

# **Struktur und Funktion der Carboxy-terminalen Domäne des Transkriptionsfaktors CbfA aus *Dictyostelium discoideum***

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)



vorgelegt dem Rat der Biologisch-Pharmazeutischen Fakultät  
der Friedrich-Schiller-Universität Jena

**von Jörg Lucas**

Jena, im November 2009

Dekan: Prof. Dr. Frank Hellwig (Friedrich-Schiller-Universität Jena)

1. Gutachter: Prof. Dr. Thomas Winckler (Friedrich-Schiller-Universität Jena)

2. Gutachter: Prof. Dr. Thorsten Heinzel (Friedrich-Schiller-Universität Jena)

3. Gutachter: Prof. Dr. Markus Maniak (Universität Kassel)

Tag der öffentlichen Verteidigung: 05.03.2010

## Inhalt

Inhaltsverzeichnis .....	i
Abbildungsverzeichnis .....	v
Abkürzungen.....	vii
1 Einleitung .....	1
1.1 Die soziale Amöbe <i>Dictyostelium discoideum</i> .....	1
1.2 Multizelluläre Entwicklung.....	2
1.3 Initiation und Regulation der frühen multizellulären Entwicklung.....	5
1.4 Phyletische Einordnung der Art <i>Dictyostelium discoideum</i> .....	7
1.5 Bedeutung von <i>Dictyostelium discoideum</i> in der Forschung.....	9
1.6 Genomische Organisation von <i>Dictyostelium discoideum</i> .....	10
1.7 Transposable Elemente in <i>Dictyostelium discoideum</i> .....	11
1.8 Aufbau des C-Modul-bindenden Faktors A.....	15
1.9 Funktion des C-Modul-bindenden Faktors A .....	20
1.10 Zielstellung .....	22
2 Material und Methoden.....	23
2.1 Laborgeräte .....	23
2.2 Chemikalien.....	24
2.3 Lösungen und Puffer .....	25
2.4 Verwendete Medien und Antibiotika .....	26
2.5 Benutzte Antikörper.....	27
2.6 Verwendete Mikroorganismen .....	27
2.7 Verwendete Kits und Enzymsysteme.....	28

## Inhalt

---

2.8	Benutzte Software .....	29
2.9	Allgemeine molekularbiologische Methoden.....	30
2.10	Kultivierung von <i>Dictyostelium discoideum</i> und anderen <i>Dictyosteliales</i> .....	30
2.11	Isolation und Synthese von gDNA, RNA, mRNA und cDNA .....	30
2.12	Erstellung von Kernextrakten .....	31
2.13	Quantitative <i>Real Time</i> -PCR (qRT-PCR).....	31
2.14	<i>Microarray</i> -Methoden .....	32
2.15	Phylogenetische Analysen .....	33
2.16	Aufreinigung rekombinanter CbfA-CTD im analytischen Maßstab .....	34
2.17	Aufreinigung rekombinanter CbfA-CTD im präparativen Maßstab.....	35
2.18	DNA-Protein-Interaktionsassay – CTD- <i>Pulldown</i> .....	37
2.19	Kultivierung und Transformation von <i>S. cerevisiae</i> .....	37
2.20	Isolierung von Plasmid-DNA aus Hefe – Plasmidrescue .....	38
2.21	Konstruktion und Screening von Hefe-2-Hybrid-Banken .....	39
3	Ergebnisse.....	40
3.1	Kurze Übersicht zum Stand der Forschung.....	40
3.2	Transkriptionelle Reaktion auf die <i>cbfA</i> -Suppression und die Überexpression von GFP-CbfA <sup>724-998</sup> .....	42
3.3	Clusteranalyse der erhobenen <i>Microarray</i> -Daten.....	48
3.4	Verifizierung der <i>Microarray</i> -Daten .....	50
3.5	Differentiell exprimierte Gene im Vergleich AX2 zu JH.D bilden Anhäufungen in funktionellen Kategorien – GOAT-Analyse .....	53
3.6	Neue Transkriptomansätze.....	55
3.7	GOAT-Analyse der CbfA-abhängigen Gene unter Betrachtung der Regulation durch GFP-CbfA <sup>724-998</sup> .....	59

## Inhalt

---

3.8	Einfluss der Carboxy-terminalen Domäne auf die Entwicklung.....	61
3.9	Kernlokalisierung von GFP-CbfA <sup>724-998</sup> und AT-Haken-Mutanten .....	63
3.10	Genregulation durch GFP-CbfA <sup>724-998</sup> und AT-Haken-Mutanten.....	65
3.11	Regulatorischer Einfluss von Lysinen im Bereich des AT-Hakens .....	66
3.12	Bindung der AT-Hakenmutanten an den Modellpromotor von <i>acaA</i> .....	69
3.13	Konstruktion von Hefe-2-Hybrid-Banken.....	71
3.14	Charakterisierung der Hefe-2-Hybrid-Banken mittels PCR.....	74
3.15	Suche nach Protein-Protein-Interaktionen mittels Hefe-2-Hybrid .....	75
3.16	Betrachtungen zur Phylogenie von CbfA .....	79
3.17	Aufbau der CTD in CbfA-orthologen Proteinen .....	81
3.18	Bakterielle Expression der Carboxy-terminalen Domäne zur kristallographischen Strukturaufklärung .....	82
3.19	Herstellung von rekombinanter CTD im großen Maßstab.....	87
3.20	Nachweis der Funktionalität des rekombinant gereinigten Proteins .....	90
3.21	Weiterführende Arbeiten zur kristallographischen Strukturaufklärung .....	91
4	Diskussion.....	96
4.1	Einsichten in die regulatorische Funktion von CbfA .....	96
4.2	Beteiligung von CbfA <sup>724-998</sup> an der CbfA-abhängigen Genregulation in JH.D.....	98
4.3	Einfluss des Kernmotives des AT-Hakens und seiner umgebenden Region auf die CTD-vermittelte Regulation.....	102
4.4	Suche nach Interaktionspartnern der Carboxy-terminalen Domäne .....	106
4.5	Phylogenie von CbfA .....	108
4.6	Röntgenkristallographische Strukturaufklärung von CbfA <sup>793-1000</sup> .....	109
4.7	Ausblick .....	112

## Inhalt

---

5	Zusammenfassung .....	114
6	Summary .....	116
	Literaturverzeichnis.....	118
	Anhang .....	128
A	Liste der verwendeten Primer .....	128
B	Plasmidliste .....	131
C	Differentiell regulierte Gene zwischen AX2 und JH.D sowie ihre Antwort auf GFP-CbfA <sup>724-998</sup> in JH.D.....	133
D	Liste der Anreicherungen biologischer Prozesse, molekularer Funktionen und zelluläre Kompartimente in den differentiell exprimierten Genen zwischen AX2 und JH.D .....	137
E	Liste der Anreicherungen biologischer Prozesse, molekularer Funktionen und zelluläre Kompartimente in den differentiell exprimierten Genen der Clusteranalyse.....	141
F	Sequenzierung von Einzelklonen der Hefe-2-Hybrid-Banken .....	145
	Liste der wissenschaftlichen Publikationen .....	146
	Danksagung .....	147
	Lebenslauf .....	149
	Ehrenwörtliche Erklärung .....	150

## Abbildungsverzeichnis

Abbildung 1: Darstellung der multizellulären Entwicklung von <i>D. discoideum</i> .....	3
Abbildung 2: Schematische Darstellung des PKA-Regelkreises zur Initiation der Entwicklung in <i>Dictyostelium discoideum</i> .....	5
Abbildung 3: Phylogenie der <i>Dictyosteliidae</i> .....	7
Abbildung 4: Proteombasierende Phylogenie der Eukaryonten.....	8
Abbildung 5: Übersicht über die Transposons in <i>Dictyostelium discoideum</i> .....	12
Abbildung 6: Schematische Darstellung der Struktur von TRE5-A.1.....	14
Abbildung 7: Schematische Darstellung des Domänenstruktur des C-Modul-bindenden Faktors A.....	16
Abbildung 8: Chemischer Mechanismus der Demethylierung von Arginin- und Lysinresten durch JmjC-haltige Proteine .....	17
Abbildung 9: Primärsequenz von der Carboxy-terminalen Domäne von CbfA.....	19
Abbildung 10: <i>Northern-Blot</i> mit Sequenz von ORF1 als Sonde gegen TRE5-A.....	41
Abbildung 11: <i>Western-Blot</i> mit Antikörper 7F3 von Gesamtzelllysaten .....	42
Abbildung 12: Zelluläre Verteilung von CbfA, durch CbfA-Suppression entstandenes CbfA <sup>1-452</sup> und ektopisch exprimiertes CbfA <sup>724-998</sup> .....	44
Abbildung 13: Normalisierung der Primärdaten einer Serie von <i>Microarrays</i> .....	46
Abbildung 14: Clusteranalyse der durch CbfA-Suppression regulierten Gene und ihre Reaktion auf GFP-CbfA <sup>724-998</sup> .....	48
Abbildung 15: Verifizierung der Arraydaten mittels qRT-PCR .....	52
Abbildung 16: GOAT-Analyse CbfA regulierter Gene .....	54
Abbildung 17: <i>De-novo</i> -Synthese von Nukleosid-5'-Monophosphaten im Vergleich der Regulation der beteiligten Gene .....	58
Abbildung 18: GOAT-Analyse der durch CTD Co-regulierten Gene .....	60
Abbildung 19: multizelluläre Entwicklung von AX2, JH.D und deren Transformanten mit GFP-CbfA <sup>724-998</sup> .....	61
Abbildung 20: qRT-PCR 8 Stunden hungernde Zellen gegen nicht hungernde Zellen.....	62

## Abbildungsverzeichnis

---

Abbildung 21: Kernlokalisierung der GFP-CbfA <sup>724-998</sup> in JH.D.....	64
Abbildung 22: Einfluss von Mutationen im Bereich des AT-Hakens auf die Regulation durch CbfA <sup>724-998</sup> .....	65
Abbildung 23: Kernlokalisierung von AT-Haken-Mutanten in JH.D .....	67
Abbildung 24: regulatorischer Einfluss von Lysinen im Bereich des AT-Hakens .....	68
Abbildung 25: Aufreinigung von CbfA-Fragmenten mit Hilfe spezifischer DNA-Proben.....	70
Abbildung 26: Aufreinigung von GFP-CbfA <sup>724-998</sup> und dessen Mutanten mit <i>acaAP</i> <sup>-739 bis -448</sup> .....	71
Abbildung 27: Gelelektrophorese von PCRs an Einzelklonen der Banken.....	73
Abbildung 28: Nachweis bekannter Fragmente in den Hefebanken.....	75
Abbildung 29: exemplarische Beispiele der Überprüfung der im Mating gefundenen Klone.....	77
Abbildung 30: phylogenetisches Verhältnis von CbfA-ähnlichen Proteinen anhand ihrer Aminosäuresequenz.....	80
Abbildung 31: Sequenzvergleich der Carboxy-terminalen Domänen von CbfA und orthologen Proteinen.....	82
Abbildung 32: Vergleich der Induktion von Fusionsproteinen mit CbfA <sup>793-1000</sup> .....	84
Abbildung 33: Induktions- und Aufreinigungsversuche im kleinem Maßstab .....	86
Abbildung 34: Abhängigkeit der Elution von (His) <sub>6</sub> -CbfA <sup>793-1000</sup> von der Imidazolkonzentration. ....	88
Abbildung 35: Bestimmung des optimalen Verhältnisses Faktor Xa zum Fusionsprotein .....	88
Abbildung 36: Aufreinigung von rekombinatem CbfA <sup>793-1000</sup> mittels Affinitätschromatographie .....	89
Abbildung 37: Funktionstest des rekombinanten Proteins durch Bindung an bekannte Promotorfragmente.....	90
Abbildung 38: Überprüfung der Strukturierung und Qualität des rekombinant gereinigten Proteins .....	92
Abbildung 39: Aufklärung der Abbaufragmente der Kristallisationsansätze .....	94



## Abkürzungen

Standard-Einheiten und ihre Abstufungen sind nicht in der Liste der Abkürzungen vorhanden. Nicht der deutschen Sprache entnommene Fachbegriffe, Gennamen und wissenschaftliche Bezeichnungen taxonomischer Gruppen sind in der gesamten Arbeit *kursiv* dargestellt.

3AT	3-Amino-1,2,4-triazol
A, Ade	Adenin, Nukleobase, hier Auxotrophiemarker in Hefe
AS	Aminosäure
bp	Basenpaare
bzw.	beziehungsweise
cAMP	Cyclisches Adenosin -3',5'- monophosphat
CbfA	C-Modul-bindender Faktor A, Protein aus <i>Dictyostelium discoideum</i>
cDNA	<i>complementary deoxyribonucleic acid</i> , Komplementäre Desoxyribonukleinsäure
cfu	<i>colony forming unit</i> , Kolonie bildende Einheit, ein Maß für die Zellzahl von Mikroorganismen
CTD	Carboxy-terminale Domäne, meist bezogen auf CbfA
Da (kDa)	Dalton, Maßeinheit der Masse ,1 Da = $\frac{1}{12}$ der Atommasse von Kohlenstoff
dA	2'-Desoxyadenosin-5'-monophosphat
DNA	<i>deoxyribonucleic acid</i> , Desoxyribonukleinsäure
dT	2'-Desoxythymidin-5'-monophosphat
ESI-MS	Elektrospray-Ionisation-Massenspektroskopie
EST	<i>expressed sequence tag</i> , meist durch cDNA-Synthese gewonnene, exprimierte Sequenzen, ein Gen kann mehrere ESTs besitzen
FDR	<i>false discovery rate</i> , gibt die Wahrscheinlichkeit an, dass ein Gen in einer Liste differentiell exprimierter Gene nicht differentiell exprimiert ist
GOAT	<i>Gene Ontology Annotation Tool</i> , Programmpaket zur Annotation von Genen
H, His	Histidin, Aminosäure, hier Auxotrophiemarker in Hefe
HRP	<i>horseradish peroxidase</i> , Meerrettichperoxidase, Bestandteil der Chemilumineszenzdetektion eines <i>Western-Blots</i>
IPTG	Isopropylthio- $\beta$ -D-galactosid
L, Leu	Leucin, Aminosäure, hier Auxotrophiemarker in Hefe
JmjC	<i>Jumonji</i> -Domäne des Carboxy-terminalen Typs, Proteindomäne
LOWESS	<i>locally weighted least squares</i> , statistische Methode der Normalisierung

## Abkürzungen

---

MBP	maltosebindendes Protein
min	Minute
MQ-Wasser	entsalztes Wasser mit geringer Leitfähigkeit Q
mRNA	<i>messenger-ribonucleic acid</i> , Boten-Ribonukleinsäure
n.b.	nicht bestimmt
Ni-NTA	Nickel-Nitrilotriacetische Säure
NLS	<i>nuclear localization sequence</i> , Kernlokalisationssequenz
n.s.	nicht signifikant
NTD	Amino-terminale Domäne, meist bezogen auf CbfA
ORF	<i>open reading frame</i> , offener Leserahmen
PCR	Polymerasekettenreaktion, Verfahren zur Amplifizierung von DNA
Pst	<i>pre-stalk cell</i> , Prä-Sporenzellen, Zelltyp von sich entwickelnden <i>Dictyostelium discoideum</i>
qRT-PCR	quantitative <i>real time</i> Polymerasekettenreaktion an einer komplementären DNA, Methode zur relativen Quantifizierung von mRNA-Mengen
RNA	<i>ribonucleic acid</i> , Ribonukleinsäure
rpm	<i>rotations per minute</i> , Umdrehungen pro Minute
SAM	<i>Significance Analysis of Microarrays</i> , Programm zur Identifizierung differentiell exprimierter Gene innerhalb eines <i>Microarray</i> -Experimentes
SDS-PAGE	Natriumdodecylsulfat-Polyacrylamidgelelektrophorese, Methode zur Proteinauftrennung
sek	Sekunde
T, Trp	Tryptophan, Aminosäure, hier Auxotrophiemarker in Hefe
tRNA	<i>transfer-ribonucleic acid</i> , Transfer-Ribonukleinsäure
vgl.	vergleiche
v/v	<i>volume per volume</i> , Volumen pro Volumen
w/v	<i>weight per volume</i> , Masse pro Volumen
u.a.	unter anderen
UTR	<i>untranslated region</i> , untranslatierte Region auf einer Nukleinsäure
X- $\alpha$ -Gal	5-Bromo-4-chloro-3-indolyl- $\alpha$ -D-galactopyranosid
X- $\beta$ - Gal	5-Bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranosid
z.B.	zum Beispiel

# 1 Einleitung

Zur Einführung der in dieser Arbeit behandelten Thematik soll ein Überblick über den untersuchten Organismus *Dictyostelium discoideum*, dessen grundlegende Eigenschaften, seine phylogenetische Stellung sowie seine Bedeutung in der Forschung gegeben werden. Es folgt eine Darstellung des Wissensstandes über das in dieser Arbeit zu charakterisierende Protein C-Modul-bindender Faktor A (CbfA, DDB\_G0279409) im Allgemeinen und dessen Carboxy-terminaler Domäne (CTD) im Speziellen.

## 1.1 Die soziale Amöbe *Dictyostelium discoideum*

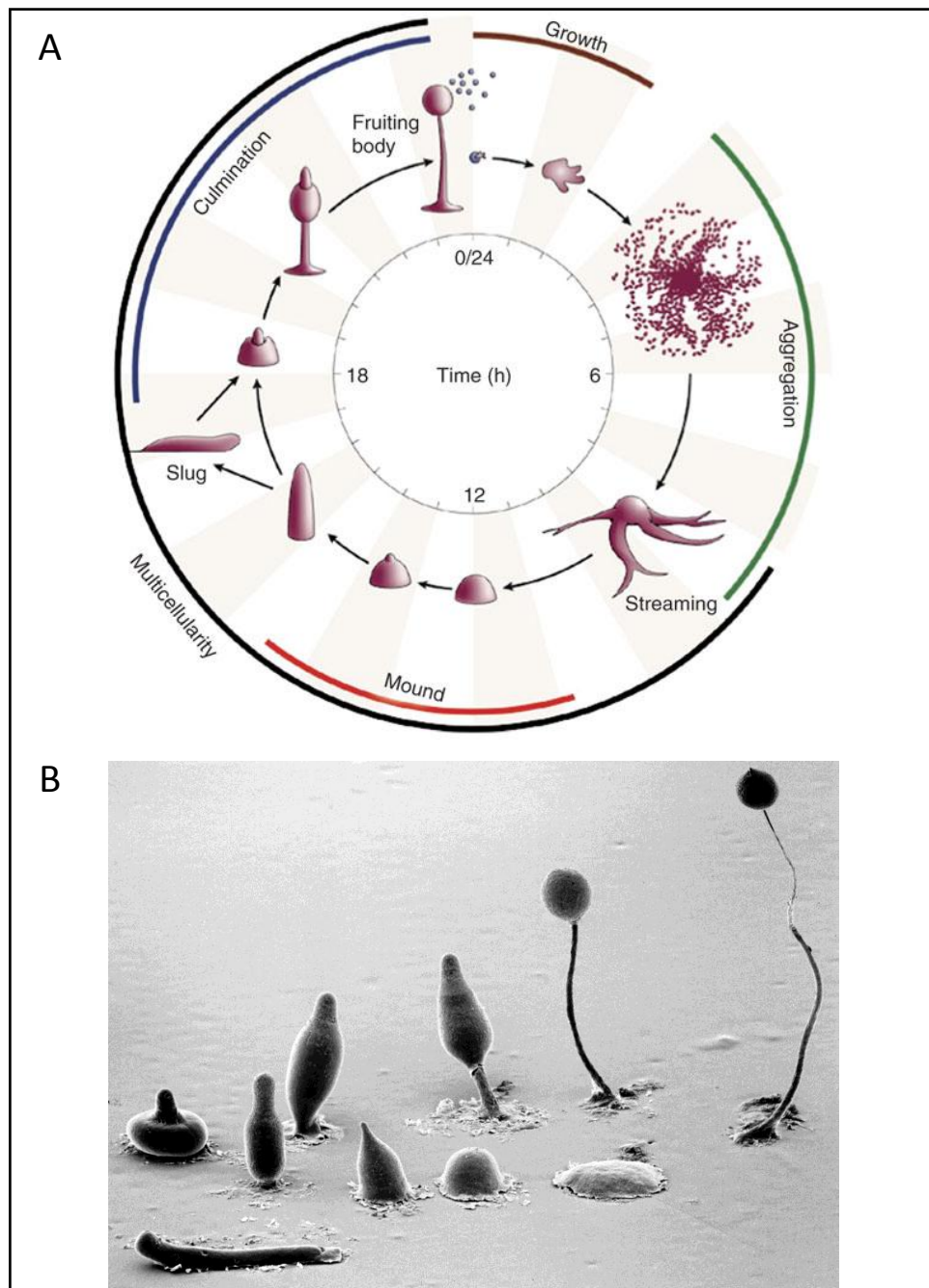
Die Gattung *Dictyostelium* wurde erstmalig Mitte des 19. Jahrhunderts anhand der Art *Dictyostelium mucuroides* beschrieben und aufgrund der Morphologie der Fruchtkörper den Pilzen zugeordnet (Brefeld 1869). Wegen des Fehlens filamentöser Stadien schien diese Zuordnung bereits früh überholt, wobei die Stellung im System des Lebens lange ungeklärt geblieben ist. Die in dieser Arbeit untersuchte Art *Dictyostelium discoideum* wurde in den dreißiger Jahren des letzten Jahrhunderts erstmalig beschrieben (Raper 1935). Es handelte sich hierbei um das Isolat NC4 aus dem Boden des Hardwood Forest, einem Mischwald in Mittelgebirgslage in North Carolina (USA). Dieses Isolat ist der direkte Vorgänger des in dieser Arbeit untersuchten Laborstammes AX2 (Sussman und Sussman 1967, Watts und Ashworth 1970). *D. discoideum* und andere Vertreter der Gattung leben auf Waldböden und ernähren sich dort von Bakterien oder anderen Mikroorganismen (u.a. Bozzaro und Ponte 1995). In Phasen ausreichender Nährstoffversorgung liegt dabei der Organismus als einzellige Amöbe vor, welche sich in 4 Stunden durch Zellteilung verdoppeln kann. Für eine Teilung müssen etwa 1.000 Bakterien phagozytiert, getötet und verdaut werden (Mutzel 1991, Clarke und Maddera 2006). Der auf hohe Makropinozytoseraten selektionierte Stamm AX2 benötigt für eine Zellteilung 10 bis 12 Stunden, ist aber zum Wachstum auf axenischen Medien geeignet (Watts und Ashworth 1970), was viele biochemische und molekulare Studien erst ermöglicht. In der vegetativen Wachstumsphase können sich die 8 bis 12 µm großen Zellen durch Aktin- und Myosindynamik amöboid bewegen. Bei dieser Bewegung werden Pseudopodien ausgebildet.

Bei einsetzendem Nährstoffmangel oder abiotischen Signalen, wie zu hoher Temperatur oder zu geringer Feuchtigkeit, kommt es zur Aggregation der Amöben. Etwa 100.000 Amöben bilden einen Fruchtkörper (Kessin 2001). Im Verlauf der Ausbildung eines Fruchtkörpers kommt es zur Ausprägung multizellulärer Charakteristika. Dazu gehören die Zelldifferenzierung, Zell-Zell-Kontakte, die Bewegung als ein Organismus sowie der programmierte Zelltod. Zellen sterben kontrolliert bei der Ausbildung des Stieles ab, wodurch die sich bildenden Sporen emporgehoben werden (Raper 1984, Kessin 2001, Schaap *et al.* 2006). Die multizelluläre Entwicklung ist ein Prozess, welcher unter Laborbedingungen etwa 24 Stunden dauert.

### 1.2 Multizelluläre Entwicklung

Der Entwicklungszyklus von *Dictyostelium discoideum*, wie in Abbildung 1 dargestellt, beginnt durch das Aussenden des Botenstoffes cyclisches Adenosin-3',5'-monophosphat (cAMP) durch sogenannte Schrittmacherzellen (Loomis 1993). Diese Zellen sind die ersten, welche durch eine mehrere Komponenten enthaltende Sensor- und Signalkaskade nachteilige Bedingungen feststellen. Sezernierte Komponenten dieses *Quorum-sensing*-Systems sind unter anderen der Pre Starvation Factor (PSF) und der Conditioned Medium Factor (CMF) (Clarke und Gomer 1995). Zu den auslösenden Umweltbedingungen gehören eine zu hohe Amöbendichte, geringe Nahrungsressourcen sowie abiotische Faktoren wie zu hohe Temperaturen oder der Ammoniumgehalt im Boden. Etwa 100.000 Zellen bilden ein Aggregat am Punkt des ersten Signales. Dabei sezernieren die einströmenden Zellen ebenfalls cAMP. Da dieses zum Großteil an ihrem dem Aggregationszentrum abgewandten Zellende geschieht, kommt es zum sternförmigen Einströmen der Zellen (Abbildung 1). Diese Phase wird als *Streaming* bezeichnet. Im Unterschied zu den plasmoidalen Schleimpilzen verschmelzen die Zellen nicht zu einem Synzytium, sie bilden ein Pseudoplasmodium. Die Aggregation ist nach 8 bis 12 Stunden beendet. Bereits in diesem Stadium kommt es zu einer Differenzierung der Zellen in zwei Hauptkategorien, welche im späteren Verlauf in mehrere Unterkategorien eingeteilt werden können. Etwa 20% der Zellen sind Prä-Stielzellen, etwa 70% sind Prä-Sporenzellen. Die restlichen 10% sind undifferenziert (Escalante und Vicente 2000). Es kann zu Umwandlungen zwischen den beiden Kategorien von Zellen kommen, wobei das Verhältnis während der gesamten Entwicklung konstant bleibt. Die Differenzierung in Prä-Stielzellen verläuft unter

anderen in Abhängigkeit vom *Differentiation-Inducing Factor* (DIF), einem chlorinierten Alkylphenon (Kay und Jermyn 1983, Insall *et al.* 1992) und cAMP.



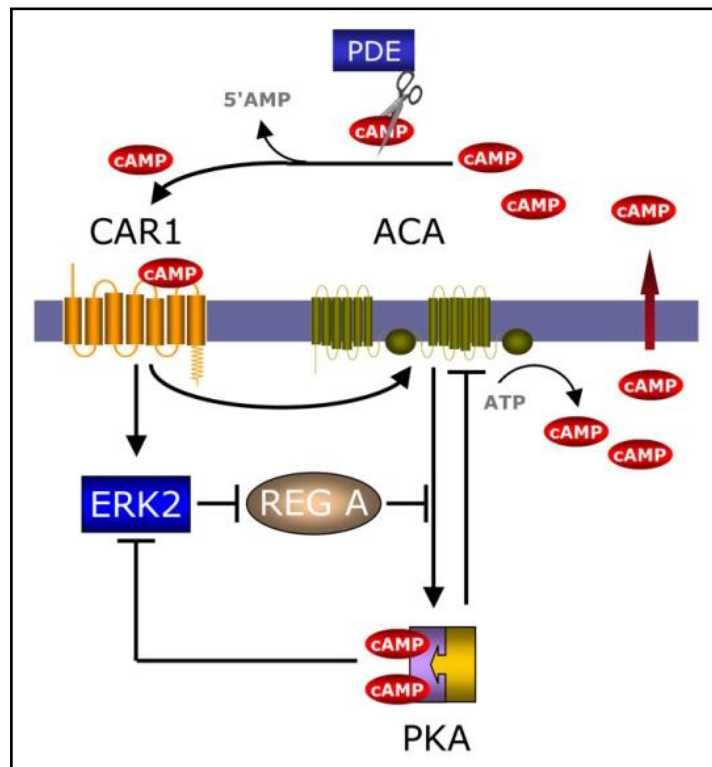
**Abbildung 1: Darstellung der multizellulären Entwicklung von *D. discoideum*; (A) schematische Darstellung nach Chisholm und Firtel (2004); (B) durch Scanning-Elektronen-Mikroskopie erstelltes Bild des Entwicklungszyklus; Copyright, M.J. Grimson & R.L. Blanton Biological Sciences Electron Microscopy Laboratory, Texas Tech University.**

Das sich an die Aggregation anschließende Stadium wird *Mound* genannt. Aus diesen Mounds kann es zur Ausbildung von *Slugs* kommen. Dabei kommt es zur vertikalen Ausdehnung der Aggregate (*Finger*), welche nach vorne überkippen. Der nacktschneckenförmige Zellverband kann sich auf einer aus Zellulose und Proteinen gebildeten Scheide photo- und thermotaktisch bewegen. Diese Bewegung wird sehr wahrscheinlich ebenfalls durch den Botenstoff cAMP gesteuert (Dormann *et al.* 1998). Die Entscheidung über die Ausbildung und Art der *Slugs* ist multifaktoriell (Kellermann und McNally 1999).

