

## 4 Diskussion

Ziel der vorliegenden Arbeit war es, Methoden für eine effiziente Proteomanalyse von Pflanzen zu entwickeln. Die Methodenentwicklung erstreckte sich hierfür über einen weiten Bereich und beinhaltet von der Materialgewinnung über die Proteinextraktion, die Proteintrennung, die Proteinvisualisierung bis hin zur anschließenden Proteinidentifikation ein breites methodisches Spektrum. Die Summe dieser Methoden wurde in einem zweiten Schritt in einem Automatisierungsprozess umgesetzt, sodass Proteomanalysen bei Pflanzen im Hochdurchsatz möglich sind.

### 4.1 Bewertung der entwickelten Methoden für die Proteomanalyse

#### 4.1.1 Bewertung der fraktionierten Proteinextraktionsmethode

Die fraktionierte Proteinextraktionsmethode, die im Rahmen dieser Arbeit entwickelt wurde<sup>150</sup>, stellt eine Neuentwicklung gegenüber der bislang hauptsächlich im Pflanzenbereich verwendeten TCA/Aceton-Präzipitationsmethode<sup>151</sup> dar.

Die Gründe für die bisherige Verwendung der zuletzt genannten, harschen Proteinextraktionsmethode liegen in erster Linie in dem Problem der erhöhten Aktivität von Proteasen in pflanzlichen Geweben<sup>95,189</sup>. Viele dieser Proteasen, die sich bei pflanzlichen Geweben in den Vakuolen befinden, werden beim mechanischen Zellaufschluss frei<sup>233</sup>. Um dem entgegenzuwirken, wurde die TCA/Aceton-Präzipitationsmethode entwickelt. Diese Methode bietet den Vorteil, dass – nachdem der Zellaufschluss unter flüssigem Stickstoff stattgefunden hat – die Probe in einem komplett und schnell denaturierenden Puffer (TCA/Aceton) aufgenommen wird. Diese Denaturierung, die mit der Proteinpräzipitation einhergeht, inhibiert somit die Proteasen, bevor sie aktiv werden können. Die Geschwindigkeit der Denaturierung bzw. der Inhibition der Proteasen ist hierbei von größter Bedeutung. Selbst Blattproben, die in stark denaturierendem 9 M Harnstoffpuffer nach der Extraktion aufgenommen werden, zeigen starke Degradationsanzeichen, da die Denaturierung der Proteasen nicht schnell genug beziehungsweise nicht vollständig gelingt<sup>95</sup>. Weitere Arbeiten zeigen, dass sogar in Gegenwart des anionischen Detergens SDS Proteasen bis zu einem gewissen Punkt Aktivität aufweisen können<sup>234</sup>.

Dieser zentrale Punkt, der auch im Rahmen der vorliegenden Arbeit als überaus kritisch wahrgenommen wurde, ließ sich durch die Verwendung eines starken, die meisten gängigen Proteasen inhibierenden Cocktails von Proteaseinhibitoren lösen. Die Fragilität und

Grenzwertigkeit der Proteaseninhibition wird in 3.1.2.4 beschrieben. Hier lässt sich klar erkennen, wie schnell aus einem „normalen“ 2-DE-Muster ein Degradationsmuster wird.

Die Bewertung von 2-DE-Mustern auf Degradation fällt mitunter besonders schwer, da Degradation in manchen stoffwechselaktiven Geweben zum natürlichen Erscheinungsbild des Proteininventars gehört und somit trunkierte Proteine keine Folge mangelnder Proteaseninhibierung darstellen müssen. Letztlich ist somit der Einsatz von Proteaseinhibitoren gerade bei pflanzlichen Proteinextraktionen immer eine Gratwanderung – zumal gerade durch den konzentrierten Einsatz von Proteaseinhibitoren starke Artefakte, d.h. Modifikationen von Proteinen entstehen können<sup>235</sup>.

Die im Rahmen dieser Arbeit entwickelte fraktionierte Proteinextraktionsmethode zeigt, dass sie nicht nur eine größere Anzahl von Proteinen extrahieren kann, sondern dass diese extrahierten Proteine in den unterschiedlichen Fraktionen aufgrund der Verwendung verschieden chemisch zusammengesetzter Puffer auch eine gänzlich unterschiedliche Proteinzusammensetzung aufweisen.

Hinsichtlich einer Effizienzsteigerung der extrahierbaren Proteine durch die Verwendung der fraktionierten Proteinextraktion konnte anhand zweier *Arabidopsis thaliana*-Gewebe (Blatt und Stängel) eine 3fache Steigerung der Proteinspots gezeigt werden (vgl. 3.2). Während Blattextrakte und Stängelsextrakte, die mittels der präzipitationsbasierten Extraktion erzeugt wurden, einmalig jeweils 1100 (Blatt) bzw. 1700 (Stängel) Proteinspots aufwiesen, zeigten die fraktioniert aufgearbeiteten Gewebe mit jeweils 2500 (Blatt Fraktion 1) und 2000 (Blatt Fraktion 2) bzw. 3600 (Stängel Fraktion 1) und 2200 (Stängel Fraktion 2) Proteinspots. Nach Abzug der positionsidentischen Proteinspots, die etwa 20 % der Gesamtpots ausmachen, ergab sich eine Summe von 3600 Proteinspots für Blattgewebe und 4600 Proteinspots für Stängelgewebe.

Darüber hinaus wurde neben den Fraktionen 1 und 2 zu jedem Gewebe eine Fraktion 3, die mittels SDS-PAGE aufgetrennt wird, und eine Fraktion 4, die sowohl als 2-DE-Gel als auch per SDS-PAGE aufgetrennt werden kann, hergestellt.

Zusammenfassend gilt es also festzuhalten, dass neben der erheblichen Steigerung der Proteinspotanzahl in Fraktionen 1 und 2 auch noch weitere, für 2-DE-Gele eventuell problematische Proteine (stark hydrophobe Proteine) durch die Erweiterung der Extraktion auf Fraktion 3 und 4 aufgetrennt werden konnten (vgl. 3.8.1.2).

## 4.1.2 Bewertung des 2-DE-Systems

Die zweidimensionale Gelelektrophorese – auf der Basis von mobilen Carrier Ampholinen zur Erzeugung von pH-Gradienten, wie sie 1975 von Klose und O'Farrell entwickelt wurde<sup>45,47</sup> – ist gegenwärtig die einzige Technik, die bis zu 10000 Proteinspots in einem Experiment auflösen kann<sup>48</sup>. Darüber hinaus ist sie selbst nach der Entwicklung von Carrier-Ampholin-unabhängigen, stabilen pH-Gradienten<sup>155</sup> und deren Weiterentwicklung<sup>192,236</sup> eine der am häufigsten diskutierten Techniken. Regelmäßig erscheinen Übersichtsartikel zu diesem Thema, in denen Stärken und Schwächen eingehend analysiert werden<sup>123,219,237-244</sup>. Im Folgenden sollen einige der häufig diskutierten Punkte anhand des in dieser Arbeit verwendeten Systems und der damit erzielten Resultate erörtert werden.

### 4.1.2.1 Carrier Ampholine oder Immobililine?

Wie in 3.3.1 experimentell gezeigt wurde, sind 2-DE-Gele, die mittels der beiden zur Zeit gängigen isoelektrischen Fokussiertechniken (CA und IPG) hergestellt wurden, kaum miteinander vergleichbar. Dieses Resultat entspricht den Ergebnissen von Nawrocki et al.<sup>245</sup> – Ergebnisse, die letztendlich eine Kompatibilität der beiden Systems zueinander ausschließen. In der Konsequenz bedeutet dies: Entweder werden CA-basierte 2-DE-Gele<sup>246</sup> verwendet oder aber man bedient sich der IPG-basierten 2-DE-Technologie<sup>192,247</sup>.

Die beiden für eine solche Entscheidung relevanten Hauptkriterien lauten einerseits „Praktikabilität“, also Handhabbarkeit der Gele und Durchführung von reproduzierbaren Experimenten, und andererseits „Qualität der Ergebnisse“. Diese beiden Kriterien in Bezug auf die beiden IEF-Systeme zu evaluieren ist jedoch nicht ganz so einfach, wie es auf den ersten Blick scheinen mag. Denn gerade das Kriterium „Qualität“ hängt in erster Linie von der Erfahrung des Experimentators ab. Aus diesem Grunde soll bei der Erörterung zunächst auf die Praktikabilität der 2-DE-Gele eingegangen werden.

IPG-Gele, mit ihren immobilisierten pH-Gradienten, können inklusive sämtlicher Puffer und der dazugehörigen technischen Ausstattung bei mehreren Anbietern kommerziell erworben werden. Die Qualität der einzelnen Gele ist mehr oder minder gleichbleibend und aufgrund der auf Plastikträgern fixierten IEF-Polyacrylamidgelen lassen sie sich sehr gut handhaben. Im Gegensatz dazu müssen sämtliche Lösungen und Puffer für eine CA-basierte IEF selbst hergestellt werden, da es zur Zeit keinen kommerziellen Anbieter derselben gibt. In einem spezialisierten Proteomics-Labor kann dies insofern von Vorteil sein, als sich somit Einfluss auf Qualität und Frische der hergestellten Lösungen nehmen lässt. Doch bedeutet es auch einen bei weitem höheren Arbeitsaufwand.

