 2-DE-Proteom-Projekt publiziert. Ziel dieses Projekts war es, eine größere Anzahl Proteine aus dem Reis- und dem *Arabidopsis*-Proteom zu ermitteln. Zu diesem Zweck wurden verschiedene Proteinextrakte aus Pflanzengewebe wie Blättern, Stängeln oder Wurzeln präpariert und mittels 2-DE getrennt^{107,108}. Die Technik, die hierbei zur Proteinidentifikation eingesetzt wurde, war zumeist die Edman-Sequenzierung. Dieser Ansatz hat sich aber aufgrund seines hohen Arbeitsaufwands und der geringen spezifischen Anwendungsbezogenheit nicht durchgesetzt. Allerdings lag es bei den beiden erwähnten Projekten mit Sicherheit auch daran, dass mit der Edman-Sequenzierung eine falsche Wahl der Proteinidentifizierungsmethode getroffen wurde. Letztlich ließen sich von den vielen auf 2-DE-Gelen getrennten Proteinen nur ungefähr 50 unterschiedliche identifizieren. Dies lag daran, dass ein Großteil der ausgewählten Proteine N-Terminal blockiert war und sich somit der Sequenzierung entzog.

Im Folgenden ging man deshalb wieder dazu über, reine Grundlagen- und Ressourcen-Arbeiten durch – wie bereits erwähnt – gezieltere, differentielle und auf spezifische Fragestellungen gerichtete Projekte zu ersetzen¹⁰⁹⁻¹¹¹.

Die Analyse von Pflanzen unter Trockenstress zählte hierbei zu den wichtigsten Bereichen der 2-DE-Analytik¹¹²⁻¹¹⁴. Darüber hinaus lassen sich weitere kleinere und mittelgroße Projekte aufzählen, die einen bestimmten Stimulus oder eine bestimmte Mutante genauer analysieren^{103,112,115}.

Der gegenwärtig aktuellste Trend im Pflanzenproteombereich zielt auf die Analyse von subproteomenen bzw. organellen Proteomen ab. Dies hat mehrere Gründe: Zunächst sind hierbei die Proteinmischungen der 2-DE-Gele weniger komplex. Darüber hinaus sind die Proteine, die man aus diesen 2-DE-Gelen identifizieren kann, spezifischer und somit auf einen klar umrissenen Bereich eingeeengt, als dies bei Gesamtproteinextrakten der Fall ist. Hieraus resultiert, dass sich viele Arbeitsgruppen alte biochemische Aufreinigungs- und Präparationsmethoden zur subzellulären Fraktionierung zunutze machen^{116,117} und diese

in Verbindung mit neueren Techniken und Chemikalien in ihre Proteomprojekte einbauen¹¹⁸.

So beschäftigen sich z. B. einige französische Gruppen mit der Aufklärung des Plasmamembran-Proteoms¹¹⁹⁻¹²¹. Dies ist eine überaus diffizile Aufgabe, da gerade die hydrophoben Membranproteine nach wie vor als sehr schwer handhabbar in der 2-DE gelten¹²²⁻¹²⁴. Eine andere Gruppe versucht, das Proteom von *Arabidopsis thaliana*-Zellwänden zu erforschen¹²⁵.

Ein weiteres stark umforschtes Organell sind die Chloroplasten. Hier gibt es gleich mehrere Gruppen, die sich intensiv bemühen, die gesamte Proteininformation dieser Organellen aufzudecken und ihre Funktionsweisen zu erklären¹²⁶⁻¹³⁰.

Weiterhin gibt es einige Publikationen im Bereich der Proteomanalyse von Pflanzenribosomen^{131,132} und von Pflanzenmitochondrien¹³³⁻¹³⁶. Eine andere Gruppe beschäftigt sich hingegen mit dem Wurzelproteom von Pflanzen, die Stickstoff mittels bestimmter bakterieller Symbionten fixieren können¹³⁷. Es zeichnet sich also ein deutlicher Trend zur Spezialisierung und Abgrenzung der einzelnen Forschungsgruppen ab.