Die Kulmination entweder direkt aus dem *Mound*-Stadium oder über den Zwischenschritt der *Slug* führt zur terminalen Differentiation der Zellen. Innerhalb der Prä-Stielzellen (pst) können dabei mindestens 4 Subpopulationen von Zellen anhand der Expression bestimmter Marker differenziert werden. PstA und Pst0 unterscheiden sich in der Expression des Gens *ecmA*, welches für ein Protein der extrazellulären Matrix und Bestandteil des Stielgerüsts codiert (Early *et al.* 1993). PstAB-Zellen exprimieren neben *ecmA* auch *ecmB*, dessen Protein ähnliche Eigenschaften wie *Ecma* besitzt (Jermyn und Williams 1991). Der vierte Zelltyp, die *Anterior-like cells*, variieren in der Expression der beiden Marker. Diese Zellen bilden die sogenannten *Upper* und *Lower Cups*, welche die Verbindung zwischen Sporen und Stiel darstellen (Jermyn und Williams 1991). Zusammen mit den anderen Stiel-Zelltypen bilden sie auch die *Basal disk*, die Verbindung zwischen Substrat und Stiel aus (Jermyn *et al.* 1996). Der Stiel selber besteht aus PstA, AB und O-Zellen. Die Zellen des Stiels vakuolisieren und sterben kontrolliert ab. Die Prä-Sporenzellen können von den Pst-Zellen durch die Expression der Gene *cotA*, *cotB* und *cotC* unterschieden werden (Escalante und Vicente 2000). Sie werden von den Stielzellen nach oben geschoben und dort durch den *Upper* und *Lower Cup* gehalten. Beim Eingang in das dormante Stadium sinkt der Wassergehalt der Sporen und es wird eine Mukosaccharid-Kapsel ausgebildet. Unter günstigen Umweltbedingungen können die Sporen innerhalb weniger Stunden wieder in die vegetative Wachstumsphase übergehen (Cotter *et al.* 1992).

### 1.3 Initiation und Regulation der frühen multizellulären Entwicklung

Die Einleitung und frühen Stadien der multizellulären Entwicklung wird durch ein diffiziles regulatorisches Netzwerk gesteuert. Teile dieses Netzwerkes sind in Abbildung 2 dargestellt. In Winckler *et al.* 2001 und 2004 konnte eine Abhängigkeit dieses Netzwerkes von dem in dieser Arbeit untersuchten CbfA gezeigt werden. Deshalb soll nachfolgend ein Überblick über dieses Thema gegeben werden.



**Abbildung 2: Schematische Darstellung des PKA-Regelkreises zur Initiation der Entwicklung in *Dictyostelium discoideum*;** Beschreibung siehe Text; beteiligte Proteine: ACA = Adenylcyclase A, CAR1 = cAMP-Rezeptor 1, PKA = Proteinkinase A (Serin-Threonin Kinase), REG A = intrazelluläre cAMP-Phosphodiesterase, ERK2 = Extracellular Signal Regulated Kinase 2 (MAP-Kinase), PDE = extrazelluläre cAMP-Phosphodiesterase

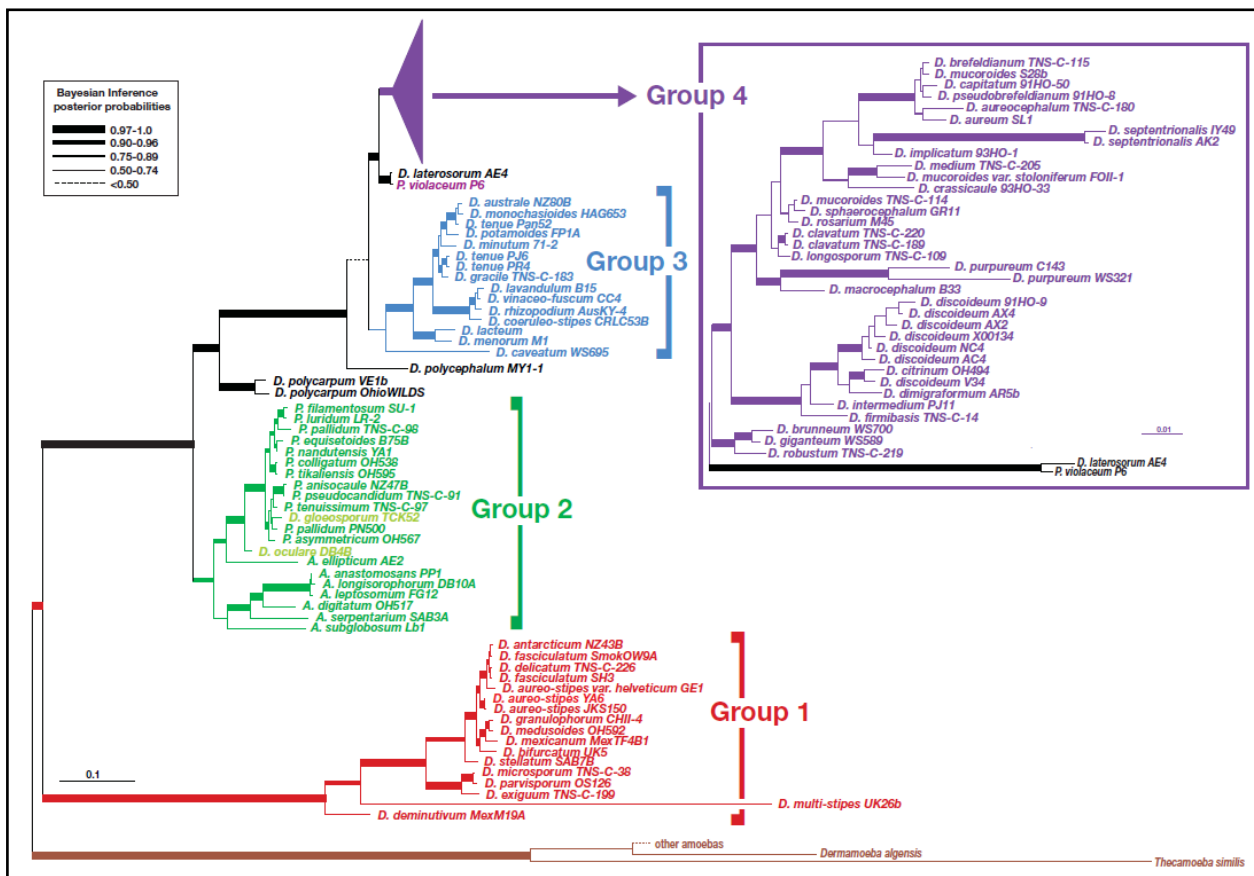
Die Initiierung der Entwicklung wird durch den Botenstoff cyclisches Adenosin-3',5'-monophosphat (cAMP) gesteuert (u.a. Devreotes 1994, Firtel 1995). Senden mehrere Zellen in der Umgebung dieses Signal aus, kommt es zur vermehrten Bindung des cAMPs an den Rezeptor CAR1 (Klein 1988). Es handelt sich dabei um einen Rezeptor mit sieben Transmembrandomänen, welcher an ein heterotrimeres G-Protein gekoppelt ist. Durch

Ligandenbindung kommt es zur Freisetzung der G $\beta$ -Untereinheit von der G $\alpha$ -Untereinheit, was zusammen mit dem Protein CRAC (Insall *et al.* 1994) zur verstärkten Bildung der Adenylylcyclase A (AcaA) führen. Die Aktivität der AcaA erreicht nach 90 Sekunden ihren Höhepunkt. Es kommt zur Umwandlung von ATP zu cAMP, welches teilweise sezerniert wird. Das Signal wird dadurch an benachbarte Zellen weiter gegeben. In der das Signal aussendenden Zelle kommt es ferner zu einer Erhöhung des intrazellulären cAMP-Spiegels (Loomis 1999). Das sich selbst verstärkende cAMP-Signal wird durch die Proteinkinase A (PKA) beendet (Gerisch *et al.* 1979). Diese Serin-Threonin-Kinase liegt in vegetativen Zellen als inaktives Heterodimer vor. Die katalytische Untereinheit PKA-C wird durch die regulatorische Untereinheit PKA-R gebunden. Kommt es nun zur Erhöhung der intrazellulären cAMP-Konzentration über 100 nM bindet cAMP an PKA-R, welches die nun aktive Form der PKA-C frei gibt. Diese inhibiert die Adenylylcyclase A. Es wird über mehrere Minuten kein cAMP mehr gebildet. Die Blockierung der Funktion der AcaA wird durch Senkung des intrazellulären cAMP-Spiegels durch das Enzym cAMP-Phosphodiesterase (RegA) aufgehoben. Durch die Umwandlung des cAMPs in AMP kommt es zur Inaktivierung der PKA (Shaulsky *et al.* 1998). Die AcaA kann nun durch extrazelluläres cAMP wieder aktiviert werden. Dies erklärt die für *Dictyostelium discoideum* typischen cAMP-Pulse in einer Frequenz von 5 bis 10 Minuten. Es kommt zur Ausbildung der Aggregate. Das zur Initiierung der Entwicklung genutzte cAMP-System wird durch einen von der MAP-Kinase ERK2 abhängigen Regulationsweg kontrolliert. ERK2 wird in Abhängigkeit von CAR1 cAMP-abhängig aktiviert (Aubry *et al.* 1997). Dadurch kommt es durch Phosphorylierung zur Inaktivierung der Phosphodiesterase RegA. Die Aktivität der PKA bleibt länger erhalten, AcaA länger inaktiv. Das System kommt zum Erliegen. Die dauerhafte Aktivität der PKA inaktiviert ERK2. Dies erfolgt über die Ras-Proteine Ras-GEF und Ras-GAP (Aubry *et al.* 1997 und Insall *et al.* 1996). Die sezernierte Phosphodiesterase PDE baut das extrazelluläre cAMP ab. Der Botenstoff cAMP ist an weiteren Prozessen in der Entwicklung, z.B. der Migration im *Slug*-Stadium (Dormann *et al.* 1998), beteiligt. Die entwicklungsabhängige Genregulation wird durch die Aktivierung der PKA eingeleitet (Thomason *et al.* 1998).



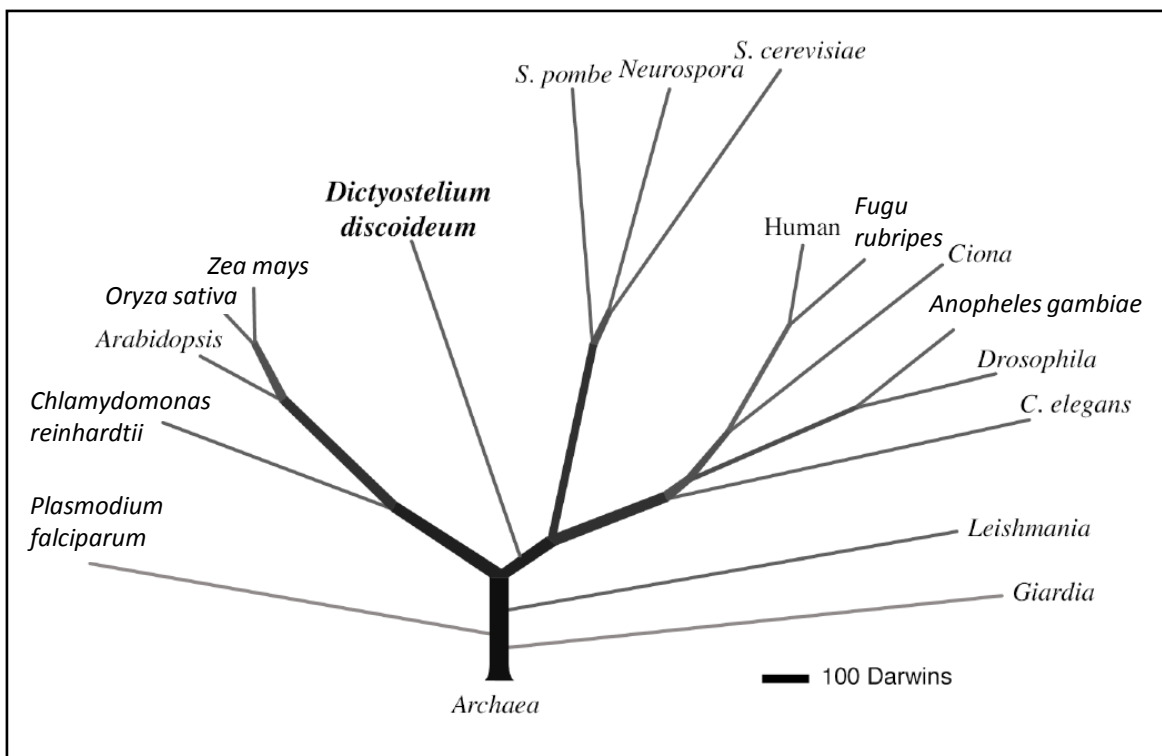
### 1.4 Phyletische Einordnung der Art *Dictyostelium discoideum*

*D. discoideum* gehört zu der Ordnung der *Dictyosteliales*, welche nur aus der Familie der *Dictyosteliidae* besteht. Diese Familie umfasst 89 Arten der Gattungen *Dictyostelium*, *Polysphondylium* und *Acytostelium*. Die ursprüngliche Einteilung der Gattungen erfolgte nach der Morphologie der Fruchtkörper. *Dictyostelium*-Arten besitzen unverzweigte oder lateral verzweigte Fruchtkörper, die Gattung *Polysphondylium* sich regelmäßig wiederholende, verzweigte Fruchtkörper. *Acytostelium*-Arten zeichnen sich durch zellfreie Stiele ihrer Fruchtkörper aus (Raper 1984, Schaap *et al.* 2006). Neuere Untersuchungen unter Nutzung molekularbiologischer Methoden zeigen vier distinkte Gruppen von Arten (Alvarez-Curto *et al.* 2005 und Schaap *et al.* 2006). Die in Abbildung 3 dargestellte Phylogenie ist auf Basis der ribosomalen RNA von 74 Arten und über 100 Isolaten erstellt worden.



**Abbildung 3: Phylogenie der *Dictyosteliidae***; gezeigt ist der durch *Bayesian inference* entstandene Baum auf Grundlage von 1.655 Nukleotiden der 17S rRNA; Außengruppe: Vertreter anderer *Amoebozoa*; Gruppe 4 im Vergleich zu Gruppen 1 bis 3 im anderen Maßstab dargestellt; Abbildung aus Schaap *et al.* (2006)

Die phylogenetisch älteste Gruppe 1 umfasst dabei ausschließlich Arten der Gattung *Dictyostelium*, welche sich teils sehr stark in ihrer Fruchtkörpermorphologie unterscheiden. Die Gruppe 2 besteht aus allen Arten der Gattung *Acytostelium*, wobei bis auf *A. ellipticum* hier von einem Monophylum auszugehen ist. Des Weiteren sind zwei *Dictyostelium*-Arten und fast alle *Polysphondylium*-Arten in dieser Gruppe vorhanden. Gruppe 3 wird ausschließlich aus Arten der Gattung *Dictyostelium* gebildet, darunter auch die im Bezug auf andere *Dictyosteliales* kannibalistische Art *D. caveatum*. Der Kern der Gruppe 4 besteht aus im Vergleich zu den anderen Gruppen phylogenetisch und morphologisch sehr nah verwandten *Dictyostelium*-Arten, darunter *D. discoideum* und die als erstes beschriebene Art *D. mucuroides*. Ferner wurden die Arten *D. laterosporum* und *P. violaceum* der Basis der Gruppe 4 zugeordnet. Beide Arten bilden violett erscheinende Fruchtkörper.



**Abbildung 4: Proteombasierende Phylogenie der Eukaryonten;** der Stammbaum basiert auf 5.279 orthologen Proteinclustern der 17 dargestellten, komplett sequenzierten Organismen; die Wurzel wird aus 159 Proteinclustern aus 6 Archaea-Proteomen gebildet; 1 Darwin = 1/2.000 des genetischen Unterschiedes zwischen Hefen und Menschen; verändert nach Eichinger *et al.* (2005)

Die Stellung der *Dictyosteliidae* im System des Lebens wurde ebenfalls erst in jüngerer Vergangenheit aufgelöst. Nach der anfänglichen Einteilung als Vertreter der Pilze (Brefeld 1869) wurden die Dictyosteliden später zusammen mit den plasmoidalen Schleimpilzen (*Myxogastridae*) und den Protosteliden zu den Schleimpilzen (*Mycetozoa*) zusammengefasst (Olive und Stoianovitch 1975, Loomis *et al.* 1995, Baldauf und Doolittle 1997). Daher stammt die auch heute noch geläufige deutsche Bezeichnung zellulärer Schleimpilz. Alle Schleimpilze haben die Eigenschaft, einen definierten Entwicklungszyklus durchlaufen zu können. Der Ursprung der Schleimpilze liegt nach der Abzweigung der *Viridiplantae* (Pflanzen) von der Entwicklungslinie der Eukaryonten, aber noch vor der Auftrennung der *Metazoa* und *Mycota* (Baldauf *et al.* 2000, Eichinger *et al.* 2005). In Abbildung 4 ist die Einordnung von *D. discoideum* in das System der Eukaryonten dargestellt. Die wissenschaftliche Bezeichnung sowohl im Deutschen als auch im Englischen lautet heute sozialen Amöben (*social amoeba*). Die Bezeichnung leitet sich von der Ausbildung multizellulärer Fruchtkörper aus unizellulären Amöben ab.

### **1.5 Bedeutung von *Dictyostelium discoideum* in der Forschung**

Der Organismus *Dictyostelium discoideum* bietet die Möglichkeit der funktionellen Analyse von Genprodukten in einem sequenzierten, kontrollierten und genetisch manipulierbaren Umfeld. Dabei sind im Vergleich zur Kultur von Säugetierzelllinien mit relativ beschränkter Ausstattung in wenigen Tagen hohe Zellmengen zu erreichen (Watts und Ashworth 1970). *D. discoideum* besitzt einen höheren Anteil homologer Proteine zu den Vertebraten als andere genutzte Modellorganismen außerhalb des Tierreiches (Eichinger *et al.* 2005). Zahlreiche in menschlichen Krankheiten involvierte Faktoren und Vorgänge sind vorhanden. Durch moderne molekularbiologische Methoden sind dabei viele Prozesse in Echtzeit sowohl verfolgbar als auch darstellbar. Einige allgemeingültige Zusammenhänge wurden erstmalig bei *D. discoideum* beschrieben. Dazu gehören die überlappende Funktion von Proteinen des Zytoskeletts (Eichinger *et al.* 1999), die funktionelle Bedeutung und Charakterisierung des Motorproteins Myosin (DeLozanne und Spudich 1987, Metha *et al.* 1999), die Vielfalt der Zwei-Komponenten-Systeme in der Signaltransduktion (Wang *et al.* 1996) und der Aufbau sowie die Funktionsweise von G-Protein gekoppelten Rezeptoren (Maeda und Firtel 1997, Plyte *et al.* 1999, Leung *et al.* 1999).

Die Ernährungsweise von *D. discoideum* in seiner vegetativen Form über Phagozytose zeigt viele Homologien zu Teilen der zellulären Immunantwort der Vertebraten. Grundlegende Erkenntnisse über diesen Prozess und die Möglichkeit von bakteriellen Krankheitserregern, diesen Vorgang zu ihren Vorteil auszunutzen, wurden an diesem System beschrieben (u.a. Farbrother *et al.* 2006, Na *et al.* 2007). Des Weiteren bietet *D. discoideum* durch seinen vom vegetativen Wachstum deutlich abgetrennten, multizellulären Entwicklungszyklus die Möglichkeit, grundlegende Entwicklungsprozesse zu untersuchen. Obwohl sich der Mechanismus der multizellulären Entwicklung mechanistisch zu dem der *Metazoa* unterscheidet, können in diesem System Prozesse wie Zell-Zell-Kommunikation, Zelltypdeterminierung und Morphogenese untersucht werden. Auch in diesem Zusammenhang sind einige grundlegende Vorgänge in *D. discoideum* erstmalig beschrieben worden. Dazu gehören die Erkennung des extrazellulären cAMPs über G-Protein gekoppelte Rezeptoren (Sun und Devreotes 1991), die Verwendung des chlorinierten Alkylphenons DIF zur Festlegung des Zelltyps (Kay und Jermyn 1983, Insall *et al.* 1992, Kubohara 1999) sowie Mechanismen der räumlichen Differenzierung (Überblick in Escalante und Vicente 2000).

### **1.6 Genomische Organisation von *Dictyostelium discoideum***

Das komplette Genom von *D. discoideum* ist bekannt (Eichinger *et al.* 2005) und über die Plattform [www.dictybase.org](http://www.dictybase.org) der Öffentlichkeit zugänglich. Das Genom des untersuchten Stammes AX4 (Knecht *et al.* 1986) besitzt etwas mehr als 34 Millionen Basenpaare. Im Gegensatz zu dem in dieser Arbeit verwendeten Stamm AX2 entstanden der Stamm AX3 und dessen direkter Nachkomme AX4 einer chemischen Mutagenese. Es ist zu erwähnen, dass der sequenzierte Stamm AX4 im Gegensatz zu AX2 eine etwa 750.000 bp lange Verdopplung eines Bereiches auf Chromosom 2 besitzt. Dieser schließt sich dem Ausgangsbereich direkt an und ist in AX2 nicht vorhanden. Das Genom ist haploid und umfasst sechs Chromosomen im Größenbereich von 3,5 bis 8,6 Millionen Basenpaaren. Ein oftmals beobachtetes, potentiell siebtes Chromosom in Präparationen vegetativer Zellen ist ein extrachromosomales Konkatermer eines palindromischen rDNA-Elementes. Es besteht aus 88.000 bp großen Einzelementen mit einer Kopienzahl von etwa 100. Das Ursprungselement des Palindroms befindet sich auf Chromosom 4 (Sucgang *et al.* 2003).

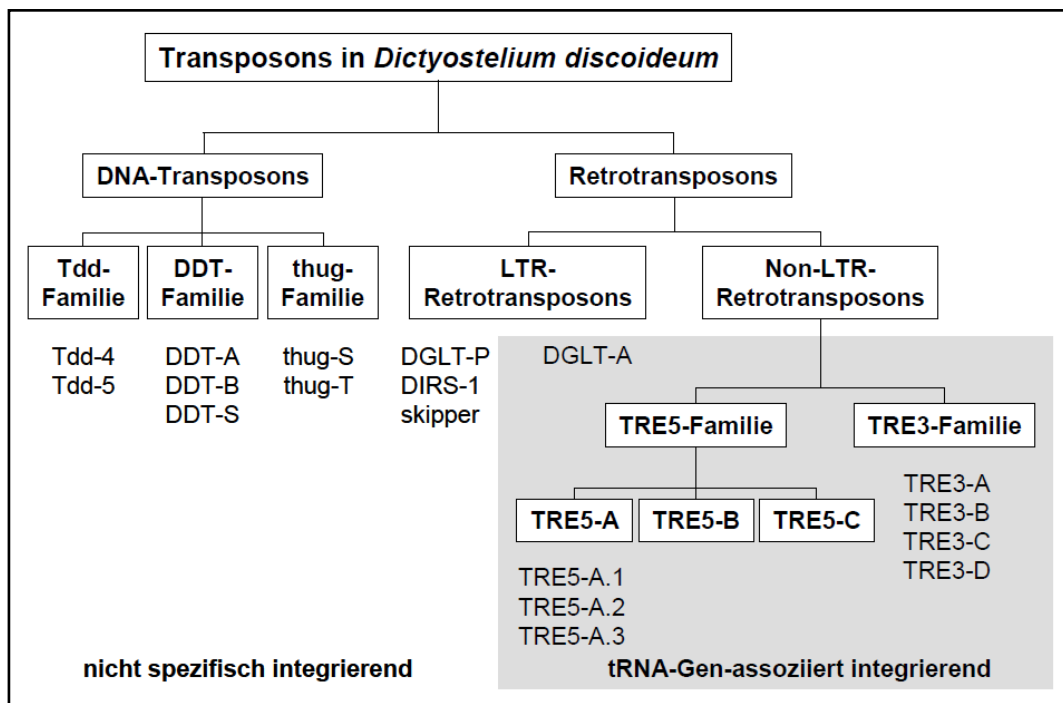
*D. discoideum* und einige andere soziale Amöben haben einen sehr geringen GC-Anteil im Genom. In *D. discoideum* beträgt er 22,4% (Dutta und Mandel 1972).

Das Genom von *D. discoideum* hat neben dem geringen GC-Gehalt einige weitere spezielle Eigenschaften, besonders im Vergleich zum menschlichen Genom. Das Genom ist um den Faktor 90 kleiner als das menschliche Erbgut, besitzt aber mit 13.605 Genen (Dictybase, nach Eichinger *et al.* 2005 etwa 12.500) mehr als die Hälfte der Anzahl menschlicher Gene (etwa 22.000, International Human Genome Sequencing Consortium 2004). In dem Genom von *D. discoideum* sind 390 tRNA-Gene enthalten. Die durchschnittliche Länge der codierenden Sequenz eines Gens ist dabei vergleichbar um 1.400 bp. Der Bereich zwischen zwei Genen, welcher zusätzlich Promotoren und andere regulatorische Elemente umfasst, ist bei *D. discoideum* 750 bp lang, im Menschen etwa 60.000 bp. Jede Zelle besitzt etwa 200 Mitochondrien, welche jeweils ein Genom von 55.600 bp Länge besitzen. Das mitochondriale Genom codiert für 41 Gene. Am stärksten wird der Unterschied der Genome durch den Anteil codierender Sequenzen im Genom sichtbar. Beim Mensch liegt er bei etwa 1%, bei *D. discoideum* bei 60%. Die Kompaktheit ist hauptsächlich dem haploiden Chromosomensatz geschuldet, welcher jegliche nachteiligen Ereignisse auf das Genom zur Entfaltung bringt. Der Anteil der durch Transposition entstandenen Sequenzen beträgt bei *D. discoideum* etwa 10% (Glöckner *et al.* 2001), beim Menschen 45% (International Human Genome Sequencing Consortium 2004).