Ein weiterer Unterschied zwischen CA-basierten und IPG-basierten IEF-Gelen besteht darin, dass die CA-basierten Gele nicht selbsttragend sind. Dies bedeutet, dass diese Gele in Glaskapillaren unterschiedlicher Stärke gegossen werden. Später erfolgt die IEF in Glaskapillaren, um abschließend das IEF-Gel aus der Kapillare auf das SDS-PAGE-Gel zu transferieren. Gerade dieser Transferschritt geht nicht selten mit einer Deformation oder einer Fraktion des IEF-Gels einher – Störungen, die zu artifiziellen Änderungen der resultierenden 2-DE-Muster führen. Diese Gefahr wurde bei IPG-Gelen elegant durch die kovalente Bindung des Gels an einen Plastikträger umgangen. Mit anderen Worten: Es steht einem ein stabiler, in der Polyacrylamid-Matrix fixierter pH-Gradient zur Verfügung; darüber hinaus ist das gesamte Polyacrylamidgel auf einem Plastikträger fixiert, der eine Deformation des Gels verhindert. Hieraus resultiert, dass die „Praktikabilität“, also die Handhabung der IPG-Gele bei weitem einfacher und somit indirekt auch reproduzierbarer mit dem IPG-System sein sollte.

Die „Qualität“ der resultierenden 2-DE-Gele hängt, wie bereits erwähnt, von der Erfahrung des Experimentators ab. Alles in allem lässt sich jedoch feststellen, dass der direkte Vergleich der CA-basierten IEF mit dem IPG-System sehr klar zu Gunsten des CA-basierten Systems ausfällt. Zum einen zeigen die Proteinspots eine sauberere Trennung, die auf einer schärferen Fokussierung beruht. Zum anderen ist die Anzahl der Proteinspots, die sich aus ein- und demselben Proteinextrakt anhand beider Systeme auftrennen lässt, bei CA-Gelen weitaus höher.

Hiefür gibt es 2 Erklärungen: Erstens erhöhen CA die Löslichkeit von Proteinen im IEF-Gel<sup>248</sup>. Da Proteine aufgrund des Verlustes der Nettoladung während der IEF zum Präzipitieren<sup>249</sup> tendieren, helfen CA, dieser Präzipitation entgegenzuwirken und das Protein in Lösung zu halten. Als Konsequenz aus diesem Befund wurden so genannte „*Mixed bed*“- oder Hybrid-IEF-Gele entwickelt, bei denen IPG-Gelen über den IEF-Laufpuffer etwa 1 % an CA zugefügt wird<sup>248,250</sup>. Die zweite Erklärung der erhöhten Proteinspotzahl aus CA-basierten IEF-Gelen beruht auf der bereits erwähnten, nicht-selbsttragenden Natur der CA-Gele. Zwar ist dies für die Handhabung der IEF-Gele von Nachteil, bedeutet aber einen Vorteil für die Qualität der Ergebnisse. Dies liegt zunächst daran, dass CA-basierte IEF-Gele bei der Umpufferung aus dem harnstoffhaltigen IEF-Puffer der Fokussierung in den SDS-Puffer für die folgende PAGE rundherum zugänglich sind. CA-Gele müssen dementsprechend nicht lange equilibriert werden, bevor sie sich auf das SDS-PAGE-Gel transferieren lassen. Wohingegen Proteinverluste während der zweimaligen 10minütigen Equilibrierung von IPG-Gelen gerade bei kleineren bzw. gut wasserlöslichen Proteinen zu erwarten sind<sup>251</sup>. Neben den Verlusten bei der Equilibrierung verliert man einen Teil der Proteine auch dadurch, dass es nicht gelingt,

sie allesamt aus dem IEF-Gel zu lösen. Auf diese Weise verliert man vorsichtigen Schätzungen zufolge zwischen 11 und 16 % der fokussierten Proteine bei CA-basierten IEF-Gelen<sup>48</sup>. Da sich diese Verluste mit der Zugänglichkeit des SDS-Puffers zu den Proteinen und darüber hinaus mit Gelmatrix-Protein-Wechselwirkungen erklären lassen<sup>252</sup>, zeigen dementsprechend auch hier CA-Gele klare Vorteile.

Zusammenfassend kann an dieser Stelle also festgehalten werden, dass IPG-Gele eindeutige Vorteile in der Praktikabilität aufweisen – eine Erleichterung insbesondere für ungeübte Experimentatoren. Hingegen ermöglichen CA-basierte Gele letztendlich qualitativ hochwertigere, d.h. reichhaltigere und besser aufgelöste 2-DE-Gele. Der Aspekt der Reproduzierbarkeit, dem ebenfalls ein hoher Stellenwert zukommt, soll an dieser Stelle nicht diskutiert werden, da mit beiden Systemen, wie aus vielen Publikationen ersichtlich wurde, reproduzierbare 2-DE-Gele hergestellt werden können<sup>48,253-256</sup>.

#### **4.1.2.2 Riesengele, große Gele, viele Gele, keine Gele**

Wie in 3.3.2 gezeigt werden konnte, ist es mitunter von Vorteil, sehr große Gele zu verwenden. Dies gilt insbesondere, wenn komplexe Proteinextrakte – also Extrakte mit einer sehr hohen Anzahl unterschiedlicher Proteinspezies – getrennt werden sollen. Ebenso gilt dies für Proteinextrakte, bei denen ein großer Anteil von Proteinspezies aufgrund Ihrer pI-Werte jenseits des pH-Gradienten (pH 3–10) des IEF-Gels liegt. So ließ sich im Verlauf dieser Arbeit zeigen, dass beispielsweise ribosomale *Arabidopsis thaliana*-Proteine, von denen viele einen pI jenseits von pH 10 aufweisen<sup>131,132</sup>, in einem 40-cm-IEF-Gel aufgetrennt werden können, während dies in 20 cm-Gelen nicht möglich war (vgl. Abb. 18c). Dies war allerdings nicht mittels einer Gleichgewichtsfokussierung möglich, sondern nur unter Verwendung der so genannten Nicht-Gleichgewichtsfokussierung<sup>194</sup>. Hierbei erfolgt aufgrund einer Laufzeitverkürzung eine ausreichende Proteintrennung im pH-Gradienten, ohne dass die Proteine ihren pI erreichen. Je größer ein IEF-Gel bei solch einer Nicht-Gleichgewichtsfokussierung ist, desto leichter fällt es, sowohl Proteine mit einem pI innerhalb des pH-Gradienten hinreichend zu trennen als auch Proteine mit einem pI jenseits des pH-Gradienten im Gel noch darzustellen.

Der zweite Fall, der den Einsatz von Riesengelen sinnvoll erscheinen lässt, tritt auf, wenn es extrem komplexe Proteinmischungen aufzutrennen gilt. Hierbei wird der Anteil an comigrierenden Proteinen – d.h. Proteinspots, die aus mehr als einer Proteinspezies bestehen – aufgrund der vergrößerten Trennfläche stark reduziert<sup>48,257</sup>. Gerade diese Entzerrung von sich überlagernden Proteinspots gestaltet die digitale Auswertung von 2-DE-Gelen mittels

Bildauswertungsprogrammen nach wie vor überaus arbeitsaufwendig. Hier ist jedoch einschränkend hinzuzufügen, dass die Häufigkeit, mit der man so komplexe, in einem 20 cm-Gel nicht mehr gut darstellbare Proteinmischungen auftrennt, bei fraktionierten Proteinextrakten eigentlich relativ gering ist. Liegen nichtsdestoweniger solche Proben vor, so bieten Riesengele eine Möglichkeit. Eine weitere – vielleicht elegantere – Alternative besteht darin, die Komplexität der Probe soweit zu reduzieren, dass sie wieder problemlos auf einem gängigen Gelformat darstellbar ist. Dies kann anhand einer Vorfraktionierung des Gewebes oder der Zelle erfolgen <sup>118</sup>. Eine weitere Möglichkeit besteht in einer spezifischen Fraktionierung des eigentlichen Proteinextrakts mittels Immunopräzipitation <sup>258</sup> beziehungsweise in einer unspezifischeren Fraktionierung durch die Auftrennung der Proteinprobe in Fraktionen mit distinktem pH <sup>259,260</sup>. Durch jede vorherige Probenfraktionierung erhöht sich zwar die Zahl an 2-DE-Gelen, doch verkürzt sich im Gegenzug auch die Zeit für die Auswertung pro Gel. Für eine solche Vorgehensweise, bei der eine pH-Vorfraktionierung der Proteinextrakte stattfindet, zeigt sich das IPG-basierte IEF-System erneut als vorteilig, da man bei den IPG-Gelen zu jeder pH-Proteinfraktion das exakt passende IPG-Gel verwenden kann <sup>257</sup>.