Dieser Trend findet seine Fortsetzung in der Analyse der nächsten Untereinheit von Proteomgruppen, den einzelnen Proteinkomplexen. Einige Gruppen versuchen gezielt, bestimmte Proteinkomplexe aus speziellen Geweben oder Organellen zu extrahieren und dann ihre Komponenten zu analysieren. Aus Chloroplasten beispielsweise wurde ein 350 kDa großer Proteasekomplex mittels Blue Native Gel Elektrophorese¹³⁸ isoliert und seine Bestandteile wurden mit MALDI-TOF-MS identifiziert¹³⁹. Eine andere Gruppe wählte einen ähnlichen Ansatz, um TIM- und TOM-Komplexe aus Mitochondrien zu analysieren¹⁴⁰.

Dieser Ansatz der spezifischen Analytik eines bestimmten Proteinkomplexes wurde anhand von Hefe (*Sacharaomyces cerevisiae*) in 2 zeitgleich publizierten Untersuchungen grundlegend verändert. Durch die Verwendung homologer Rekombination und die Möglichkeit, mittels dieser modifizierten Proteine in das Hefegenom einzubringen, gelang es, einen gänzlich unspezifischen Ansatz zu verwenden^{141,142}. Bei diesen Arbeiten wurden ca. 6200 Hefemutanten erzeugt, bei denen jeweils ein Protein so markiert und in das Hefegenom integriert wurde, dass man es samt aller daran gebundenen Proteine mittels dieser Markierung durch 2 aufeinander folgende Affinitätsaufreinigungen aus dem Proteom „herausfischen“ konnte. Auf diese Weise konnten einige tausend Proteinkomplexe identifiziert und in Korrelation zu einander gesetzt werden. Da Pflanzen in erster Linie nicht homolog, sondern willkürlich rekombinieren, ist dieser Ansatz für

höhere Pflanzen zwar ungeeignet, ließe sich aber bei Vorläuferformen wie dem Moos *Physcomitrella patens* sicherlich einsetzen¹⁴³⁻¹⁴⁵.

1.3. Zielsetzung der Arbeit

Wie in 1.2 dargestellt, liegt die Proteomforschung im Pflanzenbereich hinter der Proteomforschung im tierischen und bakteriellen Bereich zurück⁹². Dies liegt vor allem daran, dass im methodischen Bereich eine Vielzahl von ungelösten oder aber limitierenden Fragen nicht intensiv weiterverfolgt wurde. Das Ziel der vorliegenden Arbeit ist es daher, Methoden zu etablieren, die eine umfangreiche Proteomanalyse pflanzlicher Systeme ermöglichen. Zu diesem Zweck ist mit *Arabidopsis thaliana* eine Modellpflanze ausgewählt worden, anhand derer es die Methodenentwicklung durchzuführen gilt. Hierbei geht es sowohl darum, neue Methoden zu schaffen, als auch bereits existierende Methoden zu evaluieren. Die methodischen Untersuchungen betreffen dabei alle notwendigen Schritte der Proteomanalyse.

Die beiden zentralen Techniken, die hierbei Einsatz finden sollen, sind daher die 2-D-Gelelektrophorese und die MALDI-TOF-Massenspektrometrie. Die zu erarbeitenden Methoden sollen sich zu einem Konzept zusammenfügen, das in Zukunft erlaubt, Proteomforschung bei Pflanzen auf der Basis von Hochdurchsatzanalytik in automatisierter Form durchzuführen.

Die Anwendbarkeit der zu entwickelnden Methoden auf biologische Fragestellungen soll abschließend an 3 Anwendungsbeispielen aus dem Pflanzenbereich erprobt werden. Hierbei gilt es zum einen, die Hochdurchsatzanalytik auf die Analyse von *Arabidopsis thaliana*-Gewebe anzuwenden. In einem zweiten Beispiel soll die Veränderung des Proteoms des Phloemsaftes unter Trockenstress in Gurkenpflanzen beobachtet werden. Zielsetzung des dritten Beispiels ist es, das Subproteom der ribosomalen 80S Proteine aus *Arabidopsis thaliana*-Blättern zu analysieren.