### **1.7 Transposable Elemente in *Dictyostelium discoideum***

Da das in dieser Arbeit untersuchte Protein C-Modul-bindender Faktor A (CbfA) über seine Interaktion mit dem Retrotransposon TRE5-A erstmalig beschrieben wurde (Geier *et al.* 1996), soll nachfolgend auf die in *D. discoideum* vorhandenen Transposons eingegangen werden. Mobile genetische Elemente wurden seit ihrer ersten Beschreibung durch McClintock (1948) in fast allen untersuchten Organismen nachgewiesen. Als Transposons werden genetische Elemente bezeichnet, die ihre Lokalisation im Genom ändern können, aber nicht in der Lage sind, infektiöse Partikel zu bilden. Dabei unterscheidet sich der Gehalt der Transposons am gesamten Genom von Organismus zu Organismus. Während er bei *Arabidopsis thaliana* nur bei 0,1% liegt (Konieczny *et al.* 1991), sind es bei *Zea mays* bis zu 85% (Kumar und Bennetzen 1999).

*Dictyostelium discoideum* hat einen Transposongehalt von etwa 10% (Glöckner *et al.* 2001), der Anteil am menschlichen Genom liegt bei 45% (Hedges und Batzer 2005). Die zur Transposition eines Elementes benötigten Enzyme sind innerhalb des Transposons codiert oder werden durch andere Transposons gebildet. Des Weiteren sind in bisher untersuchten Systemen immer zelluläre Prozesse wie z.B. DNA-Reparaturmechanismen beteiligt (Zingler *et al.* 2005).



**Abbildung 5: Übersicht über die Transposons in *Dictyostelium discoideum*;** nach Winckler *et al.* (2005), Abbildung aus Siol (2006c) übernommen; grauer Kasten zeigt Transposons mit tRNA-assoziiierter Integration, Erläuterung im Text

Transposons können anhand ihres Mechanismus in zwei Klassen eingeteilt werden (Craig *et al.* 2002, siehe Abbildung 5). Gemeinsam haben beide Klassen, dass es bei der Integration des Transposons meist zu einer Verdopplung einer kurzen Zielsequenz durch einen versetzten Schnitt in der DNA kommt. Diese Verdopplung ist transposonspezifisch (Gilbert *et al.* 2002). Die DNA-Transposons werden aus einem Locus ausgeschnitten und integrieren in einen anderen Locus. Dabei bleibt die Kopienzahl des Elementes konstant. In *D. discoideum* können drei Familien von DNA-Transposons mit insgesamt sieben Vertretern identifiziert werden. Es existieren 235 Kopien von DNA-Transposons mit einem Anteil von zusammen 1,5% am Genom (Glöckner *et al.* 2001). Alle vollständigen DNA-Transposons codieren für eine multifunktionelle

Transposase. Die zweite Klasse von Transposons sind die Retrotransposons. Im Gegensatz zu den DNA-Transposons benötigen sie einen RNA-Zwischenschritt. Das Transposon wird in seiner vollen Länge in eine mRNA transkribiert. Für die Transposition wichtige Proteine werden aufgrund dieser Information gebildet. Darunter befindet sich bei Retrotransposons genau wie bei den strukturell ähnlichen Retroviren eine reverse Transkriptase (Malik *et al.* 1999). Diese wandelt die Information der mRNA des Transposons in eine komplementäre DNA (cDNA) um. Diese cDNA wird dann mit Hilfe transposoncodierter Proteine, aber auch Wirtsfaktoren, in das Genom integriert. Es kommt dadurch zur Vergrößerung des Genoms, da das Ausgangselement erhalten bleibt.

In *D. discoideum*, wie auch in anderen Organismen, können zwei Gruppen von Retrotransposons unterschieden werden. Die LTR-Retrotransposons werden von zwei 100 bis 500 bp langen Sequenzen flankiert, den Long Terminal Repeats. Diese besitzen sowohl eine Promotor- als auch eine Terminatorfunktion und tragen dazu bei, dass diese Elemente mit einer hohen Wahrscheinlichkeit komplett integrieren. Es sind vier Vertreter dieser Gruppe mit etwa 300 Kopien im Genom von *D. discoideum* nachgewiesen. Diese machen 4,5% des Genomes aus. Dabei entfallen 3,3% Anteil am Genom auf die 235 Kopien des in seinen Aufbau für ein LTR-Transposon eher untypischen DIRS-1-Element (Glöckner *et al.* 2001). Die *D. discoideum* LTR-Transposons bestehen aus ein bis vier offenen Leserahmen.

Die Vertreter der Non-LTR-Retrotransposons in *Dictyostelium discoideum* werden aufgrund ihrer spezifischen Integration in der Umgebung von tRNA-Genen als *tRNA gene-targeted retroelements* (TRE) bezeichnet. Anhand der Integration stromaufwärts (5' oder *upstream*) oder stromabwärts (3' oder *downstream*) von tRNA-Genen spricht man von TRE5 oder TRE3-Elementen. Beide Elementgruppen besitzen in ihren jeweiligen vollständigen Vorgänger zwei offene Leserahmen (ORF). Es sind vier Vertreter der TRE3-Familie in etwa 150 Kopien vorhanden. Sie machen 2,2% des Genoms der sozialen Amöbe aus. Die Integration erfolgt 40-150 bp *downstream* der tRNA-Gene (Winckler *et al.* 2002).

Die TRE5-Familie umfasst drei Vertreter, wobei das am besten untersuchte TRE5-A in drei Formen unterschiedlicher Länge vorkommt. In Abbildung 6 ist das 5.659 bp lange TRE5-A.1

schematisch dargestellt. Die Besonderheit der TRE5-Transposons ist der modulare Aufbau der untranslatierten Regionen (UTR). Am 5' Ende befindet sich mit dem A-Modul ein elementinterner Promotor, welcher für die Bildung der Plusstrang-RNA verantwortlich ist. Die Transkription umfasst das komplette A-Modul. Daran schließt sich das B-Modul an. In diesem Bereich befindet sich der Translationsstart des ORF1-Proteins. Zu Beginn der 3'-UTR wird das B-Modul wiederholt. Es existiert eine Indizienkette, die daraufhin deutet, dass es sich bei dem B-Modul um den ursprünglichen Promotor sowohl für die Plus- als auch für die Minusstrang-RNA des TRE5-A-Elementes handelt (Schumacher *et al.* 1994, Winckler *et al.* 2005). Ferner existiert am 3' Ende das C-Modul. Es handelt sich dabei um einen aktiven Promotor, welcher für die Transkription der Minusstrang-RNA verantwortlich ist. Entgegen der Erwartung einer negativen Auswirkung von *Antisense*-RNA auf die Transposition durch die vorhandene RNAi-Maschinerie (Martens *et al.* 2002), scheint diese die Integration von TRE5-A zu begünstigen.



**Abbildung 6: Schematische Darstellung der Struktur von TRE5-A.1;** aus Winckler *et al.* (2005); UTR = untranslatierte Region, EN = Endonuklease, RT = Reverse-Transkriptase, HC = histidin- und cysteinreiche Domäne

Die Suche nach den mit dem C-Modul interagierenden Proteinen in *D. discoideum* führte zur Entdeckung des C-Modul-bindenden Faktors A (Geier *et al.* 1996). Diese Interaktion wurde *in vitro* validiert (Horn *et al.* 1999). TRE5-A.1 codiert für zwei offene Leserahmen. Das ORF1-Protein ist 51 kDa groß und durch Protein-Protein-Interaktionen mit Untereinheiten des Transkriptionsfaktors IIIB für die tRNA-spezifische Integration verantwortlich (Siol *et al.* 2006b, Chung *et al.* 2007). Das ORF2-Protein besitzt drei funktionelle Domänen mit Endonuklease- und Reverse-Transkriptaseaktivität. Die Funktion der histidin- und cysteinreichen Domäne ist nicht geklärt. TRE5-A-Elemente sind in aktuell kultivierten Zelllinien aktiv (Beck *et al.* 2002, Siol *et al.* 2006b). Es existieren etwa 90 Kopien der TRE5-Elemente, welche circa 1,5% des Genoms ausmachen. Die Integration erfolgt 44 bis 54 bp *upstream* von tRNA Genen (Szafranski *et al.* 1999, Winckler *et al.* 2002, Siol *et al.* 2006b).

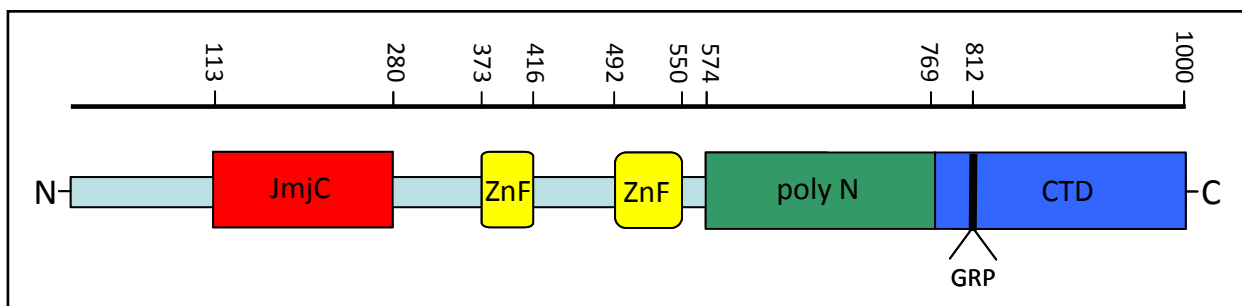


Es muss erwähnt werden, dass auch das LTR-Retrotransposon DGLT-A spezifisch 10-30 bp upstream von tRNA-Genen integriert (Hofmann *et al.* 1991, Winckler *et al.* 2005). Der Integrationsmechanismus ist dabei wie bei den TRE3-Elementen noch nicht aufgeklärt.

### **1.8 Aufbau des C-Modul-bindenden Faktors A**

Arbeiten über das Retrotransposon TRE5-A.1 in *Dictyostelium discoideum* zeigten die Existenz von drei unterschiedlichen, als A-Modul, B-Modul und C-Modul bezeichnete Sequenzbereiche in den untranslatierten Regionen des Transposons (siehe Abschnitt 1.7, Marschalek *et al.* 1992a und 1992b). Es konnte eine Polymerase-II-abhängige Promotoraktivität für diese Module gezeigt werden (Schumann *et al.* 1994). Elektrophoretische Gelretardierungsversuche zeigten die spezifische Bindung eines in Kernextrakten von *Dictyostelium discoideum* vorhandenen Proteins an das für die Bildung der Antisense-RNA des TRE5-A.1 verantwortliche C-Modul (Geier *et al.* 1996). Dieses Protein wurde als C-Modul-bindender Faktor A (CbfA) bezeichnet. Die anfänglich genutzte Proteinbezeichnung CMBF ist aufgrund der Verwechslungsgefahr mit dem Conditioned Medium Factor (CMF) in CbfA geändert worden. Das CbfA-codierende Gen *cbfA* (Gen ID nach Dictybase: DDB\_G0279409) wurde durch Microsequenzierung und anschließendes Screening einer genomischen DNA-Bank (Horn *et al.* 1999) identifiziert. Die DNA-Sequenz ist ferner über den Genbank®-Eintrag AF052006.2 zugänglich. Das Gen *cbfA* befindet sich auf dem Chromosom 3, Position 2.308.619 bis 2.312.032. Die codierende Sequenz ist 3.003 bp lang und wird durch ein 410 bp langes Intron unterbrochen. Das Intron beginnt nach dem für die Aminosäure 60 codierenden Basentriplett an der Position 181. Das Protein CbfA ist 1.000 Aminosäuren lang und hat eine Masse von 114.270 Dalton.

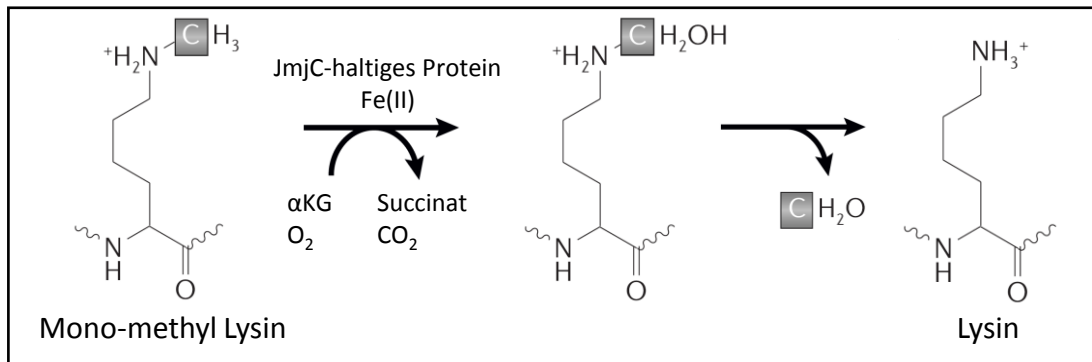
Anhand der Aminosäuresequenz lassen sich mehrere strukturelle und funktionelle Domänen innerhalb von CbfA ableiten (Abbildung 7). Am Amino-terminalen Ende befindet sich eine Jumonji-Domäne vom Carboxy-terminalen Typ (JmjC). Der Kernbereich der JmjC erstreckt sich über die Aminosäuren 113-280 von CbfA. JmjC-Domänen wurde erstmalig in dem Mausprotein Jumonji identifiziert (Takeuchi *et al.* 1995). Die Einteilung der Jumonji-Domänen nach Amino- und Carboxy-terminal bezieht sich auf das gemeinsame Vorkommen dieser Domänen in einem Protein (Klose *et al.* 2006). In Proteinen, in denen nur eine Jumonji-Domäne existiert, ist diese immer vom Carboxy-terminalen Typ. Die JmjC-Domäne zeigt strukturelle Ähnlichkeiten zu Metalloenzymen vom Cupintyp (Clissold und Ponting 2001). Die Kristallstruktur einiger JmjC-Domänen ist gelöst, u.a. die eines Vertreters der JMJD2-Familie (Chen *et al.* 2006).



**Abbildung 7: Schematische Darstellung des Domänenstruktur des C-Modul-bindenden Faktors A;** JmjC = Jumonji Domäne vom Carboxy-terminalen Typ, angegebener Bereich ist Kerndomäne; ZnF = Zinkfinger, angegebener Bereich gibt das erste und letzte identifizierte Cystein oder Histidin des potentiellen Zinkfingermotives an; poly N = Asparagin-reicher Bereich; CTD = Carboxy-terminale Domäne, beinhaltet AT-Haken mit dem zentralen GR<sub>812</sub>P-Motiv sowie einen an Threonin und Serin reichen Bereich von AS 771-789.

Über die Funktion der JmjC-Domänen gab es besonders in den Jahren 2006 und 2007 eine Flut von Veröffentlichungen. Es wurde dabei hauptsächlich die Modulation des epigenetischen Histoncodes durch JmjC diskutiert. Tsukada und Kollegen zeigten die Lysindemethylaseaktivität der JmjC-Domäne des humanen *JmjC domain-containing histone demethylase 1*-Proteins (JHDM1, Tsukada *et al.* 2006) an Histonen. Die Demethylierungsaktivität einer JmjC-Domäne an Argininresten von Histonen wurde erstmalig für das murine Protein *Jumonji domain-containing 6* (JMJD6, Chang *et al.* 2007) gezeigt. Der Mechanismus der Demethylierungsreaktion eines einfach methylierten Lysinrestes ist in Abbildung 8 dargestellt. Dabei ist der Prozess an ein-, zwei- und dreifach methylierten Lysinen und Argininen identisch (Klose *et al.* 2006, Chang *et al.*

2007). Die JmjC-Domäne überträgt in einer Eisen(II)-abhängigen Dioxygenase-Reaktion je ein Sauerstoffatom auf  $\alpha$ -Ketoglutarat und die Methylgruppe des zu demethylierenden Restes. Es entstehen Succinat, Kohlendioxid und eine hydroxylierte Methylgruppe. Diese Hydroxymethylgruppe wird spontan als Formaldehyd abgegeben.



**Abbildung 8: Chemischer Mechanismus der Demethylierung von Arginin- und Lysinresten durch JmjC-haltige Proteine;** Erläuterung im Text; Fe(II) = zweiwertiges Eisen,  $\alpha$ KG =  $\alpha$ -Ketoglutarat; Verändert nach Klose *et al.* (2006)

Neben der Histondemethylaseaktivität von JmjC-Domänen vor allem in Proteinen der JHDM-Familien sind weitere Funktionen der Domäne denkbar. Bekannt ist die Funktion als Plattform für Protein-Protein-Interaktionen zur Bildung höherer Komplexe. Das erste beschriebene Beispiel dafür ist das Hefeprotein Gis1, dessen Jumonji-Domäne mit mindestens 19 anderen Proteinen interagiert. Die Interaktionspartner sind funktionell in der DNA-Reparatur, Transkription und in der Sumoylierung von Proteinen anzusiedeln (Tronnorsjö *et al.* 2006).

An die JmjC-Domäne von CbfA schließen sich zwei Zinkfinger-Motive an. Der näher am Amino-Terminus gelegene Zinkfinger befindet sich zwischen den Aminosäuren 373 bis 416. Dieser Bereich besitzt insgesamt acht Cysteine und ein Histidin. Zur Koordinierung eines Zinkions sind vier Cysteine oder Histidine notwendig (Borden *et al.* 1995, Brayer und Segal 2008). Welche Aminosäuren von CbfA an der Ausbildung des Zinkfingers beteiligt sind, ist nicht bekannt. Der Amino-terminale Zinkfinger von CbfA kann in keine bekannte Klasse eingeordnet werden und wurde erstmalig durch die Anhäufung von Cysteinen und Histidinen in einem kurzen Sequenzbereich definiert. Es existieren nach heutiger Einteilung 20 verschiedene Zinkfingerfamilien (Brayer und Segal 2008).

Der näher am Carboxy-Terminus gelegene Zinkfinger befindet sich zwischen den Aminosäuren 492 bis 550. In diesen Bereich sind elf Cysteine vorhanden. Es besteht eine geringe Sequenzähnlichkeit zu mehreren Klassen von Zinkfingern, darunter der sogenannten PHD-Typ. Diese Art von Zinkfingern kommt in histonmodifizierenden Proteinen wie etwa Acetyltransferasen vor (Aasland *et al.* 1995).

Zinkfinger wurden ursprünglich als sequenzspezifische Bindungsmotive mit der DNA identifiziert (unter anderen für den C2H2-Typ in Miller *et al.* 1985). Sie sind in allen Reichen des Lebens vorhanden. Ein Protein kann ein bis vierzig Zinkfinger-Motive besitzen (Brayer und Segal 2008). Dabei modulieren Zinkfingerdomänen häufig die Bindung von funktionellen Einheiten, wie etwa von JmjC-Domänen, an die DNA (Klose *et al.* 2006). Das Wissen über die sequenzspezifische Bindung von Zinkfingern ist so umfangreich, dass durch ihre Anordnung eine „Adressierung“ von artifiziellen Proteinen an eine kurze DNA-Sequenz möglich ist (Papworth *et al.* 2006, „ZifBase“: Jayakanthan *et al.* 2009). Neben den Protein-DNA-Interaktionen wurden Zinkfinger als Interaktionsfläche mit RNAs (Hall 2005) und Proteinen (Brayer und Segal 2008) identifiziert.

Der Amino-Terminus von CbfA wird durch eine asparaginreiche Sequenz von der Carboxy-terminalen Domäne von CbfA getrennt. Die sogenannte *poly N-Stretch* befindet sich zwischen den Aminosäuren 574 bis 769 (siehe Abbildung 7). Er beinhaltet 104 Asparagine, welche in drei fast homopolymeren Abschnitten angeordnet sind. Diese werden durch zwei serinreiche Abschnitte unterbrochen. Über die Funktion dieser Domäne kann nur spekuliert werden. Ähnlich wie die in den menschlichen Krankheiten Huntingtin oder Spinocerebellare Ataxie 17 (SCA17) beteiligten Polyglutamin-Bereiche der entsprechenden Proteine, kann die poly N-Domäne eine Flexibilität der Proteinteile vermitteln, welche die Funktionen der anderen Domänen ermöglicht (Friedman *et al.* 2007).

Die in dieser Arbeit im speziellen untersuchte Carboxy-terminale Domäne (CTD) von CbfA beginnt nach dem Ende der asparaginreichen Domäne mit der Aminosäure 770. Die Aminosäuresequenz der CTD ist in Abbildung 9 dargestellt. Der dem Amino-Terminus von CbfA am nächsten gelegene Abschnitt der CTD umfasst einen Bereich, welcher viele Serine und Threonine besitzt. Diese serin- und threoninreichen Sequenzen sind aufgrund der Phosphorylierbarkeit dieser Aminosäuren häufig das Ziel posttranslationaler Regulation durch Kinasen und Phosphatasen (Marnef *et al.* 2009).

<u>770</u>	<u>780</u>	<u>790</u>	<u>800</u>	<u>810</u>	<u>820</u>
<u>PTTSSLSSLS</u>	<u>TSLSSSSTST</u>	QKPMDVHSKK	RPIVLDNDKP	<b>KGRP</b> PKNLKE	WTSTHKFIIS
<u>830</u>	<u>840</u>	<u>850</u>	<u>860</u>	<u>870</u>	<u>880</u>
LIELFRSSNN	AILGKPNPHY	KPIENLPPLV	QLYLSQRKAF	GGVLWAKTNS	CPLLPCIWVK
<u>890</u>	<u>900</u>	<u>910</u>	<u>920</u>	<u>930</u>	<u>940</u>
DLSVIPNTK	LLPSLIQGKK	IVVLFQDQD	QEEYVGIVGK	KSIFSFDEVN	QTLLKCGEV
<u>950</u>	<u>960</u>	<u>970</u>	<u>980</u>	<u>990</u>	<u>1000</u>
PLAQLEDLFN	TTEPEIAMKK	DIAAFNYKNQ	IEEKEEGLYV	KQELYNNKKI	I

**Abbildung 9: Primärsequenz von der Carboxy-terminalen Domäne von CbfA;** die Nummerierung der Aminosäuren bezieht sich auf Gesamt-CbfA; Unterstrichen = Threonin- und Serin-reicher Bereich; **fett** = AT-Haken mit zentralen Arginin 812

An diesen Bereich schließt sich ein AT-Haken an. Diese Einheit ist durch sein zentrales Glycin-Arginin-Prolin-Motiv (GRP) in der Lage, desoxyadenosin- und desoxythymidinreiche DNA-Sequenzen zu binden. Es kommt dabei zu einer Interaktion mit der kleinen Furche der DNA (Arvind und Landsman 1998). AT-haltige Basenpaare sind in ihrer Ausdehnung kompakter als GC-haltige. Der in der CTD von CbfA vorhandene AT-Haken kann aufgrund seiner ihn umgebenden Sequenz dem Typ-III zugeordnet werden (Arvind und Landsman 1998). In dieser Einordnung spielt vor allem die Verteilung basischer Aminosäuren und die von Prolinen eine Rolle. Die Bindung an die DNA erfolgt durch das zentrale Arginin, wobei für den Typ-II-AT-Haken des humanen *high-mobility-group*-Proteins I (HMG-I) die Beteiligung einer weiteren basischen Aminosäure in der Bindung gezeigt werden konnte (Huth *et al.* 1997). Das Wirken des AT-Hakens in der Bindung von Gesamt-CbfA und CbfA-CTD an das C-Modul von TRE5-A.1 konnte *in vitro* durch elektrophoretische Gelretardierungsexperimente gezeigt werden (Horn *et al.* 1999). Dabei konnten sowohl die Wichtigkeit des zentralen Arginins an der Stelle 812 sowie die

Bindung des AT-Hakens an der kleinen Furche der DNA nachgewiesen werden. Des Weiteren konnten zwei für die Bindung von CbfA essentielle Sequenzmotive, CMBS-1 und CMBS-2, identifiziert werden. Diese Sequenzen bestehen in ihrem Kern fast ausschließlich aus Desoxythymidin-Phosphaten, bezogen auf den Minus-Strang. Dies ist insofern überraschend, da *D. discoideum* durch seinen hohen AT-Gehalt sehr viele dieser oligo-dT-Bereiche besitzt. Für die Bindung des C-Moduls ist die CTD von CbfA ausreichend (Horn *et al.* 1999).

Abgesehen von den beiden bekannten, kurzen Strukturelementen ist die Primärsequenz der Carboxy-terminalen Domäne von CbfA einzigartig. Durchgeführte Datenbanksuchen vor der teilweisen Aufklärung der Genome anderer sozialer Amöben zeigten keinerlei Ähnlichkeit zu anderen bekannten Proteinfragmenten.

Im Genom von *Dictyostelium discoideum* existiert ein zu *cbfA* paraloges Gen. Dieses wurde aufgrund seiner Sequenzverwandtschaft als *cbfB* bezeichnet. Es codiert für einen 1.474 Aminosäuren umfassenden, offenen Leserahmen. Dieser zeigt im Bereich der JmjC-Domäne, der Zinkfinger sowie durch das Vorhandensein einer poly N-Domäne Ähnlichkeit zu CbfA. Die Carboxy-terminale Domäne von CbfB ist verlängert und zeigt keine Sequenzverwandtschaft zur CbfA-CTD. CbfB besitzt ferner eine Amino-terminale Verlängerung, welche in CbfA nicht vorhanden ist. Die Funktion von CbfB ist derzeit unbekannt.

### **1.9 Funktion des C-Modul-bindenden Faktors A**

Zur Untersuchung der *in vitro* nachgewiesenen Interaktion von CbfA mit dem C-Modul des Retrotransposons TRE5-A.1 sowie zur weiteren funktionellen Charakterisierung des Proteins wurde der Versuch unternommen, einen CbfA-negativen Stamm zu erstellen. Dies schlug mehrfach fehl, was auf eine für das Überleben der Zelle essentielle Funktion von CbfA hindeuten könnte (Winckler *et al.* 2001). Alternativ wurde deshalb ein Stamm erstellt, welcher CbfA stabil unterexprimiert. Dazu wurde die Amber-Stop-Codon Suppression genutzt. Die dabei entstandenen Stämme zeigen einen ausgeprägten Phänotyp im Vergleich zu dem Wildtyp AX2 in Abhängigkeit von der verbleibenden Menge an funktionellen CbfA (Winckler *et al.* 2001). Der für diese Arbeit genutzte Stamm JH.D hat etwa 5% der Wildtypmenge an Gesamt-CbfA. Es ist zu erwähnen, dass aufgrund der Methodik der Suppression 100% des Amino-terminalen Teils von

CbfA gebildet werden. Eine autonome Funktion dieses Proteinabschnittes kann nicht ausgeschlossen werden.

Der Phänotyp von JH.D im Vergleich zu AX2 zeigt eine erhöhte Pinozytoserate einhergehend mit einem größeren Volumen der Zellen. Die Fähigkeit zur Phagozytose ist gesenkt, was mit dem geringen Wachstum auf Bakterien korreliert. Die Verdopplungszeit in axenischer Kultur ist in der Mutante leicht verlängert. Die Menge an TRE5-A-Transkripten nimmt drastisch ab. JH.D ist nicht in der Lage, den multizellulären Entwicklungszyklus zu durchlaufen. Hungernde Zellen bilden keine Aggregate (Winckler *et al.* 2001).

Der Entwicklungsphänotyp von JH.D konnte auf die fehlende Aktivierung der an der Ausbildung des cAMP-Pulses direkt beteiligten Adenylylcyclase A zurückgeführt werden (siehe Kapitel 1.3, Winckler *et al.* 2004, Siol *et al.* 2006a). Er ähnelt den Entwicklungsphänotypen einiger anderer an der Initiation der Entwicklung beteiligter Faktoren, wie etwa dem von *acaA*<sup>-</sup> oder *pkaC*<sup>-</sup>-Zellen (Pitt *et al.* 1992, Mann und Firtel 1991). Des Weiteren gibt es Ähnlichkeiten zu *mybB*<sup>-</sup>-Zellen, wobei der direkte Zusammenhang hier noch nicht geklärt ist (Otsuka und van Haastert 1998). Das Genprodukt von *mybB* hat Einfluss auf die Expression der Adenylylcyclase A.