Eine gänzlich gelbteiler Ansatz der Proteomanalyse wurde im Laufe der letzten Jahre in der Arbeitsgruppe von John Yates entwickelt. Hierbei handelt es sich um die multidimensionale nanochromatographische Proteintrennung, die unter dem Akronym MudPIT (Multi-dimensionalprotein-identification technique) etabliert und für unterschiedliche Organismen erprobt wurde <sup>261-265</sup>.

Bei dieser Methode werden Proteinextrakte hergestellt und mittels einer zusammengesetzten nanoHPLC-Säule aufgetrennt. Die erste Hälfte dieser Säule besteht aus einem Säulenmaterial, das eine Kationenaustauschchromatographie ermöglicht, während die zweite Säulenhälfte aus C18-Material besteht, das eine Umkehrphasenchromatographie zulässt: Die extrahierten Proteine werden in einem ersten Schritt mittels eines proteolytischen Verdaus in Peptide zerlegt, die im Folgenden auf die stationäre Phase der Kationenaustauschchromatographiesäule geladen werden. Anschließend werden anhand von stufenweise steigenden Kationenkonzentrationen in der mobilen Phase distinkte Peptidfraktionen von der Kationensäule auf die C18-Säule eluiert. Von dieser C18-Säule werden sie im nächsten Schritt mit einem Gradienten steigender Acetonitrilkonzentration eluiert und über eine Nanospraynadel in ein Ionenfallen-MS geleitet. Hier werden die eluierten Peptide fragmentiert und sodann anhand der gewonnenen Sequenzinformationen identifiziert. Der große Vorteil dieser Methode besteht darin, dass sich die Datengenerierung sehr gut automatisieren lässt <sup>263,266</sup>. Als Ergebnis dieser Methode konnten 1484

unterschiedliche Proteine aus Hefe und 2302 unterschiedliche Proteine aus Reis detektiert und identifiziert werden <sup>263,265</sup> – Werte, die sich bis dahin nicht einmal annähernd unter Verwendung von 2-DE-basierten Proteomprojekten erzielen ließen. Doch hat diese Methode nach wie vor den großen Nachteil, dass sie kaum quantitative Informationen über die identifizierten Proteine liefert. Man weiß also nur um die Existenz eines bestimmten Proteins, ohne jedoch eine Aussage über seine Häufigkeit in der jeweiligen Probe treffen zu können. Eben dies schränkt die Methode gerade beim Einsatz in differentiellen Proteomprojekten gegenüber 2-DE-basierten Ansätzen stark ein, da die quantitative Information ein immanenter Parameter eines jeden 2-DE-Gels ist. So lässt sich aus einem Gelspot aufgrund seiner Fläche und Färbeintensität die relative Proteinmenge in erster Näherung gut ermitteln. Inwieweit dieser quantitative Wert richtig ist, wird in 4.1.3 näher diskutiert.

Ein Punkt, der für eine Koexistenz der 2-DE-basierten Proteintrennung und der chromatographiebasierten Technik spricht, ergab sich bei einer Analyse des Reisproteoms, die als vergleichende Analyse der 2-DE-Methode und der MudPIT-Methode angelegt war. Hierbei wurde gezeigt, dass von 562 identifizierten Proteine aus 2-DE-Gelen etwa 30 % keine Übereinstimmung zu den 2302 identifizierten Proteinen durch MudPIT identifizierten Proteinen aufweisen <sup>265</sup>. Es liegt also der Schluss nahe, dass diese beiden hoch komplementär zueinander sind.

#### **4.1.3 Bewertung der Proteinfärbungen für die 2-DE**

Eine Bewertung der im Rahmen dieser Arbeit verwendeten Färbungen hat viele Facetten, die es zu beleuchten gilt. Denn hierbei handelte es sich nicht um eine Färbung für sämtliche Anwendungen, sondern letztlich um jeweils die beste Färbung für eine spezifische Anwendung. So sollte die ideale Färbung auf der einen Seite überaus sensitiv sein, einen möglichst breiten dynamischen Bereich für eine Quantifikation abdecken und letztlich so wenig modifizierend wie möglich sein, um eine gute Massenspektrometrietauglichkeit aufzuweisen.

Unter den etablierten Färbungen (Silber, Coomassie, Zink/Imidazol) gibt es keine, die all diesen Anforderungen genügt. So wurde eine Silber-Färbung durchgeführt, wenn man möglichst hohe Sensitivität erreichen wollte <sup>267</sup>. Zum Zweck der Proteinidentifikation griff man auf Coomassie-Färbung <sup>216</sup> zurück. Und galt es, eine akkurate Quantifikation von Proteinen über einen großen dynamischen Bereich durchzuführen, so schien keine dieser Färbungen hinreichend tauglich zu sein <sup>200</sup>. Mittlerweile hat sich die Sicht der Dinge ein wenig verändert; es gibt sowohl sensitive sowie Massenspektrometrie taugliche Silber- <sup>157,268</sup>

als auch Zink/Imidazol-Färbungen <sup>160</sup>, die aber nach wie vor für quantitative Analysen unzureichend sind. Darüber hinaus etablieren sich zur Zeit Fluoreszenz-Färbungen, die verheißungsvoll erscheinen, da sie nominell alle gewünschten Parameter vereinen <sup>57,164</sup>.

In 3.4 wurde ein Großteil der zur Zeit gängigen Färbetechniken getestet und anhand von Parametern wie Sensitivität, Arbeitsaufwendigkeit und Kostspieligkeit verglichen. Hieraus resultierte eine Untergruppe von besonders geeigneten Färbungen, die wiederum in 3.6.5 auf ihre Massenspektrometrietauglichkeit überprüft wurden.

Die Ergebnisse dieses Vergleichs, zusammengefasst in den Tabellen 6 und 7, spiegeln alles in allem die bereits aus der Literatur bekannten Trends wider. So zeigt z. B. die Silber-Färbung die höchste Sensitivität <sup>197</sup>, während die Coomassie-Färbung bei dem direkten Vergleich die beste Massenspektrometrietauglichkeit zeigt <sup>216</sup>. Einschränkend ist jedoch darauf hinzuweisen, dass bei dem Vergleich der Massenspektrometrietauglichkeit zwar eine klare Reihenfolge der Tauglichkeit gefunden wurde, die absoluten Prozentzahlen jedoch nur eine Momentaufnahme des eigentlichen Experiments widerspiegeln. So bedeuten die Ergebnisse also nicht, dass sich mit Coomassie-Färbungen 53 % der Proteinspots identifizieren lassen, während ein Silber gefärbtes Gel lediglich 30 % Identifikationen ermöglicht. Die absoluten Zahlen können letztlich von Experiment zu Experiment schwanken, was zum größten Teil in der multifaktoriellen Natur der Analyseprozedur begründet liegt. Die Abstände in der Massenspektrometrietauglichkeit zwischen den einzelnen Färbungen hingegen sind relativ stabil. So konnten z. B. aus einem Coomassie gefärbten 2-DE-Gel etwa 75 % der ausgeschnittenen Proteinspots identifiziert werden, während sich in einem parallel angesetzten In-Gel-Verdau aus einem Silber gefärbten 2-DE-Gel desselben Gewebes etwa 50 % der ausgeschnittenen Proteine identifizieren ließen. Eben diese Schwankungen der absoluten Zahlen machen es so schwierig, eine klare Reihenfolge der verwendeten Färbungen aufzustellen: Denn mit steigender bzw. sinkender Identifikationsanzahl verändert sich auch der eingeführte Wert der theoretisch identifizierbaren Proteine aus Tabelle 7.

2 eindeutige Ergebnisse können anhand des Vergleichs der unterschiedlichen Färbungen herausgestellt werden: Die verwendete kolloidale Coomassie-G250-Färbung zeichnet sich zwar durch ihre sehr gute Massenspektrometrietauglichkeit aus, lässt jedoch mit einer Färbezeit von bis zu 5 Tagen viel zu viel Vorlaufzeit zwischen Experiment und der Auswertung verstreichen. Darüber hinaus hat sich gezeigt, dass Fluoreszenz-Färbungen ein enormes Potential für die Proteomforschung haben könnten, für Hochdurchsatzanalysen von Riesengelen jedoch eine allzu kostspielige Lösung darstellen.

Dementsprechend kann zur Zeit die Eigensynthese des rutheniumbasierten Fluoreszenzfarbstoffs von Rabilloud et al. <sup>165</sup> als kostengünstigste und Erfolg versprechendste



Lösung gewertet werden. Dieser Farbstoff charakterisiert sich durch seine hohe Sensitivität und lässt sich durch eine einfache Eigensynthese kostengünstig herstellen. So belaufen sich die reinen Farbstoffkosten für ein Fluoreszenz gefärbtes Sypro<sup>®</sup>-Ruby-Gel auf etwa 70 Euro pro Großgel. Im Vergleich dazu belaufen sich die Farbstoffkosten für die Färbung desselben Gels mit dem rutheniumbasierten Fluoreszenzfarbstoff auf nur 3 Cents<sup>165</sup>.