Die Bindung von CbfA an den Promotor der Adenylylcyclase A konnte *in vitro* gezeigt werden (Siol *et al.* 2006a). Dabei konnte die für die Bindung benötigte Promotorregion den Bereich -739 bis -448 bezogen auf den Translationsstart von *acaA* zugeordnet werden. Innerhalb dieser Sequenz existieren, ähnlich dem C-Modul von TRE5-A.1, zwei homopolymere dT-Bereiche, welche für die Bindung von CbfA essentiell sind. Die Interaktion zwischen Promotor und CbfA kann durch Stoffe, die in der kleinen Furche der DNA bindenden, verhindert werden. Dies lässt indirekt eine Beteiligung des AT-Hakens von CbfA an der Bindung vermuten.

Initiale Experimente zeigen, dass die heterologe Expression der Carboxy-terminalen Domäne von CbfA in der Suppressionsmutante JH.D den Phänotyp in Bezug auf die TRE5-A-Transkripte revertieren kann. Der Entwicklungsdefekt bleibt jedoch bestehen.

## 1.10 Zielstellung

Aus Vorarbeiten ist zu vermuten, dass der C-Modul-bindende Faktor A (CbfA) aus *Dictyostelium discoideum* wichtige genregulatorische Funktionen sowohl in der vegetativen Wachstumsphase als auch in der multizellulären Entwicklung von *D. discoideum* übernimmt. Ziel dieser Arbeit ist es, diese Funktionen zu identifizieren und eine funktionelle und strukturelle Analyse der an der Genregulation beteiligten Domänen von CbfA durchzuführen.

Dazu soll die differentielle Genexpression vegetativ wachsender Zellen des Wildtyps AX2 mit der dauerhaft CbfA-unterexprimierenden Mutante JH.D anhand der etablierten Methode der *Microarray*-Technologie ermittelt werden. Die erhaltenen Daten werden durch quantitative *Real Time*-PCR (qRT-PCR) verifiziert. Differentiell exprimierte Gene sollen funktionell annotiert werden. Ein weiteres Ziel ist es, die immer schneller und billiger werdenden Methoden zur Sequenzierung von Genomen und Transkriptomen auf das Problem der Identifizierung der CbfA-regulierten Gene anzuwenden. Ein Vergleich dieser Techniken mit den *Microarray*-Daten soll zeigen, inwieweit die neuen Techniken der direkten Transkriptom-Sequenzierung für die Genfunktionsanalyse in *Dictyostelium* geeignet sind. Ferner wird mit Hilfe der steigenden Zugänglichkeit von Genomdaten anderer *Dictyosteliidae* sowie durch PCR-Screenings mit degenerierten Primern nach orthologen *cbfA*-Genen gesucht.

Einzelne unveröffentlichte Experimente lassen eine autonome genregulatorische Funktion der Carboxy-terminalen Domäne (CTD) von CbfA vermuten. Aus diesem Grund soll in die bereits erwähnten *Microarray*-Experimente jeweils eine die CTD exprimierende Transformante der Stämme AX2 und JH.D mit eingeschlossen werden. Ferner ist eine Beteiligung des innerhalb der CTD gelegenen AT-Hakens an der spezifischen Bindung an die DNA belegt (Horn *et al.* 1999). Die Bedeutung des AT-Hakens für die genregulatorische Aktivität der CTD soll *in vivo* näher charakterisiert werden.

Zur strukturellen Charakterisierung und der daraus eventuell resultierenden Ableitung der Funktionsweise der bisher nur in CbfA von *D. discoideum* gefundenen CTD soll eine röntgenkristallographische Strukturaufklärung durchgeführt werden. Ferner wird nach potentiellen Proteininteraktionspartnern der CTD mittels des Hefe-2-Hybrids gesucht werden.



## 2 Material und Methoden

### 2.1 Laborgeräte

Autoklav:	Varioklav Dampfsterilisator (H + P Labortechnik) Hiclave™ HV-85L (HMC)
Blotapparaturen:	PerfectBlue Semi-Dry Elektrobloetter SEDEK M (Peqlab) X Cell Blot Modul (Invitrogen)
Elektrophoresen DNA:	PerfectBlue Gelsystem Mini (Peqlab)
Elektrophoresen Protein:	PerfectBlue Doppelgelsystem Twin M und L (Peqlab)
Elektroporator:	Gene Pulser Xcell mit ShockPod (BioRad)
Geldokumentation:	Gel Jet Imager 2004 (Intas)
Inkubatoren (Schränke):	TS606-G/4-i (WTW) B20 (Heraeus)
Inkubatoren (Schüttler)	Multitron 2 (HT-Infors)
Mikroskope:	Binokular-Mikroskop SZX9 (Olympus) Mikroskop CXX 41 (Olympus) Lichtquelle U-RFL-T (Olympus) Kamera Evolution® MP COLOR (Media Cybernetics)
PCR-Gerät:	Primus 96 advanced (Peqlab)
pH-Wert-Messgerät	pH-Meter pH 526 MultiCal® (WTW)
<i>Real-Time-PCR-System:</i>	Mx3000P (Stratagene)
Sterile Werkbank (LAF):	MRF-B (Steag)
UV/VIS-Spektrometer:	BioPhotometer (Eppendorf) DU®640 Spectrophotometer (Beckman)
Vortexer:	VF2 (Janke & Kunkel, IKA Labortechnik)
Waagen:	E2000D, BA 6100 und PT 310 (Sartorius)

Waagen (Analyse):	A 210 P, BP 210 P und MC 210 P (Sartorius)
Wasseraufbereitung:	Seradest SD 2800 (SERAL Reinstwassersysteme) Serapur Delta (SERAL Reinstwassersysteme)
Zentrifugalevaporator:	Concentrator 5301 (Eppendorf)
Zentrifugen:	Centrifuge 5415 D (Eppendorf) Centrifuge 5810R (Eppendorf) J2-MC (Beckman) Labofuge 400R (Heraeus) Multifuge 1 L-R (Heraeus) Megafuge 1.0R (Heraeus)

## 2.2 Chemikalien

Die verwendeten Chemikalien hatten einen hohen Reinheitsgrad, welcher immer mindestens „pro analysi“ entsprach. Die in Standardverfahren genutzten Chemikalien sind nachfolgend nicht aufgelistet. Sie wurden von folgenden Firmen/Marken bezogen: BD Biosciences, BDH Prolabo, FERAK, Fluka, Invitrogen, Merck, Roth, Oxoid, Serva, Sigma. In Rahmen dieser Arbeit genutzte, spezielle Chemikalien waren folgende:

Chemikalie	Bezugsquelle	Bestellnummer
Dynabeads® M-280 Streptavidin	Invitrogen	112.05D
poly (dAdT) DNA: Poly(deoxyadenylic-thymidylic) acid sodium salt	Sigma-Aldrich	P0883-10UN
Ni-NTA Agarose	Qiagen	1018240
Amylose Resin	New England Biolabs	E8021L
Hefeminimalmedium: Minimal SD Base	Clontech	630411
Hefe-Dropout: -Leu DO Supplement	Clontech	630414
Hefe-Dropout: -Trp DO Supplement	Clontech	630413
Hefe-Dropout: -Leu/-Tro DO Supplement	Clontech	630417
Hefe-Dropout: -Leu/-Trp/-His DO Supplement	Clontech	630419
Hefe-Dropout: -Ade/-His/-Leu/-Trp DO Supplement	Clontech	630428
Adenin: Adenine hemisulfate salt, >=99% cell culture tested	Sigma	A3159-5G
X- $\alpha$ -Gal: X-alpha Gal	Clontech	630407
5-Bromo-4-chloro-3-indolyl- $\alpha$ -D-galactopyranosid	Glycosynth	70039
HL5 Medium: ForMedium™ HL5 Medium	ForMedium	HLB0102
AEBSF hydrochloride	Molekula	M41447139
Aprotinin	Roth	A162.1
1,4-Dithiothreit (DTT)	Roth	6908.1
Leupeptin Hemisulfat	Roth	CN33.3

## 2.3 Lösungen und Puffer

50x Phosphatpuffer:	Na <sub>2</sub> PO <sub>4</sub> 100 mM; KH <sub>2</sub> PO <sub>4</sub> 735 mM
Lysepuffer (Kernextrakte):	30 mM Hepes-NaOH, pH 8; 10 mM KCl; 10 mM Mg-Acetat; 10% Saccharose; 2% Triton X-100; 1 mM EDTA; 1 mM DTT
Grundpuffer (GP):	20 mM Hepes-NaOH, pH 8; 1 mM EDTA; 1 mM DTT; 15% (v/v) Glycerin; 15-600 mM KCl
20x SSC:	3 M NaCl; 300 mM Na-Citrat
Waschpuffer (Microarray):	1,2 M Phosphatpuffer, pH 6,8; 2 mM EDTA; 50% Formamid 1% Na-Laurylsarcosinat; 0,2% SDS; in 2x SSC
100x Denhardts Reagenz:	2% Ficoll; 2% Polyvinylpyrrolidon; 2% BSA
Hybridisierungspuffer:	4x Denhardts Reagenz; 2 µM oligo (dA) <sub>18</sub> ; 500 µg/ml Fisch Sperma DNA; in Waschpuffer
Columnpuffer (pMAL):	20 mM Tris-HCl; 200 mM NaCl; 1 mM EDTA; pH7,4
Elutionspuffer MBP (pMAL):	20 mM Tris-HCl; 200 mM NaCl; 1 mM EDTA; 10 mM Maltose; pH7,4
Lysispuffer (pET):	50 mM NaH <sub>2</sub> PO <sub>4</sub> ; 300 mM NaCl; 10 mM Imidazol; pH 8,0
Waschpuffer (pET):	50 mM NaH <sub>2</sub> PO <sub>4</sub> ; 300 mM NaCl; 20 mM Imidazol; pH 8,0
Elutionspuffer (pET):	50 mM NaH <sub>2</sub> PO <sub>4</sub> ; 300 mM NaCl; 250 mM Imidazol; pH 8,0
10x Faktor Xa-Puffer:	200 mM Tris-HCl; 1 M NaCl; 10 mM CaCl <sub>2</sub> ; pH 7,8
10x Imidazol:	250 mM Imidazol; in 1x Xa-Puffer
Proteaseinhibitoren:	1.000x AEBSF (100 mM → 100 µM) in MQ-Wasser 200x Aprotinin (1 mg/ml → 5 µg/ml) in MQ-Wasser 1.000x Leupeptin (5 mg/ml → 5 µg/ml) in MQ-Wasser 1.000x DTT (1 M → 1 mM) in MQ-Wasser
Lösung A:	0,5 M Ethylenglykol; 10 mM Bicin; 1 M Sorbitol
Lösung B:	0,4 M Polyethylenglykol 1.000; 0,2 M Bicin

Lösung C:	150 mM NaCl; 10 mM Bicin
Lysispuffer (Hefe):	10 mM Tris-HCl (pH8,0); 1 mM EDTA; 300 mM NaCl; 1% SDS; 2% Triton X-100
10x TBS:	200 mM Tris; 1,5 M NaCl; pH 7,4
TBS-T:	1x TBS; 0,05% TWEEN-20

## 2.4 Verwendete Medien und Antibiotika

Luria Broth-Medium (LB):	10 g/l Bacto-Trypton; 5 g/l Bacto-Hefeextrakt; 10 g/l NaCl; (15 g/l Agar); pH 7,0
YPDA:	10 g/l Hefeextrakt; 20 g/l Bacto-Pepton; 20 g/l Glukose; (18 g/l Agar); pH 6,5; nach Autoklavieren 15 ml/l 0,2% (w/v) Adeninhemisulfat
SD:	26,7 g/l SD-Base; entsprechendes Dropout; (18 g/l Agar); pH 5,8
HL5:	5 g/l Proteose-Pepton; 5 g/l Thioton E Pepton; 5 g/l Hefeextrakt; 0.35 g Na <sub>2</sub> HPO <sub>4</sub> ·7H <sub>2</sub> O; 0.35 g KH <sub>2</sub> PO <sub>4</sub> ; pH 6,4 bis 6,7; nach Autoklavieren 25 ml/l 40% (w/v) Glukose
HL5 (ForMedium™):	22 g/l ForMedium™ HL5 Medium; nach Autoklavieren 25 ml/l 40% (w/v) Glukose
SM/5:	2 g/l Glukose; 2 g/l Bacto-Pepton; 0.2 g/l Hefeextrakt 0.1 g/l MgSO <sub>4</sub> ; 1.9 g/l KH <sub>2</sub> PO <sub>4</sub> ; 1.0 g/l K <sub>2</sub> HPO <sub>4</sub> ; 15 g/l Agar; pH 6,5

**Tabelle 1: Verwendete Antibiotika**

Antibiotikum	Einsatz	Stocklösung	Endkonzentration
Ampicillin	Bakterienkultur	100 mg/ml in 50% Ethanol	100 µg/ml
Kanamycin	Bakterienkultur Hefemedien	50 mg/ml in MQ-Wasser	30-50 µg/ml 15 µg/ml
Gentamycin	Bakterienkultur	20 mg/ml in MQ-Wasser	20 µg/ml
Streptomycin	Bakterienkultur	20 mg/ml in MQ-Wasser	10 µg/ml
Blasticidin	<i>Dictyostelium</i> -Kultur	4 mg/ml in MQ-Wasser	4 µg/ml
G418	<i>Dictyostelium</i> -Kultur	4 mg/ml in MQ-Wasser	5-10 µg/ml

## 2.5 Benutzte Antikörper

7F3: Gegen die Carboxy-terminale Domäne von CbfA (Aminosäuren 795 bis 998) gerichteter, monoklonaler, primärer Antikörper aus Hybridomazellen (Winckler *et al.* 2001). Anwendung im *Western-Blot* 1:1 in 3% (w/v) BSA in TBS-T.

3H7: Gegen die Amino-terminale Domäne von CbfA (Aminosäuren 78 bis 355) gerichteter, monoklonaler, primärer Antikörper aus Hybridomazellen. Anwendung im *Western-Blot* 1:1 in 3% (w/v) BSA in TBS-T.

Anti-Maus IgG, HRP gekoppelt: Sekundärer, in Pferden entwickelter Antikörper zur Detektion von murinen primären Antikörpern. Katalognummer 7076 (Cell Signaling), über New England Biolabs bezogen. Anwendung im *Western-Blot* 1:2.000 in 5% (w/v) Milchpulver in TBS-T.

## 2.6 Verwendete Mikroorganismen

**Tabelle 2: Verwendete *Escherichia coli* und *Klebsiella planticola*-Stämme**

Stamm	Beschreibung	Bezugsquelle
DH5 $\alpha$ <sup>TM</sup>	F <sup>-</sup> endA1 glnV44 thi-1 recA1 relA1 gyrA96 deo <sup>R</sup> nupG $\Phi$ 80dlacZ $\Delta$ M15 $\Delta$ (lacZYA-argF)U169, hsdR17(r <sub>K</sub> <sup>-</sup> )	Life Technologies <sup>TM</sup>
BL21 (DE3)	F <sup>-</sup> ompT hsdS(r <sub>B</sub> <sup>-</sup> , m <sub>B</sub> <sup>-</sup> ) dcm <sup>+</sup> Tet <sup>r</sup> gal endA (DE3)	Novagen
BL21 (DE3) RIL	F <sup>-</sup> ompT hsdS(r <sub>B</sub> <sup>-</sup> , m <sub>B</sub> <sup>-</sup> ) dcm <sup>+</sup> Tet <sup>r</sup> gal endA Hte[argU ileY leuW Cam <sup>r</sup> ] (DE3)	Novagen
BL21 (DE3) ROSETTA <sup>TM</sup>	F <sup>-</sup> ompT hsdS(r <sub>B</sub> <sup>-</sup> , m <sub>B</sub> <sup>-</sup> ) dcm <sup>+</sup> Tet <sup>r</sup> gal endA pRARE [Cam <sup>r</sup> ] (DE3)	Novagen
ArcticExpress <sup>TM</sup> (DE3)	F <sup>-</sup> ompT hsdS(r <sub>B</sub> <sup>-</sup> , m <sub>B</sub> <sup>-</sup> ) dcm <sup>+</sup> Tet <sup>r</sup> gal $\lambda$ (DE3) endA Hte [cpn10 cpn60 Gent <sup>r</sup> ]	Stratagene
ArcticExpress <sup>TM</sup> (DE3) RIL	F <sup>-</sup> ompT hsdS(r <sub>B</sub> <sup>-</sup> , m <sub>B</sub> <sup>-</sup> ) dcm <sup>+</sup> Tet <sup>r</sup> gal $\lambda$ (DE3) endA Hte [cpn10 cpn60 Gent <sup>r</sup> ] [argU ileY leuW Str <sup>r</sup> ]	Stratagene
RZ1032	F <sup>+</sup> , cat, (pCJ105; m13s, CMR) dut, ung1, thi-1, relA1, spoT1, mcrA (Sambrock <i>et al.</i> 1989)	
<i>Klebsiella planticola</i>		<i>Dictyostelium</i> Stock Center

**Tabelle 3: Verwendete *Saccharomyces cerevisiae*-Stämme**

Stamm	Beschreibung	Bezugsquelle
AH109	MAT $\alpha$ , trp1-901, leu2-3, 112, ura3-52, his3-200, gal4 $\Delta$ , gal80 $\Delta$ , LYS2 : : GAL1UAS-GAL1TATA-HIS3, GAL2UAS-GAL2TATA-ADE2, URA3 : : MEL1UAS-MEL1TATA-lacZ, MEL1	Clontech
Y187	MAT $\alpha$ , ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4 $\Delta$ , met <sup>-</sup> , gal80 $\Delta$ , URA3 : : GAL1UAS-GAL1TATA-lacZ, MEL1	Clontech
KFY1	MAT $\alpha$ ura3-52 his3-200 lys2-801 trp1-901 leu2-3 112 gal4 $\Delta$ gal80 $\Delta$ cyh <sup>r</sup> 2 LYS2::GAL1 <sub>UAS</sub> -HIS3 <sub>TATA</sub> -HIS3 URA3::GAL1UAS-GAL1TATA-lacZ	F. Hänel (HKI Jena)

**Tabelle 4: Verwendete *Dictyostelium discoideum*-Stämme**

Stamm	Beschreibung	Bezugsquelle
AX2	Durch natürliche Selektion über die Zwischenschritte DdB (erstmalig in Brackenbury <i>et al.</i> 1974 erwähnt) und AX1 (Sussman und Sussman 1967) aus dem Isolat NC4 (Raper 1935) entstandene, axenisch wachsender Laborstamm (Watts und Ashworth 1970). <i>axeA</i> , <i>axeB</i> , <i>axeC</i>	<i>Dictyostelium</i> Stock Center
JH.D	Aus AX2 hervorgegangene Amber-Suppressionsmutante des CbfA-codierenden Gens (Winckler <i>et al.</i> 2001). <i>axeA</i> , <i>axeB</i> , <i>axeC</i> , pJH1.2(am)-tetO:Glu(su)	

## 2.7 Verwendete Kits und Enzymsysteme

- Qiaquick® Gel Extraction Kit (Qiagen)
- Qiaquick® PCR Purification Kit (Qiagen)
- DNeasy® Tissue Kit (Qiagen)
- RNeasy® Mini Kit (Qiagen)
- RNase-Free DNase Set (Qiagen)
- Oligotex® Kit (Qiagen)
- Omniscript® Reverse Transcription Kit (Qiagen)
- Brilliant® II SYBR® Green QRT-PCR Master Mix Kit, 2-Step (Stratagene)
- PCR Extender System, vormals Triple Master Mix (5 Prime)
- Taq DNA Polymerase (5 Prime)
- Matchmaker™ Library Construction & Screening System (Clontech)
- RNA 6000 Nano Chip (Agilent)
- SpotReport-10 Array Validation System (Agilent)
- Cy™3 und Cy™5 (Amersham Biosciences)
- pGEM®-T Vector System I (Promega)
- Xa Removal Resin (Qiagen)

- Restriktionsendonukleasen, 4-100 U/μl (New England Biolabs)
- Antarktische Phosphatase, 5 U/μl (New England Biolabs)
- T4 DNA Ligase, 400 U/μl (Fermentas und Invitrogen)
- T4 Polynukleotidkinase, 10 U/μl (New England Biolabs)
- T7 DNA Polymerase, unmodified, 10 U/μl (New England Biolabs)
- Proteinase K, 900 U/μl (Fermentas)
- Faktor Xa Protease, 1 U/μl (New England Biolabs)
- Lysozym, 20.000 U/mg (Roth)

## 2.8 Benutzte Software

- **ScanArray® Express 3.0** (PerkinElmer Life Sciences)
- **Array Tools**  
<http://www.uni-koeln.de/med-fak/biochemie/transcriptomics/toolsarray.e.shtml>
- **R 1.6.2** und **R 1.9.1** über BioConductor, <http://www.bioconductor.org/>
- **Significance Analysis of Microarrays 1.21** (SAM, Tusher *et al.* 2001),  
<http://www-stat-class.stanford.edu/~tibs/clickwrap/sam.html>
- **GeneSpring 7.2** (Agilent Technologies), <http://www.chem.agilent.com>
- **Gene Ontology Analysis Tool** (GOAT) für R (Xu und Shaulsky 2005) über BioConductor,  
<http://www.bioconductor.org/>
- **Gene Expression Omnibus**, <http://www.ncbi.nlm.nih.gov/geo/>
- **Gene Ontology Project** (GO), <http://www.geneontology.org/>
- **Dictybase** (Chisholm *et al.* 2006) , <http://dictybase.org/>
- **dictyExpress** (Rot *et al.* 2009), <http://www.ailab.si/dictyexpress/run/>
- **Protparam** (Gasteiger *et al.* 2005) über Expasy,  
<http://www.expasy.org/tools/protparam.html>
- **PSIPRED** (Jones 1999, Bryson *et al.* 2005), <http://bioinf.cs.ucl.ac.uk/psipred/>
- **ClustalX 2.0.11** (Thomson *et al.* 1994, Larkin *et al.* 2007),  
<http://www.clustal.org/download/>
- **BoxShade 3.21**, [http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)
- **TREE-PUZZLE 4.0** (Schmidt *et al.* 2002), <http://www.tree-puzzle.de/>
- **PHYLIP 3.67** (Felsenstein 1989), <http://evolution.genetics.washington.edu/phylip.html>
- **Microsoft® Office Excel®** 2003 und 2007 (Microsoft Corporation)
- **pDRAW 1.1**, <http://www.acaclone.com/>
- **MxPro™ QPCR Software** (Stratagene)
- **QCapture Pro 5.0.1.26** (QImaging)
- **ImageJ 1.42** (<http://rsbweb.nih.gov/ij/>)

## 2.9 Allgemeine molekularbiologische Methoden

Die weit verbreiteten Standardmethoden zur molekularbiologischen Klonierung wurden nach Angaben der Hersteller der verwendeten Enzyme durchgeführt. Einen generellen Überblick bezüglich der Techniken geben Sambrook *et al.* 1989 und Mülhardt 2009.

## 2.10 Kultivierung von *Dictyostelium discoideum* und anderen *Dictyosteliales*

Die verwendeten *D. discoideum*-Stämme AX2 und JH.D wurden in dem axenischen HL5-Medium angezüchtet. Frisch aufgenommene JH.D-Kulturen wurden eine Passage lang mit 4 µg/ml Blasticidin auf das Vorhanden sein des *cbfA*-Suppressionskonstruktes überprüft. Transformanten im JH.D-Hintergrund sind mit 5, im AX2-Hintergrund mit 10 µg/ml G418 selektioniert worden. Zur Transformation sind verschiedene Protokolle verwendet worden, ein Überblick dies bezüglich ist in Gaudet *et al.* 2007 gegeben. Es wurde sowohl die Kultivierung als adhärenzte Zellschicht in Petrischalen als auch in Schüttelkultur durchgeführt. Die Inkubation erfolgte bei 22°C, die Schüttelkultur wurde mit 150 rpm bewegt. Die in Petrischalen wachsenden Kulturen sind bei einer Konfluenz von 70-90% geteilt worden, die Schüttelkultur wurde mit Zelldichten zwischen  $2 \cdot 10^5$  bis  $5 \cdot 10^6$  Zellen/ml benutzt. Für weiterführende Arbeiten wurden die Zellen bei 100 x g für 5 min bei 4°C pelletiert, das Pellet in 1x Phosphatpuffer resuspendiert und zu  $2 \times 10^7$  Zellen in kleine Reaktionsgefäße überführt. Nach einer einminütigen Zentrifugation bei 10.000 x g wurde der Überstand verworfen und die restliche Flüssigkeit mit einer Kanüle abgesaugt. Eine Lagerung der Pellets erfolgte sofort bei -80°C. Andere Vertreter der *Dictyosteliales* sind auf SM/5-Agarplatten in Verbindung mit *Klebsiella planticola* angezüchtet worden. Vor dem Einsetzen der multizellulären Entwicklung wurden die Zellen durch Zentrifugation von den assoziierten Bakterien getrennt und zu je  $2 \times 10^7$  Zellen bei -80°C eingefroren.

## 2.11 Isolation und Synthese von gDNA, RNA, mRNA und cDNA

Zur Isolation der Gesamt-RNA oder der genomischen DNA wurde ein Zellpellet von  $2 \times 10^7$  Zellen verwendet, welches mindestens einen Tag bei -80°C gelagert worden ist. Die Gewinnung der gDNA erfolgte mit dem **DNeasy® Tissue Kit** (Qiagen). Zur RNA-Isolation wurde das **RNeasy® Mini Kit** (Qiagen) unter Verwendung des *Animal Cell*-Protokolls genutzt. Zur Aufreinigung von mRNA



ist das **Oligotex<sup>®</sup> Kit** (Qiagen) mit dem *mRNA Spin-Column* Protokoll angewendet worden. Es dienten 100 bis 200 µg RNA als Ausgangsmaterial. Die Synthese der cDNA erfolgte aus 2µl Gesamt-RNA der Konzentration 250 ng/µl unter Verwendung des **Omniscrypt<sup>®</sup> Reverse Transcription Kits** (Qiagen) sowie eines oligo (dT)<sub>18</sub>-Primers.

### 2.12 Erstellung von Kernextrakten

Mindestens  $1 \times 10^9$  Zellen wurden 5 min bei 100 x g und 4°C pelletiert. Die Zellen sind anschließend zweimal in 200 bzw. 50 ml kalten 1x Phosphatpuffer resuspendiert und erneut pelletiert worden. Zur Lyse der Zellen wurde der Ansatz zehn Minuten in Lysepuffer plus 60 µg/ml Leupeptin auf Eis inkubiert. Dabei sind 7,5 ml Puffer pro g Zellpellet genutzt worden. Anschließend erfolgte eine Homogenisierung der Zellsuspension in eine Glashomogenisator, dreimal mit dem Stößel A (*loose*) und fünfmal mit dem Stößel B (*tight*). Zwischen den unterschiedlichen Stößeln wurde die Probe kurz auf Eis inkubiert. Die Kerne wurden bei 10.000 x g und 4°C für 20 min sedimentiert. Der Überstand ist mit einer Wasserstrahlpumpe entfernt worden. Das Kernpellet wurde nun in 2 ml Lysepuffer plus 60 µg/ml Leupeptin resuspendiert und drei weitere Male mit Stößel B homogenisiert. Die anschließende Pelletierung erfolgte unter den genannten Parametern für zehn Minuten. Das Kernsediment wurde nach folgenden Schema mit jeweils 2 ml des Grundpuffers plus 20 µg/ml Leupeptin mit unterschiedlichen KCl-Mengen behandelt: 1x GP15 (Grundpuffer plus 15 mM KCl), 3x GP50, 3x GP600. Die Zentrifugationen wurden jeweils mit 10.000 x g, 4°C für zehn Minuten durchgeführt. Durch die Anwendung von 600 mM KCl kam es zur Lyse des Zellkernes. Diese Waschfraktionen enthielten die Kernproteine und wurden gesammelt. Zur weiteren Verwendung wurde die GP600-Fraktion zweimal gegen mindestens 20 Volumen GP15 bei 4°C dialysiert und bei -80°C gelagert.

### 2.13 Quantitative *Real Time*-PCR (qRT-PCR)

Für die vergleichende Expressionsanalyse weniger Gene wurde die quantitative Echtzeit-Polymerase-Kettenreaktion an einer komplementären DNA (qRT-PCR) durchgeführt. Die Expression ist gegen das Gen, welches die GAPDH codiert (*gpdA*, DDB\_G0275153), standardisiert worden. Die verwendeten Primer QGAPDH-01 und 02 überspannen ein Intron, so dass Kontaminationen mit gDNA durch eine konventionelle RT-PCR ausgeschlossen werden

konnten. Die Durchführung der Echtzeitquantifizierung erfolgte auf einem **Mx3000P** (Stratagene) unter der Verwendung des **Brilliant® SYBR® Green QPCR Master Mix** (Stratagene). Pro Ansatz wurden ein µl einer 1 zu 10 Verdünnung der unter 2.11 beschriebenen cDNA zur Quantifizierung benutzt. Eine Liste der verwendeten Primer befindet sich im Anhang A. Einen zehnminütigen initialen Denaturierungsschritt bei 95°C schlossen sich 40 Zyklen von 95°C für 30 sek, 58°C für 30 sek, 72°C für 30 sek an. Anschließend wurde die Schmelzkurve der entstandenen Produkte und die Regulation nach der Methode von Pfaffl (2001) bestimmt. Alle Messungen sind an mindestens drei unabhängigen RNA-Präparationen durchgeführt worden. Die Berechnung der Standardabweichung erfolgte über die Excel®-Funktion „STABWN“, die des Signifikanzwertes p über „TTEST“.

### 2.14 Microarray-Methoden

Die Arbeiten betreffend der Microarray-Experimente wurden in der Arbeitsgruppe von Privatdozent Dr. Eichinger in Köln unter der Nutzung der dort vorhandenen Ressourcen durchgeführt. Der benutzte cDNA-Microarray „*Dictyostelium discoideum* cDNA Microarray 6000“ umfasst 5423 nicht-redundante ESTs des *Dictyostelium* cDNA Projektes aus der vegetativen Wachstumsphase (Urushihara *et al.* 2004), 450 selektierte Gene sowie Positiv- und Negativkontrollen, unter anderem das **SpotReport-10 Array Validation System** von Agilent (beschrieben in Farbrother *et al.* 2006). Alle Proben untersuchter Gene sind zweifach gespottet, Kontrollen mehrfach. Der Array besteht aus 14620 Spots. Eine genaue Beschreibung des *Microarrays* ist unter der *Accession Number* **GPL1972** im **Gene Expression Omnibus** (<http://www.ncbi.nlm.nih.gov/geo/>) zu finden.

Die Qualität der zu untersuchenden RNA ist mit Hilfe des RNA 6000 Nano Chips (Agilent) überprüft worden. Für jedes Experiment wurden aus 20 µg Gesamt-RNA, im Beisein des **SpotReport-10 Array Validation System** (Agilent) unter Nutzung des **FairPlay™ Microarray Labeling Kits** (Stratagene), cDNA synthetisiert. Durch die Nutzung von Aminoallyl-dUTP während der Synthese konnte die cDNA mit aktivierten **Cy™3** bzw. **Cy™5** (Amersham Bioscience) markiert werden. Für jede RNA-Kombination wurden *Dye Swaps* durchgeführt. Das präzipitierte Gemisch ist in 65 µl Hybridisierungspuffer aufgenommen worden (Noegel *et al.* 1985).

Der Ansatz wurde zehn Minuten lang auf 80°C erhitzt und anschließend auf einen mit Deckglas versehenen *Microarray* appliziert. Der Array ist 16 Stunden in einer Hybridisationskammer bei 37°C inkubiert worden. Anschließend wurde wie folgt gewaschen: je einmal 2x SSC + 0,1% SDS für 10 sek, 2x SSC + 0,1% SDS für 5 min, 0,1x SSC + 0,1% SDS für 5 min, fünfmal 0,1x SSC für 5 sek, 0,01x SSC für 5 sek. Die Trocknung des Arrays erfolgte durch Zentrifugation (fünf Minuten bei 235 x g). Fluoreszenzsignale sind mit einem **ScanArray® 4000XL** (Perkin Elmer Life Sciences) gemessen worden. Um dabei möglichst viele Proben erfassen zu können, wurde mit zwei Lasereinstellungen gescannt. Die Quantifizierung der Signalstärke erfolgte mit dem Programm **ScanArray® Express 3.0** (Perkin Elmer Life Sciences). Dieses Programm diente ferner der manuellen Überprüfung der erhaltenen Werte. *Spots* mit schlechter Qualität wurden als „*bad*“ oder „*not existing*“ markiert und so aus der Analyse ausgeschlossen.

**Array Tools** (<http://www.uni-koeln.de/med-fak/biochemie/transcriptomics/toolsarray.e.shtml>) wurde genutzt, um die Primärdaten in das jeweils benötigte Format für weitere Analysen umzuwandeln. Die Normalisierung der Daten eines Experimentkomplexes erfolgte mit dem Programm **R 1.6.2** (BioConductor), dabei ist die LOWESS-Methode (Yang *et al.* 2002) angewendet worden. Differentiell exprimierte Gene wurden mit dem Programm **Significance analysis of microarrays** (SAM, Tusher *et al.* 2001) identifiziert. Dieses verwendet einen erweiterten t-Test, welcher die erwartete und die tatsächliche Verteilung der Daten vergleicht. Die Clusteranalyse ist mit **GeneSpring 7.2** (Agilent Technologies) erstellt worden. Anreicherungsstudien bezüglich der Genontologie sind mit dem Programmpaket **Gene Ontology Analysis Tool** (GOAT) für **R 1.9.1** (über BioConductor, <http://www.bioconductor.org/>, Xu und Shaulsky 2005) durchgeführt worden. Dabei werden statistisch signifikante Anreicherungen und das Maß der Signifikanz innerhalb einer Liste von Genen im Bezug auf alle annotierten Gene errechnet.

### 2.15 Phylogenetische Analysen

Die untersuchten *Dictyosteliales* wurden über das *Dictyostelium* Stock Center bezogen. Folgende Stämme fanden Verwendung: *Dictyostelium giganteum* WS589 (Stamm ID: DBS0235820), *D. sphaerocephalum* GR11 (DBS0235889), *Polysphonylium pallidum* WS320

(DBS0236811), *Acytostelium subglobosum* Lb1 (DBS0235452), *D. fasciculatum* SH3 (DBS0235810), *D. purpureum* WS278 (DBS0235880). Die Reinigung der Genomische DNA und Synthese der cDNA erfolgten wie unter 2.11 beschrieben. Degenerierte Primer wurden von *D. discoideum*-CbfA und im Verlauf der Experimente von multiplen Alignments abgeleitet. Die DNA-Sequenz orthologer *cbfA*-Gene wurde aus überlappenden PCR-Fragmenten generiert und manuell angeordnet. Die Klonierung schwer zu amplifizierende PCR-Produkte erfolgte in das Vektorsystem pGEM<sup>®</sup>-T (Promega).

Die Nukleotidsequenz der *cbfA* und *cbfB* Gene von *D. fasciculatum* und *P. pallidum* wurden aus unveröffentlichten Daten vergleichender Genomsequenzierungen der Arbeitsgruppe von Dr. Glöckner am Leibnitz-Institut für Altersforschung in Jena assembliert. Die entsprechenden Sequenzen von *A. subglobosum* wurden von Urushihara und Mitarbeitern (Universität Tsukuba, Japan) zur Verfügung gestellt, die von *D. purpureum* *cbfA* und *cbfB* von Kuspa und Mitarbeitern (Baylor College of Medicine, Houston, USA).

Multiple Sequenzalignments sind mit dem Programm **ClustalX 2.0.11** (Larkin *et al.* 2007, <http://www.clustal.org/download/>, nach der Methode von Thomson *et al.* 1994) erstellt worden. Die phylogenetischen Bäume wurden mit dem Programm **TREE-PUZZLE 4.0** (Schmidt *et al.* 2002) unter Verwendung des JTT-Modells generiert. Die Verwendung des Programmpaketes **PHYLIP 3.67** (Felsenstein 1989) diente zur Überprüfung der Bäume und ergab im Wesentlichen gleiche Ergebnisse.

### **2.16 Aufreinigung rekombinanter CbfA-CTD im analytischen Maßstab**

Jeder zu untersuchende *E. coli*-Expressionsstamm wurde mit dem Plasmid pET33b(+)-CTD bzw. pMAL-c2x-CTD durch Standardmethoden transformiert. Durch 30 µg/ml Kanamycin (pET) bzw. 100 µg/ml Ampicillin (pMAL) selektierte Klone sind in Schüttelkultur vermehrt und anschließend konserviert worden. Aus einer ausgewachsenen Vorkultur wurden 50 ml Hauptkultur 2%ig inokuliert. Nach zwei Stunden Inkubation bei 37°C und 200 rpm erfolgte die Bestimmung der optische Dichte bei 600 nm (OD<sub>600</sub>). Kulturen, welche nicht bei 37°C induziert werden sollten, wurden 10 Minuten auf die zu untersuchende Inkubationstemperatur abgekühlt. Die Induktion erfolgte durch Zugabe von einem mM IPTG (Endkonzentration). Die Kultur wurde bis zum

Erreichen der stationären Phase inkubiert, anschließend pelletiert (9.000 x g, 4°C, 5 Minuten) und das Feuchtgewicht der Bakterien bestimmt. Die Pellets wurden bei -80°C gelagert.

Die Lyse der Zellen erfolgte durch Aufnahme des Pellets in 5 ml pro Gramm Feuchtgewicht Columnpuffer (pMAL) bzw. Lysepuffer (pET) mit Proteaseinhibitoren und 1 mg/ml Lysozym. Es wurde 30 min auf Eis inkubiert und anschließend dreimal 30 sek sonifiziert. Die lösliche Fraktion konnte durch Zentrifugation für 30 Minuten bei 4°C und 9.000 x g gewonnen werden. Nach der Überprüfung des pH-Wertes erfolgte die Aufreinigung durch die Bindung an  $\frac{1}{5}$  Volumen **Amylose-resin** (New England Biolabs, pMAL) bzw. **Ni-NTA-slurry** (Qiagen, pET). Anschließend wurde zweimal mit einem Volumen Column- (pMAL) bzw. Waschpuffer (pET) gewaschen. Die Elution erfolgte durch zehninütige Inkubation mit  $\frac{1}{2}$  Volumen Elutionspuffer MBP (pMAL) bzw. Elutionspuffer (pET). Der Überstand ist durch dreiminütige Zentrifugation mit 3.000 x g abgetrennt und entfernt worden. Die Elution wurde zweimal durchgeführt.

### **2.17 Aufreinigung rekombinanter CbfA-CTD im präparativen Maßstab**

Für alle Aufreinigungen im großem Maßstab wurde der Klon *E. coli* BL21 (DE3) pET33b(+)-CTD-Xa #4 genutzt. Die Anzucht erfolgte in allen Schritten in LB-Medium mit 30 µg/ml Kanamycin. Aus einer ausgewachsenen 100 ml Vorkultur wurden je 250 ml Hauptkultur 2%ig inokuliert und in acht Kolben mit je einen Liter Volumen bei 37°C und 200 rpm kultiviert. Nach dem Erreichen einer optischen Dichte bei 600 nm ( $OD_{600}$ ) von 0,6 nach etwa 100 Minuten wurde die CTD-Expression durch Zugabe von 1 mM IPTG (Endkonzentration) induziert. Die Kultivierung ist bis zum Erreichen der stationären Phase fortgesetzt worden. Anschließend wurden die Bakterien bei 9.000 x g, 4°C fünf Minuten pelletiert, ihr Feuchtgewicht bestimmt und das Pellet bei -80°C gelagert.

Die Zellyse erfolgte durch zweimalige Benutzung der *French Press* des Lehrstuhls für Angewandte und ökologische Mikrobiologie Jena. Dazu wurde das Pellet pro Gramm Feuchtgewicht in 4 ml Waschpuffer mit Proteaseinhibitoren und 10 µl DNase (3 U/µl) resuspendiert. Die Lyse der Zellen ist mikroskopisch überprüft worden. Der lösliche Anteil wurde durch Zentrifugation für 30 Minuten bei 4°C und 9.000 x g abgetrennt, dessen pH-Wert

bestimmt und gegebenenfalls mit 1 N Salzsäure auf acht eingestellt. Zur Entfernung der *E. coli*-Proteine sind pro Gramm Feuchtgewicht ein Milliliter Ni-NTA-*slurry* verwendet worden, welche 5 bis 20 mg/ml rekombinantes Protein bindet. Der Proteinüberstand und das Säulenmaterial wurden zusammen eine Stunde mit 15 rpm bei 4°C rotiert. Anschließend ist das Gemisch in eine Glassäule mit Fritte gegeben worden. Der Lauf durch die Säule erfolgte allein durch Schwerkraft. Das Säulenmaterial wurde zweimal mit drei Volumen Waschpuffer gewaschen. Die Elution erfolgte nach zehnmütiger Inkubation mit zwei Volumen Elutionspuffer. Dieser Schritt wurde wiederholt und die Elutionsfraktionen vereinigt.

Das Eluat ist über Nacht bei 4°C gegen mindestens 20 Volumen 1x Faktor Xa-Puffer dialysiert worden. Der Dialysepuffer wurde dabei viermal gewechselt. Das erhaltene Dialysat wurde für 30 sek mit 3.000 x g zentrifugiert und die Proteinkonzentration sowie das Volumen des Überstandes bestimmt. Die Lösung des rekombinanten Proteins wurde zu je einen Milliliter aliquotiert und pro mg Protein 2 µl Faktor Xa (1 U/µl) zugegeben. Es folgte eine Inkubation von 24 Stunden bei 4°C. Das Schneiden wurde durch Zugabe von AEBSF und DTT abgestoppt, der Faktor Xa anschließend durch Nutzung des **Xa Removal Resin** (Qiagen) nach Angaben des Herstellers entfernt. Dem Überstand nach der Zentrifugation für 3 Minuten bei 1.000 x g wurde ein fünftel Volumen 10x Imidazol zugesetzt (Endkonzentration Imidazol 40 mM). Das ungeschnittene Protein und die Affinitätsmarkierung sind durch Applizierung des Gemisches auf gleiche Volumen Ni-NTA-*slurry* wie bei der anfänglichen Aufreinigung entfernt worden. Der Überstand nach Zentrifugation enthielt das gereinigte, geschnittene Protein.

Alle Reinigungsschritte sind durch Überprüfung der Proteinkonzentrationen während der Aufreinigung und durch *SDS-PAGE* mit Coomassie-Färbung überprüft worden. Die Proteinkonzentration von Gemischen ist durch einen Bradford-Assay (Bradford 1976) bestimmt worden. Ab der ersten Elution konnte die Konzentration der hinreichend reinen Proben auch durch die Bestimmung der Extinktion bei 280 nm bestimmt werden (Pace *et al.* 1995). Entsprechende Extinktionskoeffizienten wurden durch das Programm **ProtParam** (<http://www.expasy.org/tools/protparam.html>) berechnet.

### 2.18 DNA-Protein-Interaktionsassay – CTD-Pulldown

Diese Methode orientiert sich an Siol *et al.* (2006a) und wurde entsprechend der Fragestellung modifiziert. Biotinylierte Proben sind durch eine PCR unter Beteiligung des Primers pGEM-for(biotin) erstellt worden. Das PCR-Produkt wurde über Agarosegelelektrophorese aufgetrennt und mit Hilfe des **Qiaquick® Gel Extraction Kits** (Qiagen) aufgereinigt. Die DNA-Konzentration wurde photometrisch bestimmt. Die nach Methode 2.12 erstellten Kernextrakte sind durch einen *Western-Blot* mit dem Antikörper 7F3, welcher gegen die CbfA-CTD gerichtet ist, auf ihren Gehalt an CbfA und GFP-CbfA<sup>724-998</sup> überprüft worden. Die Konzentration des bakteriellen CbfA<sup>793-1000</sup> wurde wie unter 2.17 beschrieben bestimmt.

Pro Ansatz wurden bis zu 400 µl Kernextrakt bzw. 10 oder 125 pmol bakteriell erstellte CTD (rCTD) genutzt. Die Proteingemische sind auf 430 µl Volumen mit GP50 aufgefüllt worden. Dem Protein bakteriellen Ursprungs wurde 1 mg/ml BSA zugefügt. Ferner wurden 0,5 (rCTD) bis 2,5 µg (Kernextrakt) poly (dAdT) DNA hinzugefügt und 15 min inkubiert. Anschließend wurde 5 pmol biotinyliertes PCR-Produkt in 70 µl GP50 zugegeben und eine Stunde bei Raumtemperatur mit 15 rpm rotiert. Die Aufreinigung des Komplexes erfolgte durch die Zugabe von 25 µl Streptavidin-gekoppelter **Dynabeads®** (Invitrogen). Nach 30-minütigem Rotieren wurden die *Beads* magnetisch aus der Lösung entfernt und fünfmal mit GP50 gewaschen. Der letzte Waschschrift erfolgte in einem neuen Reaktionsgefäß. Die Probe ist sofort mit 60 µl 2x Laemmli-Puffer versetzt und 5 min bei 95°C erhitzt worden. Je 15 µl wurden elektrophoretisch aufgetrennt, auf eine Membran geblottet und mit dem Antikörper 7F3 beprobt.

### 2.19 Kultivierung und Transformation von *S. cerevisiae*

Die Kultivierung von *S. cerevisiae* erfolgte nach den Standardmethoden wie unter anderem in Johnston (1994) beschrieben. Alle Wachstumsschritte wurden bei 30°C vollzogen und die Schüttelkulturen mit 150 rpm bewegt. Die Transformation im kleinen Maßstab erfolgte nach der in Gietz und Woods (2001) beschriebenen Methode.

Eine Kolonie eines maximal einen Monat alten Ausstriches eines *Cryostocks* auf YPDA-Platte wurde in 10 ml YPDA angeimpft und über Nacht bei 30°C geschüttelt. Am nächsten Tag dienten 5 ml als Inokulum für 100 ml YPDA. Nach Erreichen einer OD<sub>600</sub> von 0,6 wurden die Zellen zu

50 ml aufgeteilt und fünf Minuten bei 2.000 x g und 4°C sedimentiert. Das Pellet löste man anschließend durch Vortexen und resuspendierte es in 10 ml Lösung A. Durch zehnmütige Zentrifugation bei 2.000 x g und 4°C wurden die Zellen erneut sedimentiert und anschließend kurz getrocknet. Je Ansatz wurden die Zellen in 1 ml Lösung A durch Klopfen resuspendiert und nach Zugabe von 55 µl DMSO in Portionen von 200 µl in vorgekühlte Reaktionsgefäße überführt, welche daraufhin sofort bei -80°C eingefroren worden.

Zehn Minuten auf Eis angetaute Zellen wurden mit jeweils 2 µg zu transformierenden Plasmid gemischt. Anschließend wurde der Ansatz 5 Minuten bei 37°C und 700 rpm im Heizblock geschüttelt, mit einem Milliliter Lösung B befüllt und eine Stunde bei 30°C und 400 rpm inkubiert. Im Anschluss erfolgte ein Zentrifugation für fünf Minuten bei 2.000 x g, Resuspendierung in 1 ml Lösung C und eine erneute Zentrifugation. Nach Abgießen des Überstandes wurden die Zellen auf entsprechendes Hefeminimalmedium ausplattiert und drei Tage bei 30°C inkubiert.

### **2.20 Isolierung von Plasmid-DNA aus Hefe – Plasmidrescue**

5 ml Minimalmedium ohne Histidin, Adenin, Leucin und Tryptophan (SD minus HALT) wurden mit einer einzelnen Kolonie beimpft und für 36 Stunden bei 30°C und 200 rpm geschüttelt. Als nächstes überführte man dieser Kultur in drei 1,5 ml fassende, extradichte Reaktionsgefäße (*safe lock tubes*) und zentrifugierte eine Minute bei maximaler Geschwindigkeit. Die Zellpellets wurden anschließend in 200 µl Lysispuffer (Hefe) resuspendiert, mit 300 µl Glasperlen versehen und mit 200 µl Phenol / Chloroform gemischt. Nun wurden die Ansätze für zwei Minuten gevortext und anschließend fünf Minuten bei 4°C und maximaler Geschwindigkeit zentrifugiert. Die plasmidhaltigen Oberphasen sind vereinigt worden, wobei das Endvolumen bei etwa 600 µl lag. Durch Zugabe von  $\frac{1}{2}$  Volumen Chloroform / Isoamylalkohol, kurzes Vortexen und Zentrifugation wurde das Phenol entfernt. Die entstandene Oberphase von ca. 500 µl wurde in einem neuen Reaktionsgefäß mit 0,7 Volumen Isopropanol und 0,1 Vol 3 M Natriumacetat versetzt und fünfzehn Minuten bei 4°C und 10.000 x g zentrifugiert. Die pelletierte Plasmid-DNA wurde dreimal mit einem Milliliter 70%igen Ethanol gewaschen und jeweils drei Minuten bei 10.000 x g zentrifugiert. Das getrocknete Pellet wurde in 10 µl MQ-Wasser resuspendiert. 3 µl



dieser Suspension wurden durch Elektroporation in *Escherichia coli* DH5 $\alpha$ <sup>TM</sup> transformiert.  $\frac{1}{4}$  Volumen der transformierten Zellen plattierte man auf LB<sub>Amp</sub>-Platten zur Isolierung des Beuteplasmides aus, das restliche Volumen auf LB<sub>Kana</sub>-Platten zur Isolierung des Köderplasmides.

### 2.21 Konstruktion und Screening von Hefe-2-Hybrid-Banken

Zur Erstellung der Hybrid-Banken wurde das **Matchmaker<sup>TM</sup> Library Construction & Screening System** (Clontech) genutzt. Als Ausgangsmaterial dienten vegetativ wachsende *Dictyostelium discoideum*-Zellen des Stammes AX2. Dazu ist die mRNA wie unter 2.11 beschrieben aus den AX2-Pellets isoliert worden. 80 bis 100 ng mRNA wurden zur Synthese der cDNA genutzt. Dabei fanden die Protokolle F bis I des Abschnittes „Generating a cDNA Library“ der Anleitung PT-3955-1 (Invitrogen) Anwendung. Die Herstellung von kompetenten Hefezellen des Stammes Y187 geschah leicht modifiziert nach dem Protokoll „Preparation of Competent Yeast Cells—LiAc Method“. Die Transformation, *In-vivo*-Rekombination und Vermehrung der Bank erfolgte nach den Schritten 1 bis 3 des Protokolls A „Screen by Yeast Mating“.

Das Interaktionsmating vollzog sich nach den Angaben des Systemherstellers mit den jeweils in Tabelle 8 aufgeführten Modifikationen. Initial als positiv erkannte Interaktionen wurden unter Verwendung der in dem Anleitungsabschnitt „Analyzing Positive Interactions“ dargestellten Methoden untersucht.

## 3 Ergebnisse

### 3.1 Kurze Übersicht zum Stand der Forschung

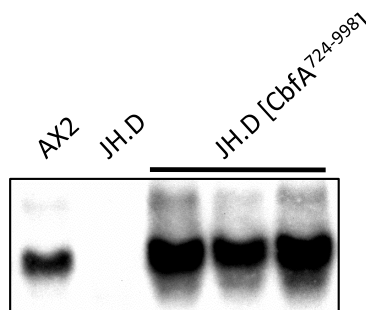
Der C-Modul-bindende Faktor A (CbfA) ist ein aus mehreren Domänen aufgebautes Protein von 1.000 Aminosäuren. Die native Molekülmasse beträgt etwa 120 kDa (Geier *et al.* 1996), die aus der Primärsequenz berechnete liegt bei 114.270 Da (Winckler *et al.* 2004). Das Genprodukt kann in allen Stadien des Lebenszykluses von *Dictyostelium discoideum* mit konstant bleibender Menge mit Hilfe eines spezifischen *Western-Blots* nachgewiesen werden. CbfA wurde über seine Bindung an den Promotor des für die Synthese von (-)-Strang-RNA des Retrotransposons TRE5-A.1 verantwortlichen C-Moduls (Schumann *et al.* 1994) erstmalig identifiziert (Geier *et al.* 1996). Es wurde aus Zellkernextrakten aufgereinigt, was eine Lokalisierung im *Nucleus* vermuten lässt. Das Protein besitzt am Amino-terminalen Ende (N-Terminus) eine Jumonji-Domäne des C-terminalen Types – JmjC – (Clissold und Ponting 2001), dessen genaue Funktionen noch unbestimmt ist. Diskutiert werden vor allem Chromatin-remodulierende Prozesse, wie sie bereits in anderen JmjC-haltigen Proteinen beschrieben sind (u.a. Tsukada *et al.* 2006). Möglicherweise dient die JmjC auch als Plattform für Protein-Protein-Interaktionen (Tronnersjö *et al.* 2006). An die Jumonji-Domäne von CbfA schließen sich zwei Zinkfinger an, welche für DNA (u.a. Elrod-Erickson *et al.* 1996), RNA (Hall 2005) oder Protein-Protein-Bindungen (Brayer und Segal 2008) verantwortlich sein können. Der folgende Abschnitt von CbfA besteht aus einer etwa 200 Aminosäuren langen Region, welche zu 50% aus Asparagin zusammengesetzt ist. Daran schließt sich die Carboxy-terminale Domäne (CTD) an. Diese besitzt ein AT-Haken-Motiv zur Bindung kompakter, dA und dT reicher Sequenzen (Arvind und Landsman 1998). Die Beteiligung des AT-Hakens an der Bindung von CbfA an das C-Modul ist nachgewiesen (Horn *et al.* 1999).

Studien an der CbfA-Suppressionsmutante JH.D, welche nur noch etwa 5% des Wildtyplevels an CbfA bildet, zeigen eine Reihe phänotypischer Veränderungen. Dies sind eine Erhöhung des Zellvolumens durch veränderte Pinozytoseraten bei gleichzeitig verminderter axenischer Wachstumsrate. Des Weiteren geringe Phagozytoseraten, die einhergehen mit geringerem

Wachstum auf Bakterien sowie eine fehlende multizelluläre Entwicklung als Antwort auf Nährstoffmangel (Winckler *et al.* 2001).