Der einzige Nachteil bei der Verwendung des rutheniumbasierten Fluoreszenzfarbstoffs besteht in der Notwendigkeit, sich einmalig einen relativ teuren Laserscanner zum Detektieren der Proteinspots zulegen zu müssen, um den Vorteil dieser Färbung vollständig nutzen zu können.

#### **4.1.4 Anmerkungen zur Etablierung des In-Gel-Verdaus**

Wie bereits dargelegt, ist der Weg, den ein Protein aus einem Gewebe bis hin zur Proteinidentifikation im Massenspektrometer zurücklegen muss, lang und wird von vielen Faktoren beeinflusst. Allein die Wahl der Färbung kann sich erheblich auf den Ausgang einer möglichen Identifikation auswirken. Weitere Komponenten, die eine erfolgreiche von einer erfolglosen Identifikation trennen, finden sich beispielsweise im Ablauf des In-Gel-Verdaus. Hierbei können zunächst irrelevant erscheinende Faktoren wie Beschaffenheit und Häufigkeit von Waschschritten, die Art der Farbstoffentfärbung, die Konzentration des verwendeten proteolytischen Enzyms oder aber die Länge des tryptischen Verdaus an sich großen Einfluss auf das endgültige Ergebnis zeigen. Viele dieser Parameter wurden im Rahmen dieser Arbeit evaluiert und so optimiert, dass sie gute Ergebnisse lieferten. Weitere Parameter, die sich auf die Ergebnisse der Proteinidentifikation aus Gelspots auswirken, liegen in der Matrixpräparation, der folgenden Probenapplikation und der abschließenden Matrixspülung begründet.

Alles in allem lässt sich viel Zeit damit verbringen, sämtliche veränderbaren Parameter, die der Identifikation eines Proteins zugrunde liegen, in der optimalen Kombination zusammenzustellen. Das Problem bei der Optimierung von komplexen Protokollen ist jedoch, dass Veränderungen des experimentellen Systems oftmals zu starken Funktionseinbußen führen können. So zeigte sich etwa, dass das zu Beginn der Arbeit anhand von Standardproteinen etablierte minimale In-Gel-Verdauprotokoll nur mäßige Ergebnisse bei der Anwendung auf Proteinspots aus 2-DE-Gelen brachte. Erst durch Readaption des Protokolls an die neue experimentelle Situation konnten erneut gute Resultate erzielt werden. Diese Readaptionen waren zum Teil gegenläufig zu den Befunden aus den vorhergegangenen Optimierungen.

Zusammenfassend gilt es festzustellen, dass das In-Gel-Verdauprotokoll seine finale Form erst dann annehmen kann, wenn die einzelnen Komponenten des experimentellen Systems wie die Festlegung auf ein Set von Färbungen, die Festlegung auf eine bestimmte MALDI-Matrix und die damit verbundenen Waschprozeduren bereits fest etabliert wurden. Rückwirkend ist festzuhalten, dass die Etablierung des In-Gel-Verdauprotokolls, vor allem in Kombination mit der 96 MTP-Automatisierung, die zeitaufwendigste und kleinschrittigste Anforderung der im Rahmen dieser Arbeit etablierten Methoden war.

#### 4.1.5 Möglichkeiten und Grenzen der MALDI-TOF-MS

Die MALDI-TOF-MS gilt, wie bereits in vielen Übersichtsartikeln beschrieben, als Methode der Wahl, sofern eine schnelle Hochdurchsatzanalytik von Proteinspots aus 2-DE-Gelen geplant ist <sup>269</sup>. Die für die Messung eines Peptidmassenfingerabdrucks <sup>80-84</sup> bei einer MALDI-Probe benötigte Zeit beträgt momentan ca. 20 Sekunden <sup>175</sup>, während bei elektrospraybasierten Methoden nach wie vor von 20–30 Minuten pro Probe auszugehen ist <sup>266</sup>. Weitere Vorteile der MALDI-TOF-MS sind zum einen ihre extrem hohe, im niedrigen attomolarem Bereich liegende Sensitivität <sup>168</sup>, ihre hohe Toleranz gegen Puffer, Salze und andere Kontaminationen <sup>209</sup> und die Möglichkeit der zeitlichen Entkopplung von Messung und Analyse aufgrund der Einbettung der Proteinprobe in eine feste Matrix <sup>270</sup>.

Weiterhin ist die Auswertung der durch MALDI-MS erzeugten Massenspektren, die vorwiegend durch einfach positiv geladene Ionen bestimmt werden <sup>76</sup>, im Vergleich zu den häufig mehrfachgeladenen ESI-Massenspektren wesentlich einfacher <sup>269</sup>.

Als Nachteil der ausschließlich MALDI-TOF-MS-basierten Proteomanalyse zählt, dass man in weiten Teilen auf die Generierung und Auswertung von Peptidmassenfingerabdrücken beschränkt ist. Die Generierung von Sequenzdaten, die bei der Identifikation von Proteinen mittels ESI-Quadrupolgeräten oder aber ESI-Ionenfallen im Rahmen der Standardanalyse erzeugt werden <sup>271</sup>, müssen bei herkömmlichen MALDI-TOF-Reflektron-Massenspektrometern <sup>272</sup> nach wie vor durch zeitaufwendige Erzeugung und Analyse von PSD-Massenspektren (*post source decay*) generiert werden <sup>273,274</sup>. Bei der MALDI-PSD-Analyse werden die mit einer Häufigkeit von bis zu 50 % anfallenden Fragmentationen der Peptide gemessen <sup>275</sup>. Trotz des hohen prozentualen Anteils der Fragmentationen und trotz der Entwicklung von Peptidderivatisierungsmethoden, die eine präferentielle Generierung der Peptidefragmentationen unterstützen sollen <sup>276,277</sup>, muss nach wie vor mit mindesten 20–30 Minuten Messzeit pro PSD-Massenspektrum auf einem konventionellen MALDI-Massenspektrometer gerechnet werden. Und hierbei ist noch nicht gewährleistet, dass das

Massenspektrum des fragmentierten Peptids tatsächlich auswertbare Sequenzinformationen liefert<sup>278</sup>.

Wie bereits dargelegt, wird diese Art der Sequenzinformationen bei ESI-basierten Massenspektrometern bei der Standardmessung erzeugt. Dieser Art der Proteinsequenzdatengenerierung in Verbindung mit der ständig steigende Zahl von EST-Datenbanken (*expressed sequence tags*), die Sequenzen von mRNA-Moleküle sammeln<sup>279</sup>, ermöglichen eine zusätzliche Identifikationsstrategie. Diese neue Strategie basiert nicht mehr auf der Analyse von Peptidmassenfingerabdrücken tryptisch verdauter Proteine, sondern verwendet die aufgrund der Peptidfragmentierung gewonnenen Sequenzinformationen der nach wie vor tryptisch zu verdauenden Proteine zur Datenbanksuche<sup>280,281</sup>. Dank dieser Strategie genügt es bei manchen Proteinen, einen einzigen Sequenzabschnitt zu erzeugen, um es mit ausreichender Wahrscheinlichkeit identifizieren zu können<sup>282</sup>.

Diese Zweiteilung der Identifikationsstrategie resultierte in der Entwicklung von so genannten Hybrid-Massenspektrometern, die Charakteristika von MALDI-TOF-Massenspektrometern und von ESI-Massenspektrometern kombinieren. So wurde z. B. eine ESI-Ionenquelle mit einem TOF-Analysator verknüpft<sup>86</sup> – und es entstand das unter dem Namen Q-TOF bekannte Instrument, das sich überaus erfolgreich in der Proteomwelt verbreitete<sup>88,283</sup>, ohne jedoch das Problem des relativ niedrigen Probendurchsatzes wirklich zu lösen. Eine weitere Konstruktion, die einer Lösung dieses Problems näher kommt, war die Entwicklung einer MALDI-Ionenfalle, welche die schnelle Ionisation von MALDI mit den Peptidfragmentierungsvorzügen einer Ionenfalle verbindet<sup>91,284</sup>. Bei diesen Geräten bestehen aber nach wie vor Probleme in der Effizienz des Ionentransfers von der Ionenquelle in den Massenanalysator<sup>285</sup>.

Ein neuer MALDI-Gerättyp scheint dieses Problem jetzt aber weitestgehend behoben zu haben: Es handelt sich um die Entwicklung so genannter TOF/TOF-Instrumente, die anstelle einer einmaligen Ionenbeschleunigung in der Ionenquelle die Möglichkeit einer zweiten Ionenbeschleunigung in der Ionendriftzone besitzen. Dieser zweiten Beschleunigungsphase lässt sich eine Kollisionszelle<sup>286</sup> vorschalten, die Peptide bestimmter Massen selektiv fragmentieren kann und somit neben der Aufnahme von Peptidmassenfingerabdrücken auch die Aufnahme von Sequenzinformation bestimmter Peptide ermöglicht<sup>90</sup>. So steht also zu vermuten, dass diese Art von Massenspektrometer die Zukunft von Proteomprojekten mitbestimmen wird.