Eine Reihe von Entwicklungsmarkern werden in dem Stamm JH.D kaum bzw. gar nicht durch das Hungern der Zellen induziert (Winckler *et al.* 2004). Dabei kann dieser Effekt jeweilig durch die Zugabe von cyclischem Adenosinmonophosphat (cAMP), ektopisch exprimierten CbfA (DDB\_G0279409), der katalytischen Untereinheit der Proteinkinase A (*pkaC*, DDB\_G0283907) oder der Adenylylcyclase A (*acaA*, DDB\_G0281545) zumindest teilweise phänotypisch geheilt werden (Winckler *et al.* 2004, Siol *et al.* 2006a). Die Beeinträchtigung des Entwicklungszykluses in der Suppressionsmutante kann direkt auf die fehlende basale Aktivität des *acaA*-Gens zurückgeführt werden. Bisherige Daten lassen vermuten, dass CbfA als Transkriptionsfaktor an dessen Transkription direkt beteiligt ist (Siol *et al.* 2006a).

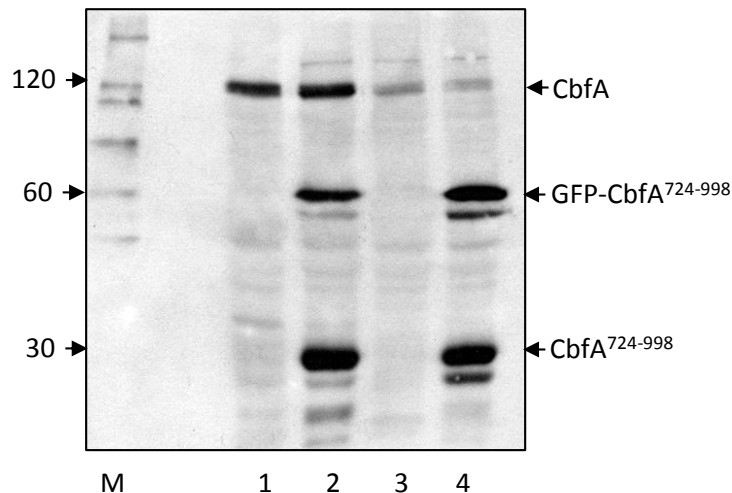
Des Weiteren zeigt JH.D eine stark verminderte Transkription beider TRE5-A-Stränge (Winckler *et al.* 2001, Beck *et al.* 2002). In diesen Zusammenhang ist zu erwähnen, dass die transkriptionelle Aktivität von CbfA in diesem Fall durch die Expression von CbfA<sup>724-998</sup> in JH.D vollständig wiederhergestellt werden kann. Mit Hilfe eines *Northern-blots* gegen die mRNA des Retrotransposon TRE5-A (Abbildung 10) kann gezeigt werden, dass der Verlust der Transkription des Transposons in JH.D durch CbfA<sup>724-998</sup> umgekehrt werden kann. Dies lässt eine eigenständige genregulatorische Funktion der CTD vermuten.



**Abbildung 10: Northern-Blot mit Sequenz von ORF1 als Sonde gegen TRE5-A; JH.D [CbfA<sup>724-998</sup>] sind 3 CTD-exprimierende JH.D-Transformanten**

### 3.2 Transkriptionelle Reaktion auf die *cbfA*-Suppression und die Überexpression von GFP-CbfA<sup>724-998</sup>

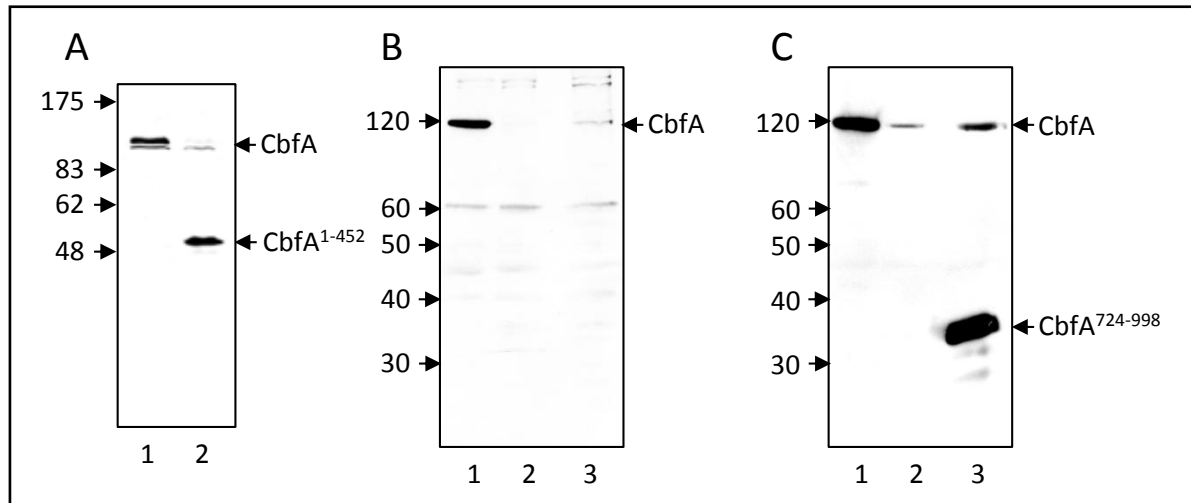
Die im vorangegangenen Abschnitt zusammengefassten Daten über den Stamm JH.D lassen zahlreiche Funktionen von CbfA als Transkriptionsfaktor sowohl in vegetativ wachsenden Zellen als auch in der Steuerung des Entwicklungsprozesses vermuten. Um die durch den beobachteten Phänotyp der CbfA-Mutante repräsentierten Änderungen zu beleuchten, sind Transkriptomanalysen durchgeführt worden. Da aufgrund der in Abbildung 10 dargestellten Beobachtungen in Zusammenhang mit der Retrotransposition des TRE5-A-Elementes von einer eigenständigen Funktion der CTD auszugehen ist, sind in Analysen des gesamten CbfA-Proteins auch Untersuchungen der Carboxy-terminalen Domäne mit einbezogen worden. Zunächst wurde die codierende Sequenz für die Aminosäuren 724-998 des CbfA mit Hilfe der Primer attrBcmbf1 und 2 aus genomischer DNA des Stammes AX2 amplifiziert und in den Vektor pDonor kloniert. Das entstandene Plasmid pEntr Cmbf1-2 fungierte als Basis für eine Rekombinationsreaktion in den Expressionsvektor pPT132. Dieser Vektor codiert neben den für Expressionsvektoren typischen Sequenzen für eine Amino-terminale Fusion des Grün fluoreszierenden Proteins (GFP) mit dem rekombinierten Genabschnitt. Der Expressionsvektor pPT132 Cmbf1-2, welcher für das Fusionsprotein GFP-CbfA<sup>724-998</sup> codiert, wurde in die Stämme AX2 und JH.D transformiert.



**Abbildung 11: Western-Blot mit Antikörper 7F3 von Gesamtzelllysaten;** M = Magic Mark (Invitrogen); AX2 (Spur 1), AX2 [GFP-CbfA<sup>724-998</sup>] (Spur2), JH.D (Spur 3) und JH.D [GFP-CbfA<sup>724-998</sup>] (Spur 4); Molekülmasse in kDa

In Abbildung 11 ist der proteinspezifische Nachweis elektrophoretisch aufgetrennter Proben (*Western-Blot*) jeweils eines Klons dargestellt. Es wurde der primäre, monoklonale Antikörper 7F3, welcher gegen die CTD gerichtet ist, eingesetzt. Bei 115 kDa ist deutlich die unterschiedliche Expression an CbfA in AX2 (Spuren 1 und 2) sowie JH.D (Spuren 3 und 4) zu sehen. Die Menge an GFP-CbfA<sup>724-998</sup> in den Transformanten von AX2 und JH.D war vergleichbar. Bei der Bande um 30 kDa handelte es sich wahrscheinlich um ein Proteolyseprodukt von GFP-CbfA<sup>724-998</sup>. Ähnliche Mengen an gespaltenem Fusionsprotein waren nach Entfernung der ersten Antikörper und Detektion mit einem GFP-spezifischen Antikörper sichtbar (Daten nicht gezeigt).

Bei dem Stamm JH.D handelt es sich um einen Stamm, welcher durch Amber-Stopp-Codon-Suppression des *cbfA*-Gens generiert wurde. Bisherige Versuche, *cbfA* komplett zu entfernen, schlugen fehl, was auf eine für das Überleben von *D. discoideum* wichtige Funktion von CbfA schließen lässt. (Winckler *et al.* 2001). Die gewählte Methode führt zu einer Verringerung der Expression durch Einführung eines Stopp-Codons in die codierende Sequenz. Es kommt zum regulären, Stopp-Codon-bedingten Abbruch der Polypeptidkettenbildung mit *cbfA*(amber)-RNA als Template am Ribosom. Der Abbruch der Expression von CbfA erfolgt an dem im Wildtyp für Glutaminsäure 453 codierenden Triplet und liegt im Bereich vor dem zweiten Zinkfinger. Durch die zusätzliche Einführung einer Suppressor-tRNA, welche das Amber-Stopp-Codon UAG erkennt, aber mit Glutaminsäure beladen ist, wird dieser Prozess zu einem gewissen Prozentsatz verhindert (Dingermann *et al.* 1990 und 1992). Dadurch wird die komplette Polypeptidkette – hier CbfA – gebildet. Dies ist im Falle von JH.D für die Lebensfähigkeit von *D. discoideum* ausreichend (Winckler *et al.* 2001). Das durch Suppression entstandene CbfA<sup>1-452</sup>, welches eine Masse von 52 kDa besitzt, kann wie auch Gesamt-CbfA mit Hilfe des Antikörpers 3H7 in Zelllysaten nachgewiesen werden (Abbildung 12 A). Bei 3H7 handelt es sich um einen monoklonalen Antikörper, der gegen die JmjC-Domäne gerichtet ist. Für die Untersuchung der genregulatorischen Eigenschaften allein der Carboxy-terminalen Domäne von CbfA in JH.D sollte eine zur Bildung eines funktionellen Gesamt-CbfAs in Frage kommende Interaktion mit dem Amino-terminalen Teil ausgeschlossen werden. Um diese Interaktion zumindestens teilweise zu widerlegen, wurden Kernextrakte von AX2, JH.D und JH.D [CbfA<sup>724-998</sup>] hergestellt.



**Abbildung 12: Zelluläre Verteilung von CbfA, durch CbfA-Suppression entstandenes CbfA<sup>1-452</sup> und ektopisch exprimiertes CbfA<sup>724-998</sup>; Western-Blot von (A) Gesamtzelllysat mit dem gegen den N-Terminus von CbfA gerichteten Antikörper 3H7 (B) Kernextrakt mit 3H7 (C) Kernextrakt mit dem gegen die CTD gerichteten Antikörper 7F3; AX2 (Spur 1), JH.D (Spur 2) und JH.D [CbfA<sup>724-998</sup>] (Spur 3); Molekülmassen in kDa**

Auf eine N-terminale Fusion des GFP an die CTD ist wegen möglicher negativer Effekte auf die Interaktion verzichtet worden. In Abbildung 12 B und C sind die *Western-Blots* gleicher Mengen der Kernextrakte mit den Antikörpern 3H7 gegen den Amino-Terminus und 7F3 gegen den Carboxy-Terminus zu sehen. Eine funktionelle Interaktion zwischen NTD und CTD würde zu einer Anreicherung des Amino-terminalen Proteinfragmentes im Zellkern führen, obwohl dieses keine eigene detektierbare Kernlokalisierungssequenz hat.

Der gegen den Amino-terminalen Teil gerichtete Antikörper ergibt nur eine spezifische Bande bei 115 kDa (Abbildung 12 B). Dies entspricht CbfA. Es sind keine weiteren Fragmente erkennbar. Im Teil C der Abbildung ist neben der CbfA-spezifischen Bande in Spur 3 (JH.D [CbfA<sup>724-998</sup>]) die erwartete CTD-Bande bei 31 kDa zu sehen. Dementsprechend sind die nachfolgenden beschriebenen Ergebnisse mit hoher Wahrscheinlichkeit nicht auf eine funktionelle Interaktion des durch Stopp-Codon-Suppression gebildeten Amino-terminalen Teils mit der CTD zurückzuführen, sondern sind direkte Folgen der zusätzlichen Expression der CTD.

Für die vergleichende Transkriptomanalyse der Stämme AX2, JH.D, AX2 [GFP-CbfA<sup>724-998</sup>] und JH.D [GFP-CbfA<sup>724-998</sup>] wurde Anfang 2008 in der Arbeitsgruppe von Privatdozent Dr. Eichinger in

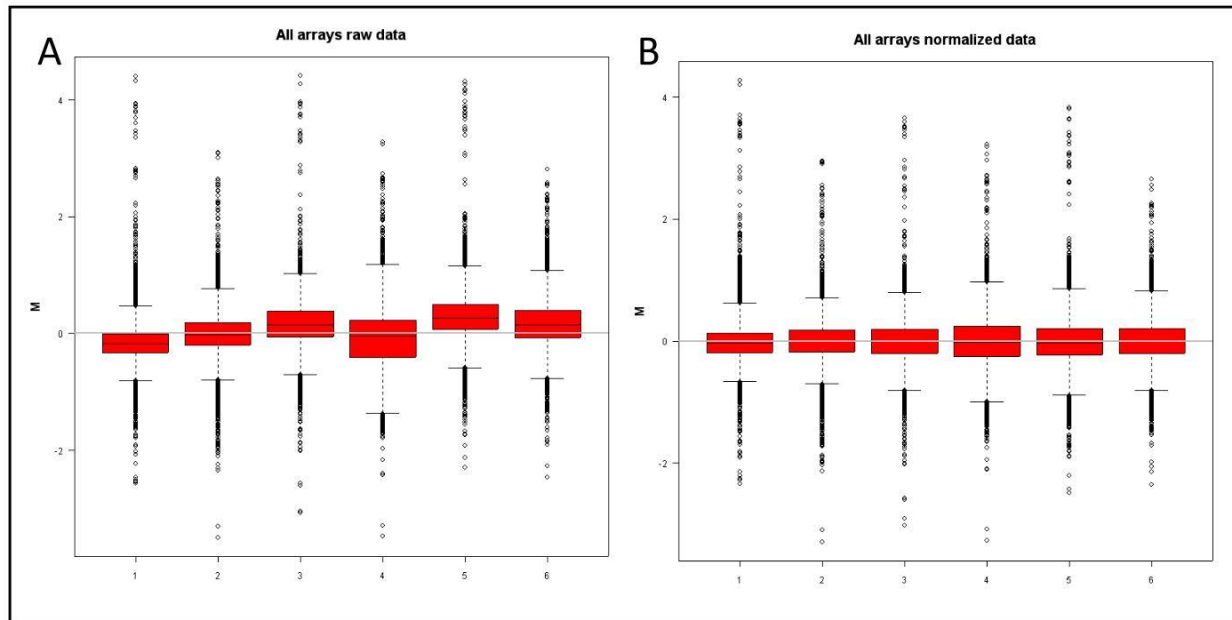
Köln eine Serie von *Microarray*-Experimenten durchgeführt. Um Aussagen sowohl über CbfA, als auch über die CTD treffen zu können, wurden folgende Kombinationen untersucht:

<u>AX2</u>	<u>JH.D GFP-CbfA<sup>724-998</sup></u>	<u>AX2 GFP-CbfA<sup>724-998</sup></u>
JH.D	JH.D	AX2

In der Literatur wird allgemein eine entgegengesetzte Darstellung benutzt. Die als Experiment definierte Mutante JH.D würde dann oben stehen und die Kontrolle AX2 unten. Die gefundenen Regulationen entsprechen dann dem Kehrwert der in dieser Arbeit dargestellten Werte. Da aber in späteren Experimenten die Komplementation der aberrenten Genexpression in JH.D durch die Carboxy-terminale Domäne gezeigt wird, ist es sinnvoller, die Genexpression in AX2 und JH.D [GFP-CbfA<sup>724-998</sup>] auf JH.D zu beziehen.

Für jede Serie von Experimenten wurden drei unabhängige RNA-Präparationen vegetativ wachsender Kulturen genutzt. Diese sind im Vorfeld des *Microarray*-Experimentes mit Hilfe des RNA 6000 Nano Chips (Agilent) auf ihre Qualität überprüft worden. Diese Technologie liefert neben zahlreich weiteren Informationen über die untersuchte Probe einen Zahlenwert (*RNA Integrity Number*) von 1-10, wobei 10 der höchsten Qualität vollkommener RNA entspricht. Alle benutzten Proben hatten einen Wert größer 8. Es wurde jede RNA-Probe jeweils mit einer entsprechenden anderen Probe untersucht. Dabei sind die Experimente mit beiden möglichen Kombinationen der Farbstoffe Cy3 und Cy5 durchgeführt worden (*dye swap*), um eventuelle farbstoffspezifische Effekte auszuschließen. Für jeden Vergleich wurden folglich sechs *Microarrays* hybridisiert. Der genutzte *Microarray* „*Dictyostelium discoideum* cDNA *Microarray* 6000“ enthielt 450 handverlesene Gene sowie 5.423 exprimierte Sequenzen (EST) aus dem Stadium des vegetativen Wachstums (Urushihara *et al.* 2004, Farbrother *et al.* 2006). Genauere Informationen über den Array sind im Material-Teil (2.14) oder auf der Seite des **Gene Expression Omnibus** (<http://www.ncbi.nlm.nih.gov/geo/>) unter der Referenznummer **GPL1972** zu finden. Jeder Array ist mit zwei unterschiedlichen Lasereinstellungen gescannt worden, um ein möglichst breites Spektrum an Intensitäten abzudecken. Nach der manuellen Kontrolle der Zuordnung der fluoreszierenden Spots durch das Programm **ScanArray Express 3.0** (Perkin Elmer) zu den aufgetragenen Genen, sind die Signalstärken quantifiziert worden. Dadurch konnte jedem auf dem Array vorhandenen Gen das Verhältnis der Farbintensität der beiden

Proben zugeordnet werden. Ferner sind nicht vorhandene oder als unzureichend erkannte *Spots* markiert worden. Die entstandene Liste der Primärdaten wurde für jede Serie vereinigt und mit dem Programm **R 1.6.2** (BioConductor) normalisiert. Dadurch werden die aufgrund der verschiedenen Lasereinstellungen und Farbstoffkombinationen unterschiedlich verteilten Primärdaten vergleichbar. Es wurde dabei die LOWESS-Methode (Yang *et al.* 2002) angewendet. In Abbildung 13 ist die Normalisierung für den Datensatz AX2 zu JH.D zu sehen.



**Abbildung 13: Normalisierung der Primärdaten einer Serie von *Microarrays*; (A) Primärdaten (B) Daten nach LOWESS-Normalisierung; Box-Whisker-plots von AX2 zu JH.D**

Zur Identifizierung differentiell exprimierter Gene ist das Programm **Significance analysis of microarrays** (SAM, Tusher *et al.* 2001) genutzt worden. Dieses verwendet einen erweiterten t-Test, welcher die erwartete und die tatsächliche Verteilung der Daten vergleicht. Des Weiteren gibt dieses Programm eine Abschätzung über die Anzahl falscher, als differentiell exprimiert identifizierter Gene an. Diese *false discovery rate* (FDR) lag in allen Experimenten unter 0,13%.

Für den Vergleich AX2 zu JH.D konnten 473 nicht redundante Gene als differentiell exprimiert identifiziert werden. Davon waren 235 Gene in AX2 höher exprimiert als in JH.D, 238 geringer. Legt man einen Schwellenwert von einer 1,5-fachen Regulation an, so ändern sich die Anzahl der regulierten Gene auf 80 in AX2 höher exprimierte Gene und 82 geringer exprimierte Gene.



Das für CbfA codierende Gen gehört zu den manuell ausgewählten Genen auf dem Microarray. Die Sonde AF052006 entspricht dem 3'-Ende von *cbfA*. Sie wird nicht differentiell exprimiert. Dies belegt, dass die CbfA-Unterexpression in JH.D nicht durch den Abbau der *cbfA*(amber)-mRNA, sondern durch partielle Stopp-Codon-Suppression entsteht.

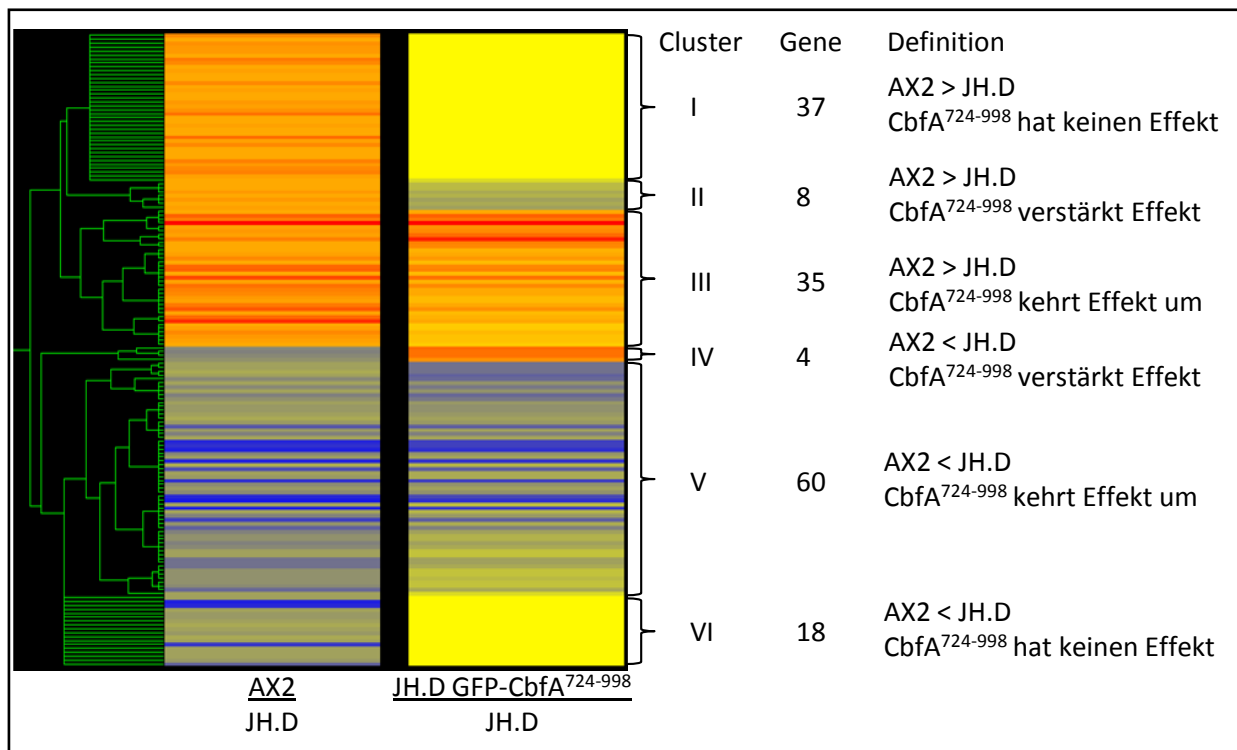
Der Vergleich JH.D [GFP-CbfA<sup>724-998</sup>] zu JH.D ergab 914 differentiell regulierte Gene. 408 (Schwellenwert 1,5x 79) dieser Gene waren in JH.D [GFP-CbfA<sup>724-998</sup>] höher exprimiert als in JH.D, 506 (88) geringer. Die *cbfA*-Sonde AF202006 zeigt eine Überexpression der für CTD codierenden RNA im Vergleich mit der endogenen *cbfA*-RNA um den Faktor 2,331 an.

Die dritte Serie von Experimenten resultierte in 1.391 unterschiedlich exprimierten Genen. Dabei waren in AX2 [GFP-CbfA<sup>724-998</sup>] 568 Gene stärker exprimiert als in AX2, 823 geringer. Unter Einbeziehung des Schwellenwertes von 1,5 sinken die Zahlen auf 153 in AX2 [GFP-CbfA<sup>724-998</sup>] im Vergleich zu AX2 überexprimierte Gene und 82 unterexprimierte Gene. Dies sind 114 auch in der JH.D-Umgebung durch die CTD 1,5-fach regulierte Gene. 49 über dem Schwellenwert von 1,5 regulierte Gene werden in JH.D nur um den Betrag von 1,2 bis 1,5 durch GFP-CbfA<sup>724-998</sup> differentiell reguliert, 4 sind nicht signifikant reguliert. Neben den 167 sowohl in AX2 und JH.D durch GFP-CbfA<sup>724-998</sup> regulierten Genen sind in AX2 68 zusätzliche vorhanden, welche in der JH.D-Umgebung durch GFP-CbfA<sup>724-998</sup> nicht differentiell reguliert werden. Diese Beobachtung kann durch einen dominant negativen Effekt der Überexpression der CTD in der AX2 Umgebung, welche 100% CbfA bildet, erklärt werden. Dementsprechend wurden diese Gene nicht in die weiteren Analysen einbezogen. Das Maß der Überexpression der CTD lag in dem Vergleich AX2 [GFP-CbfA<sup>724-998</sup>] zu AX2 bei 2,358 bezogen auf die endogene *cbfA*-RNA.

Es ist zu erwähnen, dass die Experimente in der oben erwähnten Reihenfolge durchgeführt worden sind. Die statistische Zuverlässigkeit der Daten stieg über die Zeit an. Dadurch erhöhte sich die Anzahl differentiell exprimierter Gene mit Regulationen unter 1,3x, während die Menge an höher regulierten Genen vergleichbar blieb. Ferner ist die biologische Bedeutung von geringen Regulationen fraglich sowie schwer nachweisbar. Deshalb werden nachfolgend nur Gene betrachtet, welche durch CbfA mindestens 1,5-fach reguliert werden.

### 3.3 Clusteranalyse der erhobenen *Microarray*-Daten

Das Ergebnis einer *Microarray*-Analyse ist eine Liste von Genen und ihrer differentiellen Expression aufgrund eines Experimentes. Sie ist nur der Startpunkt für weitere Analysen. Eine Möglichkeit ist dabei die Clusteranalyse, wobei die regulierten Gene aufgrund ihrer Expression an mehreren Zeitpunkten oder in unterschiedlichen Stämmen hierarchisch geordnet werden. Man nimmt an, dass ähnlich oder gleich regulierte Gene einen funktionellen Zusammenhang haben (u.a. Farbrother *et al.* 2006, Na *et al.* 2007). Für die folgende Analyse wurden nur die Daten der Experimente AX2 zu JH.D sowie JH.D [GFP-CbfA<sup>724-998</sup>] zu JH.D genutzt und das Programm **Gene Spring 7.2** (Agilent) verwendet. Der Endpunkt der Analyse ist in Abbildung 14 dargestellt. Die Liste der verwendeten Gene, ihr Expressionsprofil sowie die Zuordnung zu den Clustern befindet sich im Anhang B.



**Abbildung 14:** Clusteranalyse der durch CbfA-Suppression regulierten Gene und ihre Reaktion auf GFP-CbfA<sup>724-998</sup>; Stärke der Regulation ist farbcodiert, Regulation größer 1 (höhere Expression des Gens in AX2 als in JH.D) rot, gleich 1 gelb, kleiner 1 (geringere Expression des Gens in AX2 als in JH.D) blau

Es konnten sechs Cluster zugeordnet werden. Die Gene der Cluster I bis III sind in AX2 höher exprimiert als in JH.D. Da die Suppressionsmutante nur etwa ein zwanzigstel der Wildtypmenge an CbfA bildet, kann man vereinfachend sagen, dass diese Gene durch CbfA aktiviert werden. Dementsprechend werden die Gene in Cluster IV bis VI durch CbfA reprimiert. Jedem dieser Gene ist seine Reaktion auf GFP-CbfA<sup>724-998</sup> in der JH.D Umgebung zugeordnet. Gene, welche in dieser Serie von Experimenten nicht als differentiell exprimiert erkannt wurden, haben per Definition eine Regulation von 1. Ihre Codierung in Abbildung 14 ist gelb. Dabei ist es möglich, dass der Analyseweg eine Regulation ungleich 1 bestimmt, diese aber nicht für statistisch signifikant hält.