#### 4.1.6 Automatisierung der verwendeten Methoden

Wie im Rahmen der vorliegenden Arbeit gezeigt werden konnte, birgt die Automatisierung von Prozessen einige Komplikationen. Hierbei sei auf die Etablierung der semiautomatischen In-Gel-Verdauprozedur in 3.6.4 hingewiesen, bei der durch die Umstellung des Verdauprotokolls von 0,5 ml Eppendorfgefäßen auf 96er MTPs eine erneute Umstellung des Protokolls notwendig wurde. Dass Automatisierungen im Rahmen von Proteomprojekten unumgänglich sind, ist selbstverständlich. Ein Problem besteht jedoch darin, dass sich nicht sämtliche Prozesse der 2-DE-basierten Proteomanalyse hinreichend automatisieren lassen<sup>287-289</sup>. Hierzu zählt insbesondere die Proteinextraktion, die nach wie vor ein manueller Prozess ist. Darüber hinaus ist letztlich die gesamte 2-DE-Prozedur ein Bereich, der sich schwerlich vollständig automatisieren lässt. Angefangen mit dem Transfer des IEF-Gels auf die SDS-PAGE-Gele bis hin zur abschließenden Färbung erfordern alle Schritte personelle Durchführung. Darüber hinaus ist die eigentliche Bildauswertung, die anhand von Softwareprogrammen<sup>167,290</sup> durchgeführt werden kann, nach wie vor nicht wirklich automatisch und erfordert zeitaufwendige, manuelle Editierung. Alles in allem ist also festzuhalten, dass die Hauptlimitierung der Automatisierung zur Zeit noch auf Seiten der 2-DE liegt.

Hingegen lässt sich die MALDI-Probenpräparation (vgl. 3.6.1) relativ gut automatisieren – zumal es momentan mehrere kommerzielle Lösungen gibt, die eine robotikbasierte MALDI-Probenpräparation, inklusive In-Gel-Verdau und Probenapplikation auf das MALDI-Target anbieten. Darüber hinaus vereinfacht die Einführung des vorstrukturierten *Anchor Chip*-Probentellers die Automatisierbarkeit der MALDI-Messung. Hierdurch ließ sich das lange Zeit bestehende Problem einer ortgenauen und homogenen Matrixapplikation auf den MALDI-Probenteller auf sehr elegante Weise lösen<sup>169</sup>. Auch die eigentliche MALDI-MS-Messung konnte erfolgreich automatisiert werden. Dies erfolgte durch die Einführung von Algorithmen, die eine Fuzzy-Logik-basierte Datenakquirierung gewährleisten. Diese Programme sind in der Lage, Messparameter wie Laserintensität anhand der Qualität des zuletzt erzeugten Massenspektrums zu regulieren und automatisch optimiert anzupassen<sup>291</sup>. Dementsprechend lässt sich an dieser Stelle zusammenfassen, dass ein Großteil der zur Automatisierung nach der 2-DE-Gel-Herstellung notwendigen technischen Entwicklungen bereits etabliert sind<sup>288,289</sup>.

## **4.2 Bedeutung der aus den Anwendungsbeispielen gewonnenen Resultate für Pflanzenproteomprojekte**

Nachdem bislang in erster Linie die technischen Aspekte der in dieser Arbeit entwickelten und evaluierten Methoden diskutiert wurden, soll im Folgenden auf die biologisch-inhaltlichen Resultate und deren Bedeutung eingegangen werden.

### **4.2.1 Was bedeuten 3600 Proteinspots aus *Arabidopsis thaliana*-Blattextrakten? Und wo sind die anderen?**

Das erste Ergebnis anhand 2-DE-Gelen von *Arabidopsis thaliana*-Standardgeweben war, dass Blattgewebe bei der fraktionierten Proteinextraktion ca. 3600 unterschiedliche Proteinspots aufweist, während Stängelgewebe ca. 4600 unterschiedliche Proteinspots zeigt (vgl. 3.4.1). Vergleicht man diese Resultate mit der Zahl von 25498 Genen, die anhand der Genvorhersageprogramme nach der Sequenzierung des *Arabidopsis thaliana*-Genoms<sup>292</sup> vorhergesagt wurden, so ist Folgendes festzustellen: Selbst wenn jeder Proteinspot ein einmaliges Protein wäre, so würde das Blattgewebe nur 14 % der vorhergesagten *Arabidopsis thaliana*-Gene aufweisen. Es stellen sich also folgende Fragen: Wie viele unterschiedliche Proteine sind in einem aus 4 Zelltypen aufgebauten Gewebe exprimiert? Und wie viele dieser unterschiedlichen Proteine lassen sich auf einem 2-DE-Gel detektieren?

Für humane Zellen hat Julio Celis eine Schätzung abgegeben, die davon ausgeht, dass eine Zelle nicht mehr als 6000 unterschiedliche Gene translatiert<sup>239</sup>. Weiterhin schätzt er, dass nur etwa ein Viertel der Proteinspots auf einem 2-DE-Gel unterschiedliche Proteine sind, während es sich beim Rest der Spots lediglich um Modifikationen der unterschiedlichen Proteine handelt<sup>59</sup>. Die Richtigkeit dieser Zahlen vorausgesetzt, ließe sich umgekehrt schließen, dass 6000 translatierte Proteine pro Zelltyp – bei einer durchschnittlichen Modifikationsrate von 4 pro Protein – in etwa 24000 Proteinspots auf einem 2-DE-Gel eines einzigen Zelltyps ausmachen sollten. Wird ein Gewebe wie Blatt verwendet, sollte sich die Anzahl der Proteinspots noch einmal erhöhen, da ja nicht nur ein Zelltyp vorhanden ist. Bei einer vorsichtigen Weiterschätzung kommt man also leicht auf eine Zahl von 40000 Proteinspezies und mehr.

Doch warum lassen sich lediglich 3600 Proteinspots aus dem Blattgewebe erkennen? Ein überaus wichtiger Aspekt ist die Diskrepanz zwischen dem dynamischen Bereich der Proteinexpression einer Zelle und dem dynamischen Darstellungsvermögen von Proteinen auf 2-DE-Gelen<sup>293</sup>. So macht etwa ein 50 kDa Protein, das mit 10000 Kopien in einer Zelle vorliegt, bei einer vollständigen Proteinextraktion aus einer Million Zellen eine Menge von

0,83 ng aus. Mit anderen Worten: Dieses 50 kDa Protein ist mit etwas Glück gerade auf einem Silber gefärbten 2-DE-Gel zu sehen. Kopienzahlen von mindestens 10000 Proteinen pro Zelle bedeuten, dass weniger stark exprimierte Proteine unter normalen Extraktionsbedingungen nicht sichtbar sind. Ausgehend von der Tatsache, dass die meisten Proteine in 4 verschiedenen modifizierten Formen vorliegen, erhöht sich die Zahl der notwendigen Proteinkopien, da jetzt nicht mehr nur ein Proteinspot mit 0,83 ng getrennt wird, sondern derer 4 mit Bruchteilen eben dieser 0,83 ng.

Dieser Sachverhalt des dynamischen Bereichs der Proteinexpression lässt sich durch eine weitere Schätzung mit der relativ geringen Proteinspotzahl auf einem 2-DE-Gel in Einklang bringen. So wird postuliert, dass etwa 70–80 % der exprimierten Proteine den so genannten „housekeeping“-Proteinen angehören<sup>294,295</sup>. Bei diesen „housekeeping“-Proteinen, die in hohen Kopienzahlen vorliegen, handelt es sich in der Regel um Enzyme, die für die Aufrechterhaltung des zellularen Stoffwechsels notwendig sind. 2 Fragen, die sich hiernach stellen, lauten: Wie viele unterschiedliche Proteinspezies machen diese hochexprimierten „housekeeping“-Proteine aus? Und wie viele verschiedene Proteinspezies kann man in der Klasse der niedrigexprimierten Proteine vermuten? Lefkovitz et al. beantwortet die erste Frage mit einer Hypothese, dass 90 % der exprimierten Proteine einer Zelle aus 70–200 unterschiedlichen Proteinspezies bestehen<sup>296</sup>, während bei Goffea et al. zu lesen ist, dass ca. 50 % der etwa 6000 Gene von der Hefe *Saccharomyces cerevisiae* niedrig exprimierte Proteine darstellen<sup>4</sup>. Fasst man all diese Schätzungen und Erkenntnisse zusammen, so wundert es also nicht weiter, dass ein *Arabidopsis thaliana*-Gesamtblattextrakt nicht mehr als 3600 Proteinspots aufweist.

Um diese theoretischen Überlegungen experimentell zu überprüfen, wurden wie in 3.8.1 beschrieben, 768 Proteinspots aus Blatt Fraktion 1 und 672 Proteinspots aus Blatt Fraktion 2 ausgeschnitten sowie massenspektrometrisch analysiert. Die Ergebnisse dieser Messung zeigen 2 einfache numerische Trends, die von den zuvor formulierten theoretischen Erwägungen abweichen. So konnten von den 768 ausgeschnittenen Proteinspots aus Blattfraktion 1 406 Stück identifiziert werden. Von diesen 406 Proteinspots zeigten 177 unterschiedliche Datenbankeinträge. Dies bedeutet, dass sich etwa 2,3 Proteinspots pro Datenbankeintrag bei den Fraktion 1-Proteinspots finden. Die Zahlen für Fraktion 2 sehen mit etwa 1,7 Proteinspots pro Datenbankeintrag ähnlich aus und weichen relativ deutlich von der zuvor postulierten Zahl von 4 Proteinspots pro Protein ab.

Bei einer globaleren – also blattgewebs- und nicht mehr fraktionsbezogenen – Betrachtung der Identifikationsergebnisse ist festzustellen, dass von den 3600 detektierbaren Proteinen aus beiden Gelen etwa 674 Proteinspots identifiziert wurden, die 266 unterschiedliche

Datenbankeinträge aufweisen. Bei diesen 266 Proteine handelt es sich in der Tat hauptsächlich um cytoplasmatische oder aber um aus den Chloroplasten stammende „housekeeping“-Proteine. So konnten unter ihnen etwa alle Enzyme gefunden werden, die am Calvin-Zyklus<sup>297</sup> beteiligt sind. Weitere identifizierte Proteine sind Enzyme, die an Prozessen der Photosynthese beteiligt sind, wie die Superoxyddismutase oder aber die Ascorbat-Peroxydase. Darüber hinaus finden sich aus allen Bereichen des pflanzlichen Stoffwechsels Enzyme, die für die Aminosäuresynthese, Assimilation von Sulfat oder aber für die Degradation von Proteinen notwendig sind.

Hingegen finden sich keine Proteine, die an der Signalperzeption und deren Weiterleitung beteiligt sind. Dazu gehören zum einen Rezeptormoleküle, die in der Regel Transmembrandomänen aufweisen<sup>298</sup> und zum anderen Guanin-Nukleotid-bindende Proteine (G-Proteine)<sup>299</sup> ebenso wie Kinasen der Serin-, Threonin- und Tyrosinkinasefamilie<sup>300</sup>. Des weiteren konnten keine Proteine der Klasse der Transkriptionsfaktoren gefunden werden<sup>301,302</sup>. Das Fehlen von Kinasen und Transkriptionsfaktoren wurde letztlich schon weiter oben erklärt: Diese Proteine treten in der Zelle in sehr geringen Kopienzahlen auf, sodass sie zwar in den Proteinextrakten vorhanden, aber mittels einer Silber-Färbung nicht nachweisbar sind.

Die in 2-DE-Blattextraktgelen ebenfalls nicht detektierbare Klasse der Transmembranproteine, die Datenbankrecherchen zufolge etwa 20–30 % der offenen Leseraster (ORFs) von Genomsequenzen ausmachen<sup>226,227,303</sup>, stellt ein grundsätzliches Problem für die 2-DE-Proteintrennung dar<sup>123</sup>. Zum einen sind viele von diesen stark hydrophoben Proteinen selten exprimiert, zum anderen präzipitieren sie häufig entweder direkt bei der Applikation der Proteine auf das IEF-Gel<sup>233</sup> oder aber bei der Annäherung an ihren isoelektrischen Punkt<sup>188,249</sup>. Dementsprechend stellte sich im Verlauf der vorliegenden Arbeit die Frage, ob diese Proteine in unseren Proteinextrakten nicht existieren oder ob sie im Verlauf der Trennung verloren gehen. Dass sich die Transmembranproteine in den Blattextrakten finden, konnte letztendlich anhand der Analyse von einigen Proteinbanden aus Fraktion 3-SDS-PAGE-Gelen gezeigt werden. Hierbei wurden die Sedimente aus Fraktion 2 mittels eines SDS-Probenpuffers aufgearbeitet und in SDS-PAGE-Gelen aufgetrennt (vgl. 3.8.1.2). 3 der 10 ausgeschnittenen Proteinbanden dieser Trennung wiesen Proteine auf, die voraussichtlich 3 bzw. 6 Transmembrandomänen zeigten.

Hieraus resultiert, dass es nach wie vor Bereiche gibt, die mittels 2-DE-Gelanalyse von Gesamtproteinextrakten bestimmter Gewebe nicht erfasst werden. Zu diesen Bereichen gehören zum einen niedrig exprimierte Proteine, mit Kopienzahlen von weniger als etwa 1000–10000 Kopien pro Zelle. Für diese Proteinart ist die Sensitivität der gängigen Farbstoffe einfach nicht ausreichend. Eine Strategie, die zur Darstellung solcher Proteine auf 2-DE-

Gelen beiträgt, ist die Sub- bzw. Vorfraktionierung von Geweben und Proteinextrakten, die zur Anreicherung von spezifischen Proteinspezies führt. Ein Beispiel einer solchen Anreicherung wurde im Rahmen dieser Arbeit in 3.8.3 gegeben und wird in 4.2.3 weitergehend diskutiert.

Darüber hinaus bilden stark hydrophobe Proteine ebenfalls eine Gruppe, die nicht ohne weiteres in 2-DE-Gelen darstellbar ist. Doch auch hierbei lässt sich Licht am Ende des Tunnels erblicken: Es werden stetig Fortschritte im Rahmen der Proteinpräparation und Proteinsolubilisierung erzielt und es erscheinen Publikationen, die zeigen, dass bestimmte solcher Transmembranproteine doch in 2-DE-Gelen darstellbar sind<sup>153,304</sup>.

Trotz der aufgezeigten Schwierigkeiten beanspruchen Gesamtproteinextrakte und die hieraus resultierenden 2-DE-Gele einen überaus wichtigen Stellenwert in der Proteomanalyse. Zum einen erhält man anhand dieser verhältnismäßig leicht generierbaren Proteinextrakte Übersichtsmuster, die gerade bei differentiellen, nicht-hypothese-getriebenen 2-DE-Analysen, wie sie in 4.2.2 eingehend besprochen werden, erste wichtige Anhaltspunkte über zugrunde liegende Differenzen. Eben diese ersten Anhaltspunkte können nach entsprechender Analyse zu gezielten Auswertungen des zu untersuchenden Systems führen.

Zum anderen eröffnen gerade Gesamtextraktanalysen – aufgrund von Gemeinsamkeiten auf der Effektorebene – Einsichten über Mechanismen, die unterschiedlichen und in ihrem Wesen nichtverwandt erscheinenden Sachverhalten zugrunde liegen. Mit anderen Worten: Das biologische System sollte nicht nur über seine regulierenden Proteine, in den spezifischen Subkompartimenten, analysiert werden, sondern auch anhand der Muster gemeinsamer Auswirkungen unterschiedlicher biologischer Prozesse.

#### **4.2.2 Differentielle 2-DE-Analyse**

Die differentielle Proteomanalyse, die sich auch funktionelle Proteomanalyse nennen ließe, bietet eine Erfolg versprechende Herangehensweise an proteomische Fragestellungen jeglicher Art. Es lässt sich sowohl eine klassische Perturbationsanalyse durchführen, bei der ein Normalzustand gegen einen stimulierten, veränderten oder sonst wie beeinflussten Zustand verglichen wird. Oder aber es wird eine Entwicklungsstudie durchgeführt, um verschiedene Entwicklungsstadien ein- und desselben biologischen Systems zu vergleichen.

Das Angenehme an differentiellen Proteomanalysen ist, dass sie sich aus unterschiedlichsten Ausgangssituationen durchführen lassen. Liegt keinerlei Anhaltspunkt bezüglich der erwarteten Unterschiede zwischen zu vergleichenden Zuständen vor, so ist es möglich, einen nicht-hypothesengetriebenen Ansatz zu verfolgen. Bei einer solchen Vorgehensweise werden



zumeist Gesamtextrakte von zu vergleichenden Zuständen hergestellt und auf differentielle Proteine untersucht<sup>101,106</sup>. Normalerweise werden die erzielten Ergebnisse sodann anhand bereits bekannter Daten in einen physiologischen Kontext eingeordnet. Hierfür können die großen, mit mRNA-Expressionsdaten gespickten Datenbanken eine wichtige Hilfe bieten. Allerdings gilt es zu bedenken, dass die mRNA-Expression nicht zwingend mit der Proteinexpression korreliert<sup>32-34,305</sup>.