In Cluster I sind 37 Gene zusammen gefasst, welche durch CbfA positiv reguliert werden. Sie zeigen keine Reaktion bezogen auf ihre Expression durch die Einführung von GFP-CbfA<sup>724-998</sup>. Die Carboxy-terminale Domäne (CTD) ist nicht ausreichend für die Wiederherstellung des Expressionsniveaus durch CbfA, wobei jedoch nicht ausgeschlossen werden kann, dass sie an der Regulation beteiligt ist.

Cluster II umfasst acht positiv durch CbfA regulierte Gene, welche durch die CTD in JH.D noch weniger exprimiert werden. Dies bedeutet, dass die Überexpression von GFP-CbfA<sup>724-998</sup> den durch Reduktion von CbfA erzielten Effekt noch verstärkt.

Der mit 35 Genen vertretene Cluster III verhielt sich in seiner Reaktion auf die CTD entgegengesetzt zum Cluster II. Die verloren gegangene Regulation durch CbfA in JH.D wurde durch GFP-CbfA<sup>724-998</sup> wieder erreicht. Dabei ist zu erwähnen, dass die Expression oft im gleichen Umfang wiederhergestellt wurde. Dies ist an einer sehr ähnlichen Farbcodierung der beiden Spalten in Abbildung 14 zu erkennen. Die Funktion von CbfA kann durch die CTD alleine erfüllt werden. Zu diesem Cluster gehört auch das durch die Sonde SLJ671 vertretene Gen DDB\_G0293690, welches mit einer Identität von 98% als Fragment eines Retrotransposons vom TRE5-A-Typ identifiziert werden kann. Es wird in AX2 2,3-fach mehr exprimiert als in JH.D, GFP-CbfA<sup>724-998</sup> steigert die Expression in JH.D um den Faktor 4,7. Dies entspricht den *Northern-blot* in Abbildung 10.

Cluster IV hat mit vier Genen die kleinste Anzahl. Sie werden durch CbfA negativ reguliert, dass Vorhanden sein von GFP-CbfA<sup>724-998</sup> steigert deren Expression.

Fast 40% aller analysierten Gene befinden sich im Cluster V. Diese Gene werden durch CbfA negativ beeinflusst, dies bedeutet, dass ihre Expression in JH.D größer ist als in AX2. Entsprechend Cluster III wird hier die Funktion durch GFP-CbfA<sup>724-998</sup> wieder erlangt.

Cluster VI beinhaltet 18 Gene. Diese durch CbfA negativ regulierten Gene zeigen keine expressionelle Reaktion auf die CTD. Hier gelten die gleichen Aussagen wie für die Cluster I Gene.

### 3.4 Verifizierung der *Microarray*-Daten

Zum Nachweis der Richtigkeit der erlangten Daten wurden zwei unterschiedliche Ansätze gewählt. Aufgrund der Struktur der benutzten *Microarrays* sind für viele Gene mehr als eine Sonde vorhanden. Deshalb wurden jeweils zehn mit mindestens zwei Sonden vertretene Gene, welche differentiell exprimiert werden, überprüft. Für alle Serien war die Art der Regulation der unterschiedlichen Sonden eines Genes identisch. Bei jeweils acht Genen betrug die maximale Abweichung 10% (Tabelle 5). Des Weiteren sind drei nicht signifikant unterschiedlich exprimierte Gene überprüft wurden. Auch hier verhalten sich die verschiedenen Sonden ähnlich. Diese Analyse basiert auf dem gleichen Experiment wie die Primärdaten.

**Tabelle 5: Vergleich der Regulation von Genen mit mehreren Sonden**

Vergleich	Ø Abweichung	maximale Abweichung
AX2 vs. JH.D	11%	48% (DDB_G0277761 mit Sonden VSD732 1,541x und SSC695 2,278x)
JH.D GFP-CbfA <sup>724-998</sup> vs. JH.D	9%	35% (DDB_G0271916 mit Sonden U48298 1,249x und SSL284 1,682x)
AX2 GFP-CbfA <sup>724-998</sup> vs. AX2	13%	39% (DDB_G0272560 mit Sonden X52688 1,359x und SLI144 1,884x)

Der zweite Ansatz zur Verifizierung nutzt eine andere Methode zur Quantifizierung von Expressionsmustern. Bei der quantitativen Echtzeit-Polymerase-Kettenreaktion an einer komplementären DNA (qRT-PCR) können im Gegensatz zum *Microarray*, in dem bis zu einem gesamten Transkriptom untersucht werden kann, nur einige Gene betrachtet werden. Es sind

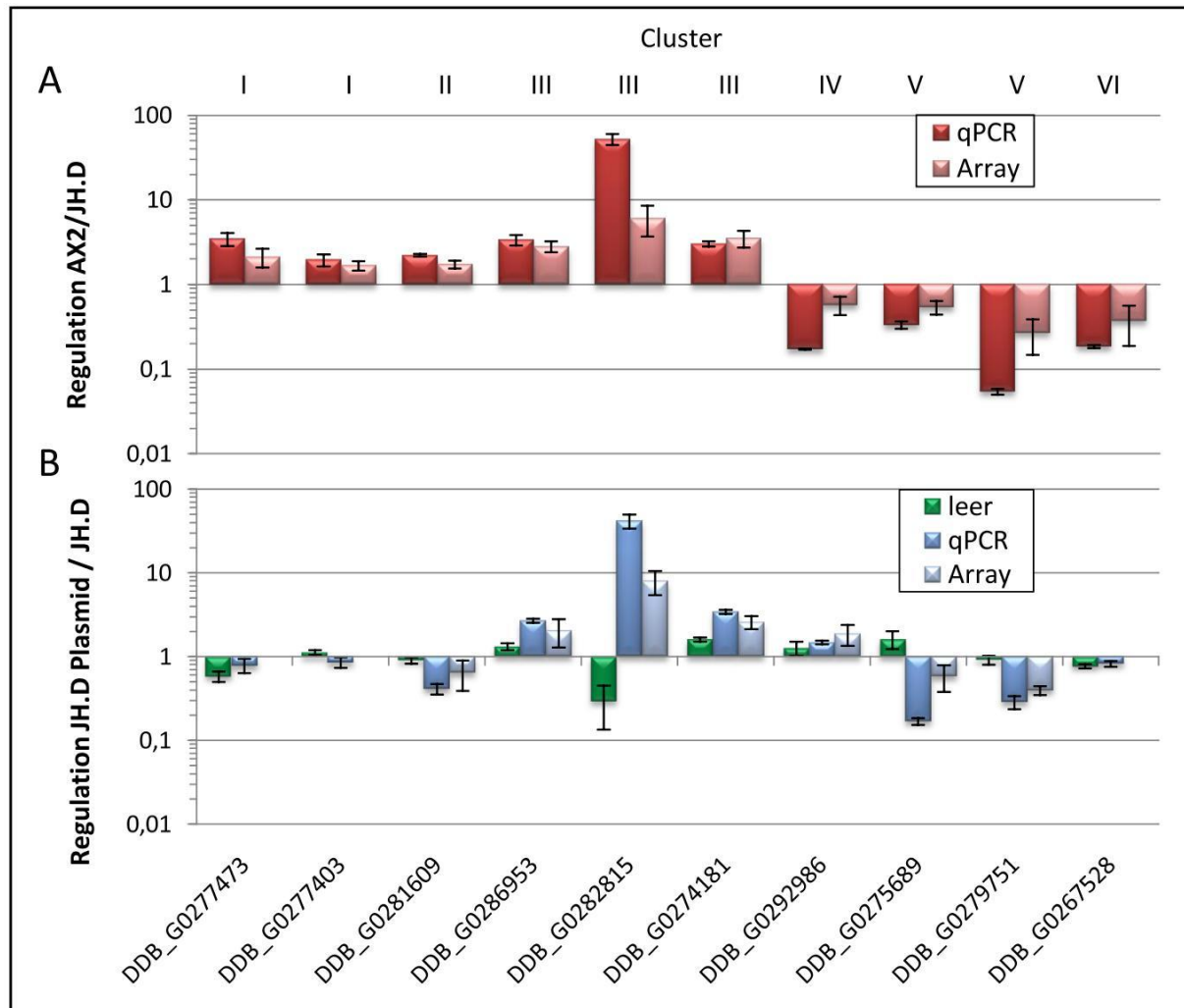
deshalb für jedes erwähnte Cluster (Abschnitt 3.3) ein bis drei Gene ausgewählt worden. Für Cluster unter 20 Genen ist ein Gen gewählt worden, für Cluster über 20 sind es zwei. Für das Cluster III wurde zusätzlich das Gen *orfSGP* (DDB\_G0282815) betrachtet, da es in Proteomuntersuchungen als das am stärksten von CbfA regulierte Protein identifiziert wurde (Jennes 2006). Die Auswahl der restlichen Gene erfolgte willkürlich, wobei ohne Kenntnis ihrer Funktion darauf geachtet wurde, dass am 3'-Ende eine spezifische PCR-Bande um 180 bp entsteht. Für einige Gene erstreckt sich die Probe über ein Intron, sodass die Qualität der cDNA-Präparation überprüft werden konnte.

Alle Messungen wurden gegen das Gen für die Glycerinaldehyd-3-Phosphat-Dehydrogenase (*gpdA*, DDB\_G0275153) standardisiert. Die für dieses Gen auf dem Microarray vorhandene Sonde SLD171 war in allen Experimenten nicht differentiell exprimiert. Die Regulation wurde mit einer modifizierten delta-delta-Methode nach Pfaffl (2001) bestimmt. Im Teil A der Abbildung 15 ist die Verifizierung des Vergleiches AX2 zu JH.D zu erkennen. Die Regulationsart war bei beiden Methoden für alle überprüften Transkripte gleich. Eine von ihrem Betrag höhere Regulation in der qRT-PCR ist ein sehr oft beobachtetes Phänomen, welches noch nicht zufriedenstellend geklärt werden konnte (u.a. Na *et al.* 2007).

Der Teil B der Abbildung 15 zeigt mit dem in grün dargestellten Balken den Einfluss einer Transformation auf die Expression ausgewählter Gene. Es gibt eine Reihe von Veröffentlichungen, die zeigen, dass einige Änderungen des Expressionsprofils eines Organismus allein auf den Prozess der Transformation bzw. Transfektion zurückzuführen sind (u.a. Jacobsen *et al.* 2009). Um dies für die selektierten Gene auszuschließen, wurde ein leerer Expressionsvektor in JH.D transformiert und der entstandene JH.D [leer] mit JH.D verglichen. Die transkriptionellen Änderungen sind bis auf eine Ausnahme gering. Das Cluster III-Gen DDB\_G0282815 wird in JH.D 3,4-mal stärker exprimiert als in JH.D [leer]. Die Expression in JH.D [GFP-CbfA<sup>724-998</sup>] ist aber um den Faktor 42 höher als in JH.D. Der mathematische Vergleich JH.D [GFP-CbfA<sup>724-998</sup>] zu JH.D [leer] ergibt eine über 140-fache, plasmidbereinigte Regulation. Um eine bessere Vergleichbarkeit mit den Arraydaten zu ermöglichen und aufgrund der geringen Änderungen der anderen untersuchten Gene wird im weiteren Verlauf der Arbeit der Effekt der Leertransformation vernachlässigt. Die Abbildung 15 B zeigt ferner die Überprüfung der Daten

## Ergebnisse

für die Funktion von GFP-CbfA<sup>724-998</sup> in JH.D. Für Cluster I und VI ist die Regulation der Arraydaten entsprechend dem Punkt 3.3 per Definition 1. Die durch qRT-PCR ermittelten Werte haben hier einen Betrag kleiner 1,3. Für die restlichen Cluster ist die Art der Regulation durch beide Methoden gleich. Es gelten die gleichen Aussagen wie für AX2 zu JH.D.



**Abbildung 15: Verifizierung der Arraydaten mittels qRT-PCR; (A)** für den Vergleich AX2 zu JH.D **(B)** JH.D [leer] zu JH.D (grün) bzw. JH.D [GFP-CbfA<sup>724-998</sup>] zu JH.D (blau); gezeigt sind Mittelwerte und Standardabweichung von 6 *Microarray*-Experimenten beziehungsweise die Quantifizierung von 3 unabhängigen RNA-Präparationen mittels qRT-PCR; die Darstellung erfolgt aufgrund besserer Übersichtlichkeit logarithmisch

### 3.5 Differentiell exprimierte Gene im Vergleich AX2 zu JH.D bilden Anhäufungen in funktionellen Kategorien – GOAT-Analyse

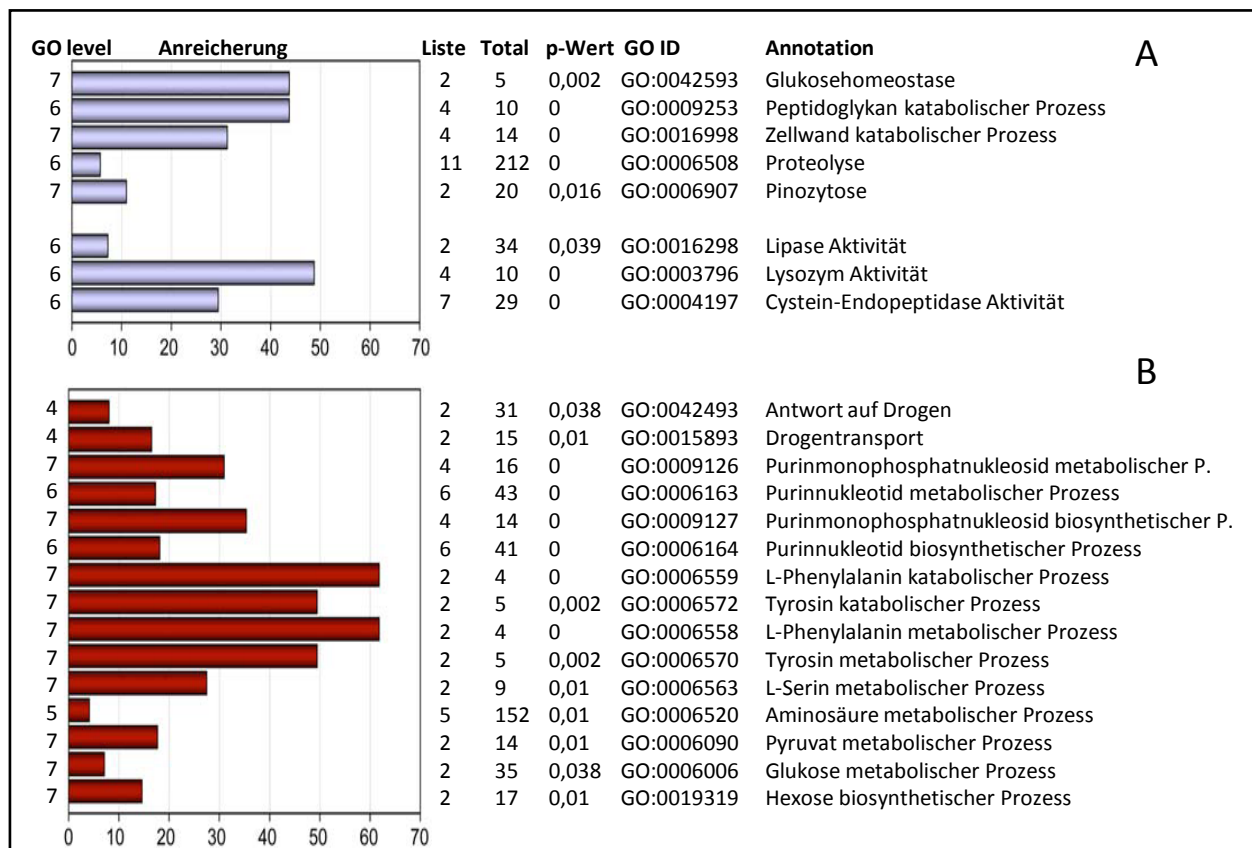
Eine oftmals sehr schwierige Aufgabe von Transkriptom- und Proteomanalysen ist die Zuordnung von molekularen Funktionen und biologischen Prozessen zu einer Liste von regulierten Genen. Einen großen Fortschritt stellt dabei das „gene ontology project“ (GO; <http://www.geneontology.org/>) dar, welches eine einheitliche Plattform für die Annotation ganzer Genome liefert. Von den etwa 12.500 *D. discoideum*-Genen sind zum Zeitpunkt dieser Arbeit 7407 Genprodukte annotiert (<http://www.geneontology.org/GO.current.annotations.shtml>, Stand 16.09.2009). Die Beschreibung teilt sich in die Kategorien molekulare Funktion, biologischer Prozess und Lokalisierung des Genprodukts auf. Eine Zuordnung der Annotationen sowie eine statistische Auswertung erfolgte mit dem Anwendungspaket **Gene Ontology Analysis Tool** (GOAT, Xu und Shaulsky 2005) für **R 1.9.1** (BioConductor).

Aufgrund des Datensatzes AX2 zu JH.D können Aussagen über die Funktion der von CbfA regulierten Gene getroffen werden. In Abbildung 16 sind einige Prozesse und molekulare Funktionen mit ihrer Anreicherung unter den differentiell exprimierten Genen dargestellt. Die komplette Liste der Annotationen mit statistischer Signifikanz  $p < 0,05$  und mindestens zwei Genen befindet sich im Anhang D. Die im Teil A der Abbildung genannten Anreicherungen positiv von CbfA regulierter Gene konnten neben der Glukosehomeostase unter den Begriffen Aufnahme und Verdauung von Bakterien zusammengefasst werden. Die in Verbindung zur Glukosehomeostase assoziierten Gene *cf45-1* (DDB\_G0269248) und *cf50-1* (DDB\_G0273175) sind Komponenten des *counting factor complexes*, welcher die Größe der Aggregate während der Fruchtkörperbildung steuert (Brock *et al.* 2002 und 2003). Eine Überprüfung des Datensatzes auf die zwei anderen bekannten Komponenten dieses Signalkomplexes zeigte eine differentielle Expression des Genes *cf60* (DDB\_G0284363) in ähnlichem Ausmaß wie die beiden oben genannten. Countin (*ctnA*, DDB\_G0274597) wurde nicht differentiell exprimiert. Eine Abfrage der Gene über **dictyExpress** (<http://www.ailab.si/dictyexpress/run/> Rot *et al.* 2009) zeigte die höchste Expression bei der Initiierung der Entwicklung während bzw. am Ende der vegetativen Wachstumsphase. Untersuchungen in JH.D in Bezug auf die basale Expression der

## Ergebnisse

Adenylylcyase A deuten auf einen ähnlichen Zeitpunkt des Einflusses von CbfA auf die Entwicklung hin (Siol *et al.* 2006a).

Auch unter den reprimierten Genen (Teil B) lassen sich mehrere Anhäufungen feststellen. Dies sind neben den Aminosäurestoffwechsel auch zwei Hauptenzyme der Glukoneogenese, Phosphoenolpyruvat-Carboxykinase (*pckA*, DDB\_G0271678) und Fruktose-1,6-bisphosphatase (*fbp*, DDB\_G0270836). In Anbetracht der differentiell exprimierten Gene des Nukleotidstoffwechsels, kann in der Bereitstellung von Glukose als der Ausgangspunkt für die Synthese von Ribosen durch den Pentose-Phosphat-Weg hier ein Wirken in der *De-novo*-Synthese von Nukleotiden vermutet werden.



**Abbildung 16: GOAT-Analyse CbfA regulierter Gene;** Auswahl an (A) aktivierter Gene (B) reprimierter Gene; GO level gibt hierarchische Ordnung der Annotation an, wobei 1 der allgemeinsten entspricht, Anreicherung ergibt sich aus dem Verhältnis der Gene in der Liste und aller Gene (Total) mit dieser Annotation in Bezug auf die Gesamtzahl der Gene in der Liste und der annotierten Gene, p-Wert entspricht der statistischen Signifikanz, GO ID und Annotation definieren den Prozess bzw. die Aktivität



Unter den angereicherten Genen befinden sich fünf Gene der *De-novo*-Purinnukleotidsynthese (*guaA*, *guaB*, *purB*, *purC/E*, *purD*) und eines der *De-novo*-Pyrimidinsynthese (*pyr4*). Überprüft man die restlichen in den Synthesewegen vorhandenen Gene, so stößt man für die Purinnukleotidsynthese auf zwei weitere, 1,3-fach in JH.D höher regulierte Gene (*purF* und *purH*). Die Gene *purL* und *purN* sind auf dem *Microarray* nicht vorhanden, *purA* wird nicht signifikant in seiner Expression beeinflusst. Die an der Pyrimidinnukleotidsynthese beteiligten Enzyme *pyr1-3*, *pyr56* und *pyrK* werden alle 1,4-fach in JH.D stärker exprimiert.

Des Weiteren sind vier ABC-G-Drogentransporter (*abcG2*, *abcG3*, *abcG10* und *abcG21*) in AX2 geringer exprimiert als in JH.D. In die Analyse der Abbildung 16 B fließen nur *abcG2* und *abcG3* ein, da *abcG10* und *abcG21* zum Zeitpunkt der Arbeit nicht in der GO-Ausgangsliste vertreten waren. Eine manuelle Berechnung der Anreicherung ergibt 25 (4 gefundene abcG-Transporter zu 24 im Genom im Verhältnis zu 82 reprimierten Genen bei etwa 12.500 Genen gesamt).

### 3.6 Neue Transkriptomansätze

Die Transkriptomanalyse mittels *Microarrays* ist eine weit verbreitete und sehr ausgereifte Technik (u.a. Churchill 2002, Leung und Cavalieri 2003). Wie jede andere Methode besitzt auch sie Grenzen. Zwei limitierende Parameter dabei sind das Vorhandensein eines Arrays für die jeweilige Fragestellung sowie der Fakt, dass nur Aussagen über Gene getroffen werden können, die auf dem Array vertreten sind. Im Zeitalter immer schnellerer, billigerer und paralleler werdender Sequenzieretechniken entwickeln sich neue Ansätze zur Analyse von Transkriptomen. In Zusammenarbeit mit der Arbeitsgruppe von Dr. Glöckner am Leibnitz-Institut für Altersforschung in Jena werden deshalb die Sequenzieretechniken „454“, eine Form der Pyrosequenzierung, und „Solexa“, eine Ligase-vermittelte Sequenziermethode, auf die Fragestellung der CbfA-regulierten Gene angewandt.

Für die Erstellung des „454“-Datensatzes wurde mRNA aus vegetativ wachsenden AX2 und JH.D Kulturen isoliert. Diese mRNA diente als Ausgangsmaterial zur Sequenzierung, wobei die Anzahl der Zuordnungen gelesener Sequenzen (*Reads*) eines Genes als direkt proportional zu seiner Expression angesehen wird. Vergleicht man nun die auf die Gesamtzahl aller *Reads* normalisierten Werte eines Experimentes zu seiner Kontrolle, erhält man einen Wert, welcher

## Ergebnisse

vereinfachend der Regulation in einen Arrayexperiment entspricht. Für den „Solexa“-Datensatz wurde Total-RNA als Ausgangsmaterial verwendet. Die Sequenzierung und anschließende Zuordnung der Sequenzen zu den Genen erfolgte in der Arbeitsgruppe von Dr. Glöckner. Einige Eckdaten der Experimente sind in Tabelle 6 festgehalten. Beide Sequenzieretechniken erfassen wesentlich mehr Gene als der zuvor verwendete *Microarray*.

**Tabelle 6: Vergleich Transkriptomanalysen AX2 zu JH.D;** oberer Teil enthält Eckdaten über die Experimente; <sup>§</sup>erfasste Gene im Array setzen sich aus 5.423 cDNAs und 450 manuell gewählten Genen zusammen, wobei einige Gene bis zu 4-mal vertreten sein können; \*zur Vergleichbarkeit sind nur Gene in der oberen Hälfte der Gesamtreads ausgewertet; <sup>#</sup>Gene mit mindestens 2,5-facher Regulation unabhängig von der Gesamtread-Anzahl; Festlegung der Grenzen erfolgt nach dem 2. biologischen Replikat; unterer Teil Vergleich der Regulation der repräsentativen Gene der Cluster, Darstellung x-fache Regulation in AX2 zu JH.D

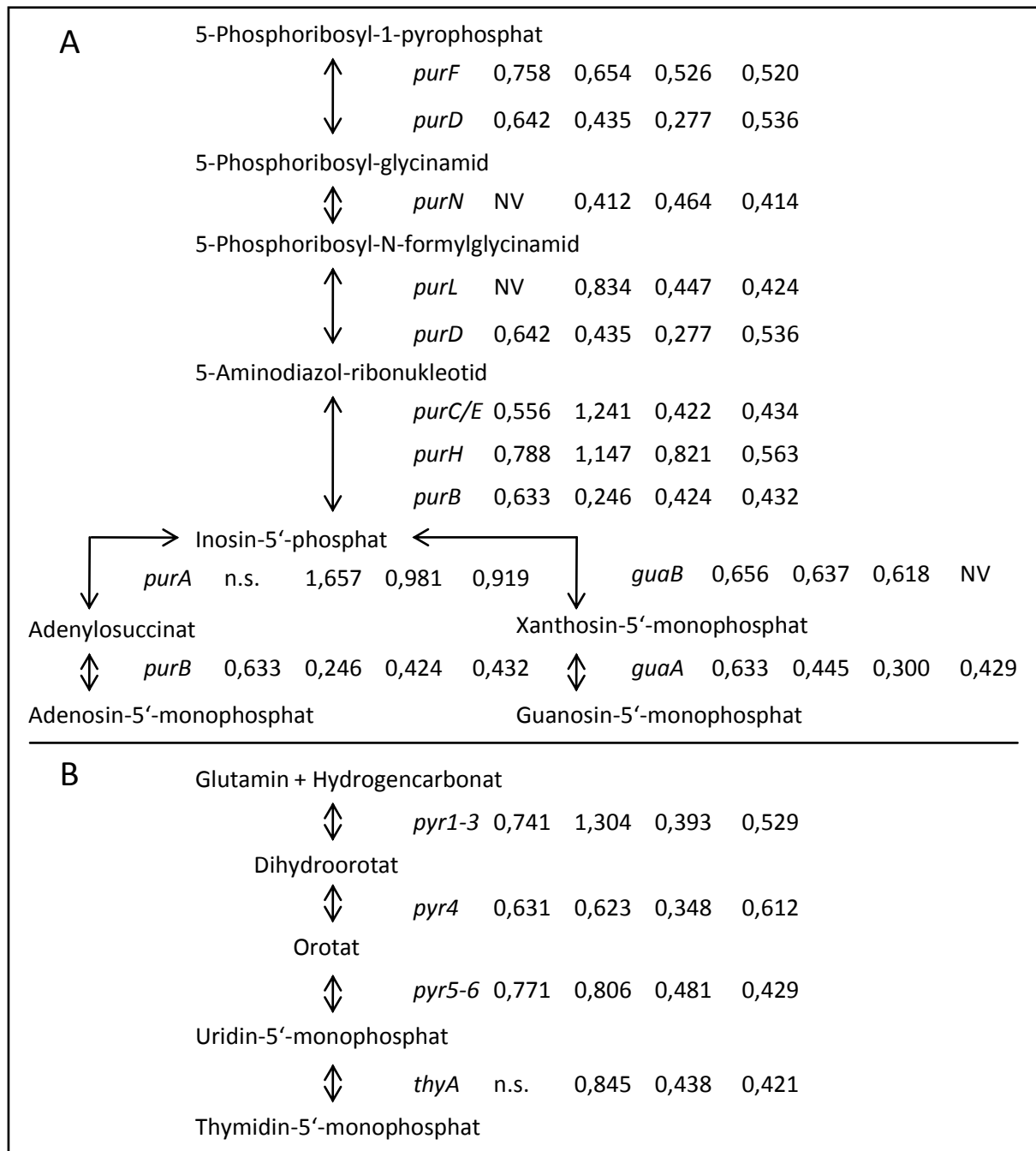
Technik	Array	454	Solexa
erfasste Gene (AX2; JH.D)	ca. 5.000 <sup>§</sup>	8.789 (7.787; 7.387)	10.045 (9.222; 9.744)
<i>reads</i> AX2		175.143	465.624.841
<i>reads</i> JH.D		144.789	509.433.802
∅ (min.; max.) <i>reads</i>		36 (1; 4008)	73.335 (11; 12.047.741)
50% (10; 90) Gene haben		> 5 (55; 0) <i>reads</i>	>7.554 (95.759; 330) <i>reads</i>
1,5x regulierte Gene (+; -)*	162 (80; 82)	482 (280; 202)	323 (200; 123)
2,5x regulierte Gene (+; -) <sup>#</sup>	30 (14; 16)	1359 (697; 672)	1671 (384; 1233)
Replikate	3 biologische je 2 technische	1 biologisches	1 biologisches (2. biologisches in Arbeit)
I-DDB_G0277473	2,120	3,269 (35 <i>reads</i> )	3,672 (42.737 <i>reads</i> )
I-DDB_G0277403	1,671	1,144 (132 <i>reads</i> )	2,407 (255.575 <i>reads</i> )
II-DDB_G0281609	1,727	7,021 (10 <i>reads</i> )	16,251 (70.709 <i>reads</i> )
III-DDB_G0286953	2,818	5,876 (52 <i>reads</i> )	7,349 (471.129 <i>reads</i> )
III-DDB_G0282815	6,091	48,044 (2252 <i>reads</i> )	33,747 (204.801 <i>reads</i> )
III-DDB_G0274181	3,529	3,093 (67 <i>reads</i> )	3,389 (217.795 <i>reads</i> )
IV-DDB_G0292986	0,572	0,321 (71 <i>reads</i> )	0,068 (143.105 <i>reads</i> )
V-DDB_G0275689	0,541	0,467 (22 <i>reads</i> )	0,147 (131.376 <i>reads</i> )
V-DDB_G0279751	0,267	0,130 (7 <i>reads</i> )	0,069 (15.328 <i>reads</i> )
VI-DDB_G0267528	0,373	0,299 (122 <i>reads</i> )	0,078 (110.714 <i>reads</i> )

Für die Nichterfassung aller etwa 12.500 Gene sind zahlreiche Gründe anzuführen. Einige Gene sind durch Genverdoppelung entstanden und sind sich in ihrer Sequenz so ähnlich, dass eine Unterscheidung nicht möglich ist. Dies gilt auch für die verschiedenen Familien der mobilen genetischen Elemente, welche knapp 10% des *D. discoideum* Genoms ausmachen (Winckler *et al.* 2002). Eine weitere Erklärungsmöglichkeit ist, dass die Expression des Genes so gering ist, dass es auch angesichts einer halben Milliarde *reads* die mRNA nicht detektierbar ist. Kandidaten hierfür sind alle Gene, welche nur in speziellen Situationen exprimiert werden, wie z.B. Gene der multizellulären Entwicklung oder durch physiologischen Stress induzierte Gene.