Ein anderer, eher hypothesengetriebener Ansatz basiert auf der differentiellen Analyse zweier Zustände unter selektiver Berücksichtigung von spezifischen Proteinklassen. Diese lassen sich entweder selektiv mittels Immunopräzipitation aus den jeweiligen Proteinextrakten isolieren oder aber selektiv unter Verwendung spezifischer Antikörper auf 2-DE-Gelen darstellen<sup>306,307</sup>. Die Selektion kann sich dabei sowohl auf Proteine mit bestimmten Domänen<sup>308</sup> oder aber auf Proteine mit bestimmten Modifikationen wie Glycosilierung<sup>309</sup> oder aber Ubiquitinierungen<sup>310</sup> erstrecken. Der Vorteil des hypothesengetriebenen Ansatzes besteht eindeutig darin, dass die selektiv vorfraktionierten, differentiellen Proteinmuster eine wesentlich geringere Komplexität aufweisen, als dies bei nicht-hypothesengetriebenen Gesamtextraktionen der Fall ist. Eben diese verminderte Komplexität kann die Auswertung der differentiellen 2-DE-Gele enorm erleichtern.

Der im Rahmen dieser Arbeit gewählte differentielle Perturbationsansatz lässt sich als Mischansatz bezeichnen, da es sich hierbei um einen nicht-hypothesengetriebenen Ansatz bezüglich der erwarteten Identität der differentiell regulierten Proteine handelt. Gleichzeitig handelte es sich jedoch um einen einschränkenden Ansatz bezüglich des Subproteoms, das durch die differentielle Stimulation der zu vergleichenden Pflanzen untersucht wird. So wurden in 3.8.2 die differentiell regulierten Proteine des Phloemexsudats, der Trockenstress ausgesetzten Gurkenpflanze *Cucurbita sativa* mit dem einer Kontrollpflanze verglichen. Die Eingangsanalyse erfolgte auf SDS-PAGE-Ebene und ließ keine differentiell exprimierten Proteine erkennen, sodass der zweite Ansatz einen Vergleich der Proben auf 2-DE-Ebene beinhaltete. Aufgrund der Wahl eines Subproteoms der Gurkenpflanze, waren die beiden resultierenden 2-DE-Gele verhältnismäßig einfach – selbst ohne Verwendung von digitaler Bildauswertesoftware – zu vergleichen. Das Ergebnis dieses Vergleichs zeigte eine Vielzahl von induzierten Proteinspots unter Trockenstress. So werden z. B. Proteaseinhibitoren hochreguliert. Diese könnten verstärkt exprimiert werden, um eine übermäßige Proteaseaktivität im Rahmen steigender Denaturierung aufgrund erhöhter Proteinkonzentration in der langsam austrocknenden Pflanze zu verhindern. Weitere mögliche Erklärungen für die in Tabelle 11 aufgeführten differentiellen Proteine werden zur Zeit noch in Kooperation mit der Arbeitsgruppe von Julia Kehr erarbeitet.

### 4.2.3 Subproteomics: Wie sauber ist ein Komplex? Funktionelle Proteine oder Artefakte?

Das letzte im Rahmen dieser Arbeit erörterte Anwendungsbeispiel (vgl. 3.8.3) resultiert aus den Befunden der subfraktionierten ribosomalen Proteinextrakte aus *Arabidopsis thaliana*-Blättern. Der erste Grund für die Wahl eines Subproteoms für die 2-DE-Trennung basiert auf dem in 4.2.1 diskutierten Problem des dynamischen Bereichs der Proteinexpression in Zellen: Zum einen liegen niedrig exprimierte Proteine bei Gesamtextraktionen in zu geringen Mengen vor, als dass sie sich mittels der gängigen Färbemethoden in 2-DE-Gelen darstellen ließen. Zum anderen würden diese niedrig exprimierten Proteine – sofern man in der Lage wäre, sie aufgrund einer Sensitivitätsverbesserung der Färbung darzustellen – von den hochexprimierten Proteinen überfärbt und diese Gele unauswertbar. Der zweite Grund für die Wahl des 80S ribosomalen Subproteoms als Analyseobjekt beruht auf der geringen Anzahl von 80S ribosomalen Proteinen in den 2 Fraktionen der Blattextrakte. So fanden sich von den 73–80 erwarteten unterschiedlichen ribosomalen Proteinen<sup>232</sup> lediglich 7. Das Fehlen dieser Proteine kann unterschiedliche Ursachen haben. So sind die meisten ribosomalen Proteine stark basisch, wodurch sie bei einer Gleichgewichtsfokussierung aus dem Gel herauslaufen können. Eine solche Erklärung erscheint naheliegend, jedoch wenig wahrscheinlich, da die verwendeten IEF-Gele ein so genanntes CAP-Gel am basischen Ende aufweisen<sup>311</sup>. Dieses CAP-Gel weist eine erhöhte Acrylamidkonzentration auf, sodass auch stark basische Proteine aufgrund der verkleinerten Porengröße des CAP-Gels zumindest teilweise zurückgehalten werden. Diese Proteine sind dann zwar nicht mehr sauber aufgetrennt, lassen sich aber als Streifen aufgelaufener Proteine im 2-DE-Gel erkennen. Ein zweiter und bei weitem plausiblerer Grund für das Fehlen ribosomaler Proteine beruht auf der Proteinextraktion an sich. Die im Rahmen dieser Arbeit etablierte Methode verwendet eine ultrazentrifugationsbasierte Trennung von löslichen und ungelösten Komponenten. Die Zentrifugalbeschleunigung beträgt dabei 226000 g – ein Wert, bei dem sowohl das intakte 80S Ribosom als auch seine beiden Untereinheiten von 40S und 60S sedimentieren und somit in keinem der Extrakte auftauchen dürften. Um diese Diskrepanz zu überbrücken und ein Beispiel für eine Subproteomanalyse zu liefern, wurde eine gradientenzentrifugationsbasierte Methode mit dem Ziel etabliert, mehrere Organellen bzw. Komplexe aus unterschiedlichen Blattgeweben zu isolieren. In dem hier durchgeführten Ansatz wurde zunächst einmal die Präparation von 80S Ribosomen optimiert. Denkbar ist aber auch die Präparation von

Chloroplasten, Mitochondrien oder Zellkernen in einem fraktionierten Trenngang mittels Gradientendichtezentrifugation durchzuführen.

Die 80S-ribosomale Präparation erfolgte in 2 verschiedenen Aufreinigungsstufen, die jeweils mittels 2-DE-Gelen aufgetrennt wurden (vgl. 3.8.3). Als Ergebnis der massenspektrometrischen Analyse wiesen beide Extrakte eine hohe Anreicherung von sowohl 60S- als auch 40S-ribosomalen Proteinen auf. So zeigte etwa der weniger aufgereinigte Extrakt 1 (E1) insgesamt 38 unterschiedliche ribosomale Proteine, während der intensiver aufgereinigte Extrakt 2 (E2) sogar 54 der erwarteten 70–80 unterschiedlichen ribosomalen Proteine verzeichnete. Ein Vergleich der nichtribosomalen Proteine, die auf den 2-DE-Gelen beider Extrakte identifizierbar waren, zeigte ein gegenläufiges Bild: In E1 konnten 42 unterschiedliche nichtribosomale Proteine identifiziert werden, in E2 hingegen nur 19. Beim weitergehenden Identitätsvergleich dieser nichtribosomalen Proteine kommt man zu folgendem aufschlussreichem Ergebnis: Insgesamt stimmen lediglich 6 nichtribosomale Proteine überein. Dementsprechend ließen sich neben 24 zusätzlichen ribosomalen 80S Proteinen noch 13 zusätzliche nichtribosomale Proteine anreichern, die in dem weniger intensiv aufgereinigten Extrakt E1 nicht gefunden wurden. Dies unterstreicht eindeutig, dass die Aufreinigung des Ribosoms in E2 sehr effektiv und umfassend war.

Nichtsdestotrotz drängen sich nunmehr folgende Fragen auf: Inwieweit stellen diese nichtribosomalen Proteine, die in den einzelnen Extrakten der Ribosomen angereichert werden, reine Kontaminationen dar bzw. können diese Proteine nicht auch mit dem Ribosom funktionell assoziierte Faktoren darstellen? Es findet sich beispielsweise eine große Anzahl von Chaperonen zwischen den nichtribosomalen Proteinen von E1. Unter ihnen sind einige – wie HSP 70 (Heat Shock Protein 70) – dafür bekannt, dass sie sich an die wachsende Polypeptidkette der translatierten mRNA heften, um eine korrekte Faltung der Proteine zu gewährleisten<sup>312</sup>. Darüber hinaus findet man HSP 60, ein weiteres Chaperon, das eine ähnliche Funktion wie HSP 70 ausüben kann<sup>312</sup>. Ein weiterer interessanter Befund besteht in der Anreicherung von Proteasomproteinen in E1. Das Proteasom ist ein Proteinkomplex, der für die proteolytische Beseitigung von Proteinen und Polypeptiden verantwortlich ist<sup>313</sup>. Um zu überprüfen, inwieweit die zusätzlichen nichtribosomalen Proteine allgemeine Kontaminationen darstellen oder ob sie spezifisch für die Ribosomenpräparation sind, wurden sie mit im Rahmen der fraktionierten Proteinextraktion aus Blattgewebe (vgl. 3.8.1) identifizierten Proteinen verglichen. Hierbei stellte sich heraus, dass sowohl die nichtribosomalen Proteine aus E1 und E2 mit 54 % bzw. 42 % Übereinstimmung eine große Anzahl zuvor nicht in den Blattfraktionen 1 und 2 identifizierter Proteine aufweisen. Dies ist erneut ein Indiz für die Spezifität dieser Proteine.

Um sich Klarheit über eine mögliche Assoziation zwischen diesen Proteinen und dem Ribosom zu verschaffen, könnte man abschließend eine *In vivo*-Lokalisation von fluoreszent markierten Fusionsproteinen in Verbindung mit Fluoreszenzmikroskopie durchführen<sup>314</sup>. Anhand solch einer funktionellen Analyse könnten somit Proteine, die tatsächlich mit dem Ribosom assoziiert sind, von solchen unterschieden werden, die als Artefakte nur zufällig mit den Ribosomen cosedimentieren.

### **4.3 Zukunft der Proteomanalyse im Allgemeinen und Speziellen**

Nachdem bislang sowohl technische als auch inhaltliche Aspekte der Proteomanalyse in engem Bezug zu den im Rahmen dieser Arbeit untersuchten Techniken und Fragestellungen diskutiert wurden, soll abschließend ein Ausblick auf Erfolg versprechende neue Techniken und Entwicklungen der Proteomforschung gewagt werden.

Die klassische Proteomanalyse, die auf Basis von 2-DE-Gelen und MALDI-MS durchgeführt wird, hat im Laufe der letzten Jahre illustre Erweiterungen erfahren. So sind einige der Bestrebungen, eine gelfreie Proteintrennung durchzuführen, von partiellem Erfolg gekrönt worden: Wie bereits ausgeführt (vgl. 4.1.2.2), wurde in der Gruppe um John Yates eine multidimensionale chromatographiebasierte Proteintrennung entwickelt, anhand derer sich in Verbindung mit einem ESI-Ionenfallen-Massenspektrometer mehrere tausend unterschiedliche Proteine aus verschiedenen Organismen identifizieren ließen<sup>263-265</sup>. Die 2 zentralen Probleme dieses Ansatzes beruhen zum einen auf dem Verlust jeglicher quantitativer Informationen der Expressionszustände einzelner Proteine<sup>315</sup>. Zum anderen geht in weiten Teilen die Möglichkeit verloren, translationale oder posttranslationale Modifikationen<sup>316</sup> zu erkennen, geschweige denn zu identifizieren.

Dies ist umso schwerwiegender, als eben diese beiden Aspekte im Laufe der letzten Jahre für die Proteomforschung immer größere Bedeutung gewonnen haben. So ist zum Beispiel bekannt, dass bei einer Vielzahl von Krankheiten der Expressionsgrad bzw. die Modifikationen einzelner Proteine überaus wichtig sind<sup>317</sup>. Die zunehmende Bedeutung dieser Aspekte lässt sich darüber hinaus an der steigenden Publikationszahl ablesen.

Gerade in Bezug auf die akkurate Quantifikation von Proteinexpression wurden einige interessante Ansätze verfolgt – unter anderem die auf Basis der 2-DE-Proteintrennung funktionierende DIGE (*Differential Gel Electrophoresis*). Die diesem Ansatz zugrunde liegende Strategie beruht auf der präelektrophoretischen Färbung von 2 zu vergleichenden Proteinmischungen anhand unterschiedlicher Fluoreszenzfarbstoffen<sup>318</sup>. Diese beiden Farbstoffe unterschieden sich hinsichtlich Molekulargewicht und Ladung nicht, wohl aber aufgrund ihrer fluorochromen Eigenschaften. Nachdem die beiden Proteinmischungen jeweils

separat mit einem der beiden Farbstoffe markiert wurden, werden sie im Verhältnis eins zu eins gemischt und gemeinsam über 2-DE-Gele aufgetrennt. Die hieraus resultierenden Gele lassen sich als Nächstes in einem Laserscanner mit unterschiedlichen Anregungswellenlängen auswerten. Zu den großen Vorteilen dieser Methode zählt, dass systematische Unterschiede, die von Lauf zu Lauf auftreten können und somit den Vergleich zweier Proben erschweren, durch diesen Ansatz ausgeschlossen werden. Ein weiterer Vorteil von DIGE ist, dass der verwendete Fluoreszenzfarbstoff eine Linearität der Quantifizierbarkeit über 5 Größenordnungen aufweist<sup>319</sup>.

Andere massenspektrometrie-basierte Methoden, die zur differentiellen Quantifizierung von Proteinexpressionszuständen genutzt werden können, beruhen auf der Verwendung stabiler Isotopen<sup>320</sup>. Bei einem Ansatz, der in den letzten Jahren für Aufsehen gesorgt hat, handelt es sich um die so genannte ICAT-Methode (*Isotop-coded affinity tags*)<sup>321</sup>. Sie beruht auf dem differentiellen Markieren cysteinhaltiger Peptide. Es werden erneut proteolytisch verdaute Proteine zweier Zustände mit der ICAT-Reagenz inkubiert, wobei eine Peptidmischung mit der normalen ICAT-Reagenz, die andere Proteinmischung hingegen mit einer modifizierten, um 8 Da schwereren ICAT-Reagenz markiert wird. Nach Vereinigung beider Peptidgemische werden anhand einer Affinitätschromatographie nur die cysteinhaltigen Peptide isoliert und im Massenspektrometer gemessen. Die Massenspektren solcher Mischungen zeigen jeweils um 8 Da pro Cystein auseinander liegende Signale, deren Verhältnis berechnet werden kann. Zusätzlich lassen sich diese Peptide fragmentieren, sodass die zugehörige Proteinidentität ebenfalls ermittelt wird. Inwieweit sich diese beiden Methoden durchsetzen werden, hängt davon ab, wie hoch die Kosten zu veranschlagen sind und natürlich welche Ergebnisse sie in nächster Zeit zu produzieren vermögen.

Auch die Technologie der Protein Chips gehört zu den Bereichen, die verstärkt mit dem Begriff Proteomics in Verbindung gebracht werden<sup>322,323</sup>. In diesem Falle handelt es sich einerseits um Aneinanderreihungen spezieller Antikörper, die zum Erkennen bestimmter Proteine aus komplexen Mischungen verwendet werden<sup>324</sup>. Andererseits handelt es sich um aneinander gereihte Reaktionsgefäße, in denen rekombinant exprimierte Proteine und mögliche Substrate gemischt werden, um deren Spezifität oder aber Funktion zu überprüfen<sup>325</sup>. Beide Ansätze sind noch weit davon entfernt, perfekt zu funktionieren, konnten aber während der letzten Jahre beachtliche Fortschritte verzeichnen.

Nicht zuletzt gewinnt die Analyse von Proteinkomplexen – und damit verbunden die Analyse von Protein-Protein-Wechselwirkungen – immer größere Bedeutung im Rahmen der Proteomanalyse. So wurden z. B. 2 Hochdurchsatzprojekte mit der Zielsetzung durchgeführt, alle Proteinkomplexe aus Hefe zu isolieren, um anschließend ihre Bestandteile zu

identifizieren <sup>141,142</sup>. Ein anderer Ansatz zur Messung von Protein-Protein-Interaktionen beruht dagegen auf der Verwendung des automatisierten Two-Hybrid-Systems in Hefe <sup>326,327</sup>. Hierbei werden keine komplexen, dafür aber direkte Wechselwirkungen zwischen Proteinen gemessen.

Nun mag die Auflistung und Beschreibung all dieser neuen Techniken den Eindruck erwecken, als stießen Ansätze wie 2-DE in Verbindung mit MALDI-MS bereits an ihre Grenzen. Doch kann davon nicht die Rede sein. Denn es existieren noch unzählige Verbesserungsmöglichkeiten im Bereich der Proteinextraktion, der Subfraktionierung von Organellen und Subproteomen, der Extraktion und Trennung von hydrophoben Proteinen auf 2-DE-Gelen, der Färbung und massenspektrometrischen Analyse von 2-DE getrennten Proteinen, der Trennung von nativen Proteinen zur eventuellen Verwendung von funktionsbestimmenden Assays, der Trennung von nativen Proteinkomplexen, die etwas über die Zusammensetzung zellulärer Maschinen verraten könnten, und so weiter und so fort.

Abschließend sei folgender Gedankengang erlaubt: Zum Zeitpunkt, als das Humanengenomprojekt gestartet wurde, war die grundlegende Technik der DNA-Abbruchsequenziermethode nach Sanger <sup>328</sup> bereits 5 Jahre etabliert, sodass sich die Methodenentwicklung auf die Steigerung der Durchsatzgeschwindigkeit konzentrieren konnte. In der Proteomforschung hingegen ist dieser Punkt noch lange nicht erreicht. Hier steckt man noch in den Anfängen. Diskutiert wird über Definitionen und strategische Vorgehensweisen, sodass man auf die nächsten Jahre gespannt sein darf.