Zu der Anzahl differentiell exprimierter Gene in den neuen Transkriptomanalysen lässt sich nur schwer eine Aussage treffen. Die in der Tabelle angegebenen Werte sind durch rein empirische Grenzsetzungen gewonnen wurden, welche sich an Veröffentlichungen von Mane *et al.* (2009) und Wall *et al.* (2009) orientieren. In beiden Veröffentlichungen werden komplexere statistische Verfahren genutzt, welche aufgrund nicht übertragbarer Qualitätsdaten innerhalb unseres Experimentes nicht anwendbar sind. Deutlich ist allerdings zu erkennen, dass die Zahl der regulierten Gene sich erhöht. Dabei muss bedacht werden, dass die Arraydaten sich aus sechs Einzelexperimenten zusammensetzen und nur statistisch signifikante Gene gezählt werden. Benutzt man nur einen Datensatz, so werden auch hier im Durchschnitt 640 Gene 1,5-fach reguliert. Des Weiteren bergen natürlich die 4-5.000 zusätzlich untersuchten Gene ein Reservoir von in der Microarray-Analyse nicht berücksichtigten Zielgenen. Vergleicht man die Sequenzieretechniken, so fällt die 3.000-mal höhere Anzahl von zugeordneten Solexareads auf.

Im unteren Teil der Tabelle 6 sind die Regulationen, der für die Cluster aus Kapitel 3.3 als repräsentativ gewählten Gene, verglichen. Alle Regulationen haben die gleiche Richtung, unterscheiden sich aber teilweise im Betrag. Die Änderungen der durch RNA-Sequenzierungen entstandenen Datensätze sind stärker und liegen meist auf einem Niveau mit den qRT-PCR-Daten (Vergleich Abbildung 15 A). Das Gen DDB\_G0277403 bildet eine Ausnahme, da hier die differentielle Expression gemessen durch den *Microarray* höher als die innerhalb des 454-Experimentes ist. Zum exemplarischen Vergleich der durch die unterschiedlichen Techniken erzielten Daten dient die Abbildung 17, in der schematisch die *De-novo*-Synthese von Nukleosid-5'-monophosphaten dargestellt ist.

## Ergebnisse



**Abbildung 17: De-novo-Synthese von Nucleosid-5'-Monophosphaten im Vergleich der Regulation der beteiligten Gene; (A) Purin-5'-Monophosphatsynthese (B) Pyrimidin-5'-Monophosphatsynthese; gezeigt ist die Genregulation AX2 zu JH.D; erster Wert Microarray, zweiter Wert 454-Sequenzierung, dritter Wert Solexa-Sequenzierung, vierter Wert qRT-PCR (Mittelwert aus 3 Experimenten); NV = nicht vorhanden, n.s. = nicht differentiell exprimiert; Pfade durch **Metacyc** generiert (Caspi *et al.* 2007)**

Die in Abschnitt 3.5 erwähnten Gene der *De-novo*-Purinnukleotidsynthese, welche nicht auf dem Microarray vorhanden sind (*purL* und *purN*), werden in dem Solexa-Experiment zweimal weniger in AX2 als in JH.D exprimiert. Im 454-Experiment gilt dies auch für *purN*. Alle anderen genannten Gene der Nukleotidsynthese verhalten sich im Solexa-Experiment ähnlich wie in den *Microarray*-Daten. Der Betrag der negativen Regulation im Vergleich AX2 zu JH.D ist bis auf das Gen *purH* in der Solexa-Sequenzierung immer stärker. Das im *Microarray*-Datensatz nicht differentiell exprimierte Gen *purA* ist im Solexa-Experiment ebenfalls nicht reguliert, in den 454-Daten zeigt es eine positive Regulation durch CbfA. Zur zusätzlichen Einschätzung der Werte sind qRT-PCRs durchgeführt worden. Die mit dieser Methode erstellten Daten ähneln denen durch Solexa-Sequenzierung generierten.

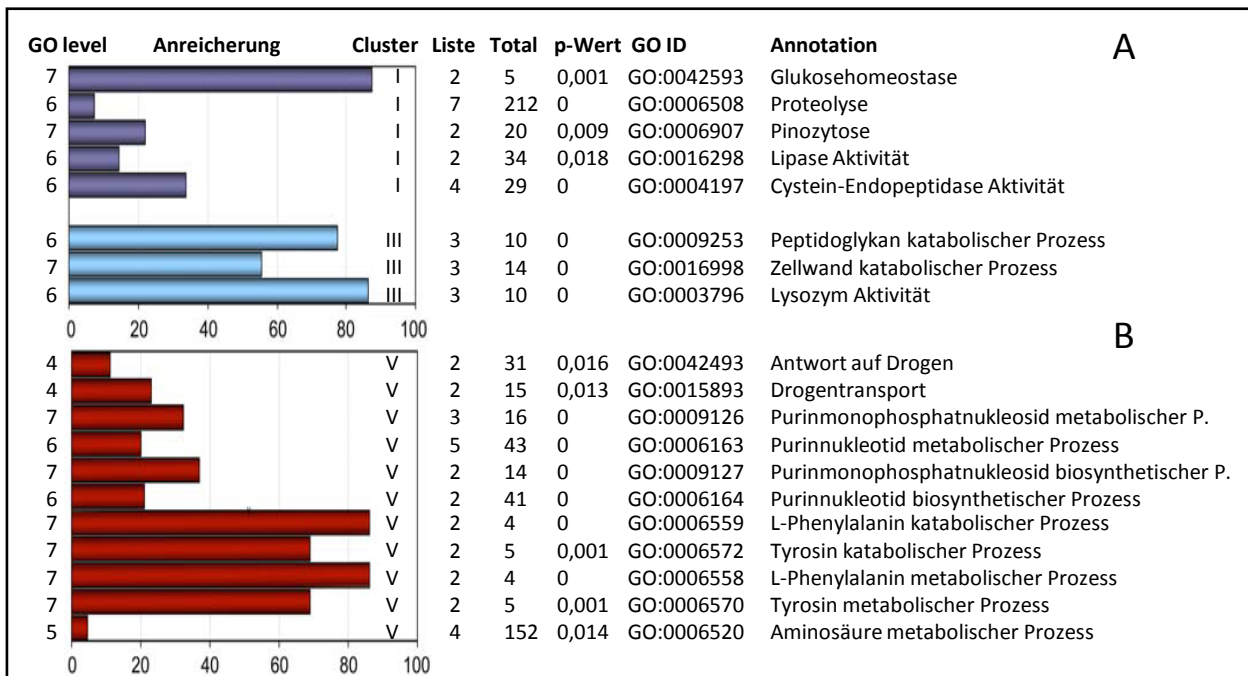
Die relative Expression der an der *De-novo*-Synthese von Pyrimidin-5'-Monophosphaten beteiligten Genprodukte ist in AX2 im Vergleich zu JH.D ebenfalls geringer. Dabei ähneln sich jeweils die durch *Microarray* und 454-RNA-Sequenzierung sowie die durch Solexa-RNA-Sequenzierung und qRT-PCR generierten Daten.

Bei den angewandten Techniken handelt es sich auch immer um biologische Replikate. Die beobachtete Varianz weist auf die Notwendigkeit der Untersuchung mehrerer solcher Replikate zum Treffen möglichst sicherer Aussagen hin.

### **3.7 GOAT-Analyse der CbfA-abhängigen Gene unter Betrachtung der Regulation durch GFP-CbfA<sup>724-998</sup>**

Die Daten, welche durch das *Microarray*-Experiment und durch die Transkriptomanalysen durch Sequenzierung erzielt worden, weisen deutlich auf eine oder mehrere Funktionen von CbfA in vegetativen Zellen hin. Für die Expression des Retrotransposons TRE5-A kann die durch einen *Northern-blot* (Abbildung 10) gezeigte Funktionalität der Carboxy-terminalen Domäne durch das *Microarray*-Experiment bestätigt werden. Neben dem TRE5-A-Gen DDB\_G0293690 deuten die unter 3.3 definierten Cluster auf weitere CTD-abhängige Gene hin. Entsprechend der unter 3.5 dargestellten Anreicherungsanalyse der CbfA regulierten Gene wurde diese Methode auch auf diese Genliste angewandt.

## Ergebnisse



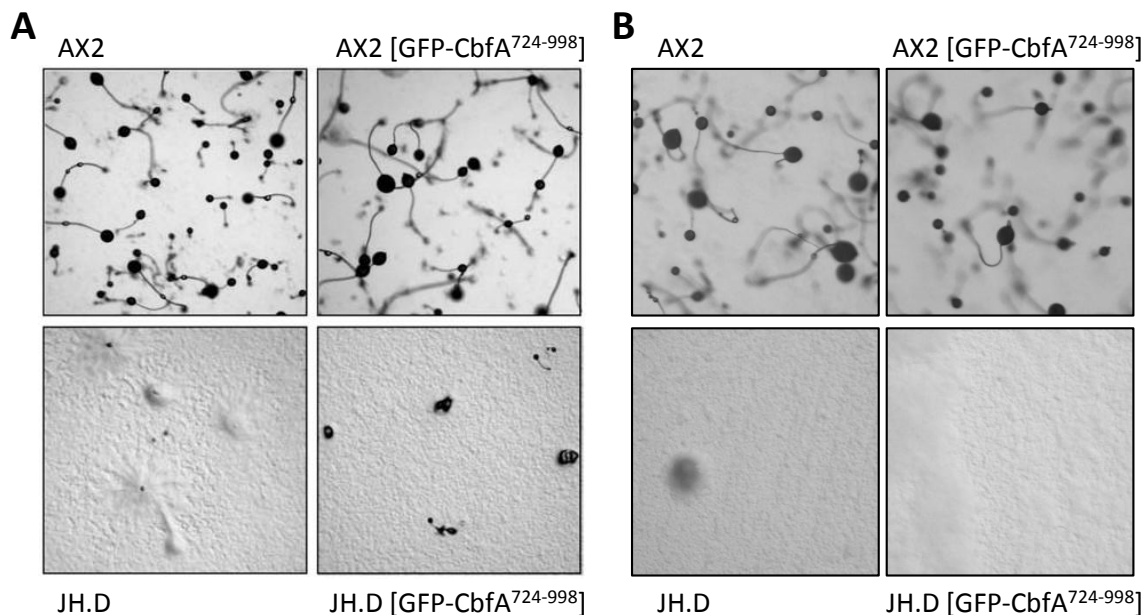
**Abbildung 18: GOAT-Analyse der durch CTD Co-regulierten Gene;** GO-Auswahl an **(A)** CbfA-aktivierten Cluster I und III **(B)** CbfA-reprimierten Cluster V; GO level gibt hierarchische Ordnung der Annotation an, wobei 1 der allgemeinsten entspricht, Anreicherung ergibt sich aus dem Verhältnis der Gene in der Liste und aller Gene (Total) mit dieser Annotation in Bezug auf die Gesamtzahl der Gene in der Liste und der annotierten Gene, Cluster ergibt sich aus Zuordnung im Kapitel 3.3, p-Wert entspricht der statistischen Signifikanz, GO ID und Annotation definieren den Prozess bzw. die Aktivität

Es konnten für vier Cluster signifikante Anreicherungen gefunden werden. Einige sind in der Abbildung 18 dargestellt, eine komplette Liste befindet sich im Anhang E. Für Cluster II und IV kann keine Anreicherung festgestellt werden, was aufgrund von nur vier bzw. acht Genen zu erwarten ist. Für die Bedeutung der CTD sind hier insbesondere die Cluster III und V interessant. Die durch Suppression von CbfA in JH.D bedingten Änderungen werden durch GFP-CbfA<sup>724-998</sup> umgekehrt. Einige Proteinprozesse, die am Abbau von Bakterien beteiligt sind, können durch die CTD allein aktiviert werden. Zu erwähnen ist in diesen Zusammenhang, dass die am *counting factor complex* beteiligten Genprodukte nicht durch den C-Terminus von CbfA reguliert werden. Sie befinden sich im Cluster I zusammen mit anderen Aktivitäten, die am Abbau von Bakterien beteiligt sind. Im Gegensatz dazu werden Teile des Drogentransportes, des Purinnukleotid- und des Aminosäurestoffwechsels durch die CTD autonom reprimiert. Die eventuell für die Bereitstellung von Ribose benötigte Glukoneogenese ist nur in der Expression von *pckA* von der CTD abhängig. Das Gen *fbp* wird nur durch Gesamt-CbfA reguliert.

### 3.8 Einfluss der Carboxy-terminalen Domäne auf die Entwicklung

Aufgrund der Ausrichtung des *Microarrays* können im Verlauf dieser Experimente keine Aussagen über entwicklungspezifische Prozesse getroffen werden. Bereits erwähnte Strukturen wie z.B. der *counting factor complex* besitzen zwar Einfluss auf die Entwicklung, werden aber hauptsächlich in vegetativen Zellen exprimiert.

Bereits in früheren Arbeiten wurde eine Abhängigkeit des Genes für die Adenylylcyclase A (*acaA*) von CbfA belegt (Siol *et al.* 2006a), was den möglichen Grund für den Ausfall praktisch aller getesteten Entwicklungsgene bedingt (Winckler *et al.* 2004). Um die Auswirkungen der CTD auf die Entwicklung zu testen, wurden mehrere positive Klone der Transformationen von AX2 und JH.D mit GFP-CbfA<sup>724-998</sup> in Schüttelkultur vermehrt und sowohl auf Phosphatplatten als auch auf Bakterienrasen ihre Fähigkeit zur makroskopischen Entwicklung überprüft. Als Vergleich dienten die nicht transformierten Ausgangsstämme. In Abbildung 19 sind mikroskopische Aufnahmen der multizellulären Entwicklung nach Abschluss des Zyklus in AX2 abgebildet.

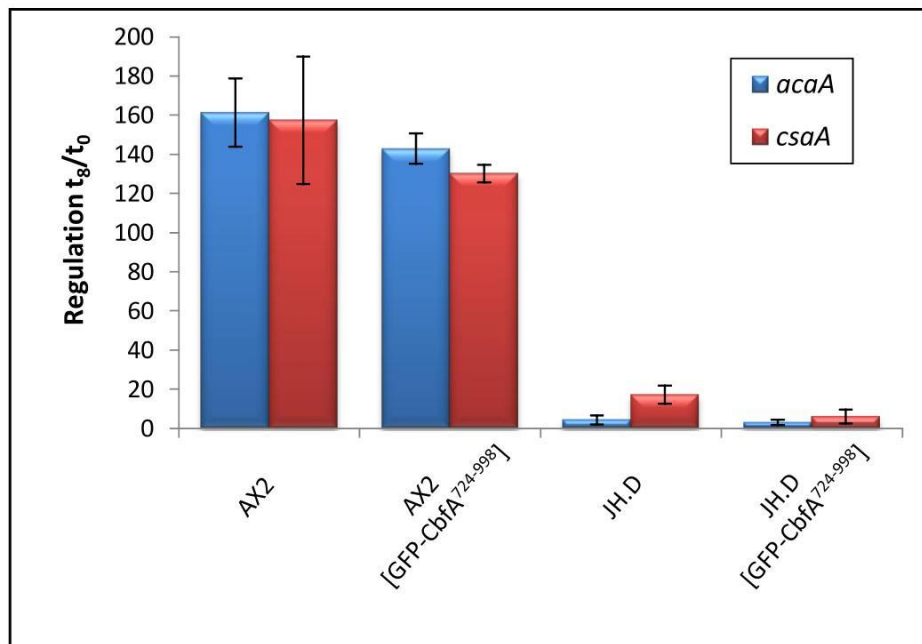


**Abbildung 19:** multizelluläre Entwicklung von AX2, JH.D und deren Transformanten mit GFP-CbfA<sup>724-998</sup>; (A) auf Phosphatplatten nach 30 Stunden (B) auf *Klebsiella planticola* nach 34 Stunden; alle Aufnahmen mit 57x Vergrößerung

## Ergebnisse

Eine längere Inkubationszeit resultierte nach 72 Stunden in einigen kleinen Fruchtkörpern sowohl in JH.D als auch in JH.D [GFP-CbfA<sup>724-998</sup>], was mit früheren Beobachtungen in Einklang zu bringen ist (Winckler *et al.* 2001). In beiden angewandten Tests waren die Ergebnisse ähnlich. Die Expression von GFP-CbfA<sup>724-998</sup> in JH.D hat keinen sichtbaren Einfluss auf den Aggregationsdefekt. Es sind außerdem keine Effekte in der AX2 Umgebung zu beobachten.

Entsprechend den mikroskopischen Beobachtungen ist auch bei diesem Experiment kein Einfluss von GFP-CbfA<sup>724-998</sup> in JH.D auf die Induktion der Entwicklungsmarker zu erkennen. Die Adenylylcyase A wird acht Stunden nach Beginn der Entwicklung drei- bis viermal mehr wie in der vegetativen Phase exprimiert. In der AX2-Umgebung hat diese Induktion den Faktor 140 bis 160. Das *csaA*-Transkript wird in JH.D sogar stärker induziert als in CTD-exprimierenden JH.D-Zellen. Ein ähnlicher Trend ist in AX2 und AX2 [GFP-CbfA<sup>724-998</sup>] zu beobachten, wobei beide Änderungen aufgrund der hohen Standardabweichung nicht signifikant sind. Diese Schwankungen lassen sich durch eine schwierige Quantifizierung der sehr geringen Expression der Entwicklungsmarker zum Zeitpunkt null erklären.



**Abbildung 20: qRT-PCR 8 Stunden hungernde Zellen gegen nicht hungernde Zellen;** angegeben sind Mittelwert und Standardabweichung aus drei Experimenten



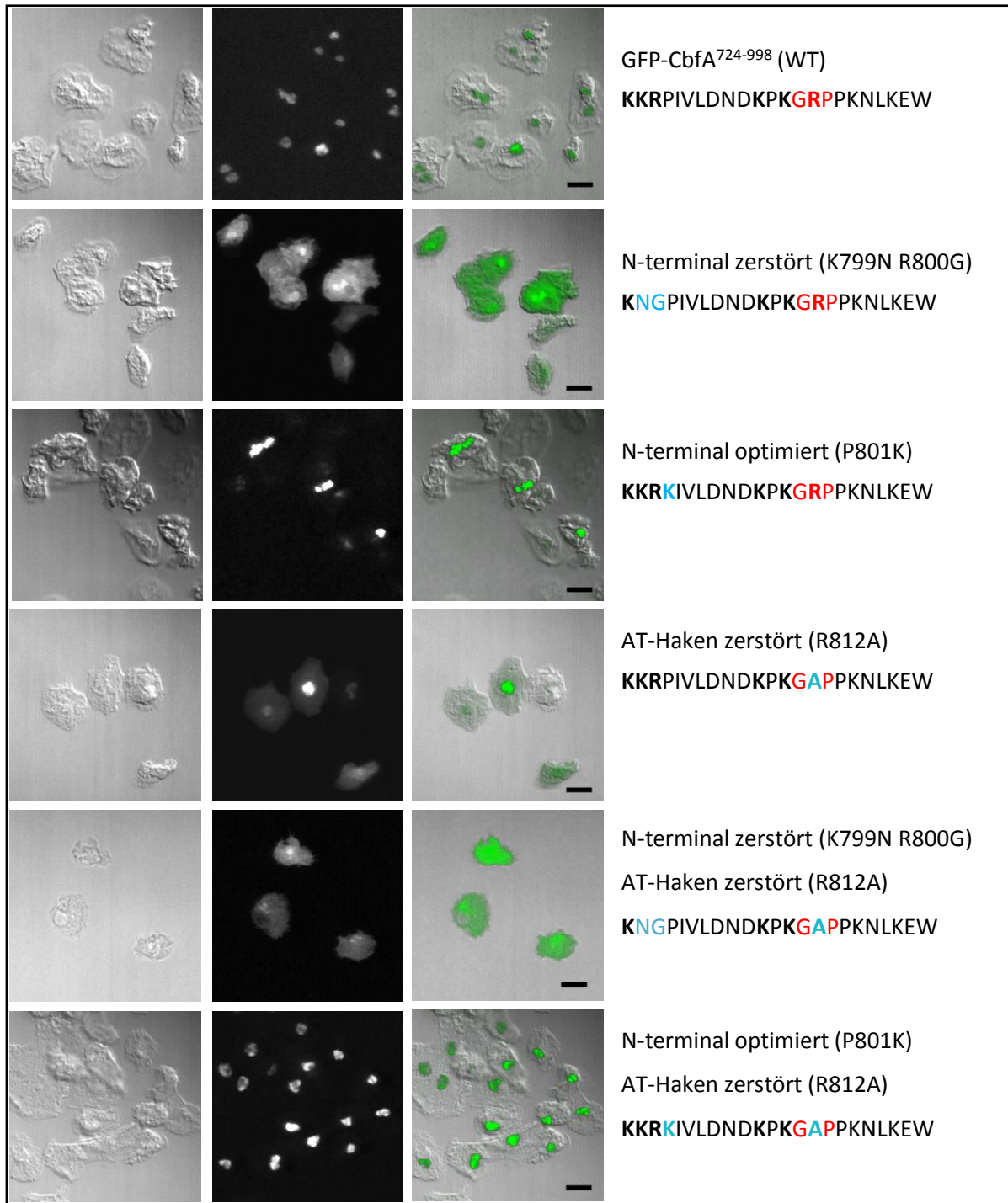
### 3.9 Kernlokalisierung von GFP-CbfA<sup>724-998</sup> und AT-Haken-Mutanten

Die in Kapitel 3.3 und 3.7 dargestellten Ergebnisse lassen eine autonome regulatorische Funktion der CbfA-CTD vermuten. Die Daten legen nahe, dass sowohl CTD-reprimierte als auch CTD-aktivierte Gene existieren. In der einzigartigen Carboxy-terminalen Domäne lassen sich aufgrund ihrer Sequenz zwei Strukturelemente ausfindig machen. Dies ist zum einen die zweigeteilte Kernlokalisierungssequenz K<sub>798</sub>KRPIVDNDKPKGR<sub>812</sub> (NLS unterstrichen) sowie ein sogenannter AT-Haken mit seinem zentralen GR<sub>812</sub>P-Motiv. Dieses Element kann adenin- und thymidinreiche Sequenzen aufgrund ihrer Kompaktheit binden (Arvind und Landsman 1998).

Es ist zu erwähnen, dass CbfA eine zweite potentielle NLS besitzt (K<sub>488</sub>KIKCHRCEKRFKKF<sub>502</sub>), welche sich im Bereich des zweiten Zinkfingers befindet. Ziel der nachfolgend beschriebenen Arbeiten ist es, durch gezielten Austausch funktionell wichtiger Aminosäuren die Beteiligung des AT-Hakens an der CTD-vermittelten Regulation zu belegen. Da sich zentrale Elemente dieser Struktur mit der Kernlokalisierung überschneiden, müssen diese Themen zusammenhängend betrachtet werden. Alle durch Einzelstrangmutagenese manipulierten Plasmide sind in den Stamm JH.D transformiert worden. Erhaltene Klone wurden auf vorhandene Fluoreszenz, sowie auf die Expression des Fusionsproteins GFP-CbfA<sup>724-998</sup> mittels *Western-Blot* überprüft. Die Fluoreszenzverteilung fixierter Zellen repräsentativer Klone wurde mittels konfokaler Laserscanning-Mikroskopie dokumentiert. Dies erfolgte in Zusammenarbeit mit der Abteilung Molekulare und Angewandte Mikrobiologie des Leibniz-Institut für Naturstoff-Forschung und Infektionsbiologie – Hans-Knöll-Institut (Jena).

Die Lokalisation der GFP-markierten CTD sowie ihrer Mutanten ist in Abbildung 21 dargestellt. Es wurden als erstes die Komponenten der NLS betrachtet. Durch die Umwandlung der Aminosäuren 799 bis 800 von Lysin und Arginin zu Asparagin und Glycin wird ein großer Anteil des Fusionsproteines nicht mehr in den Kern transloziert. Die funktionell entgegengesetzte Mutation Prolin 801 zu Arginin hat im Vergleich zum Wildtyp keine sichtbare Auswirkung auf die Verteilung. Das entstandene KKRK-Motiv entspricht einer „klassischen“ NLS (u.a. Cristophe *et al.* 2000). Die Umwandlung von Arginin 812 zu Alanin führt zu einer deutlichen Erhöhung des cytoplasmatischen Anteils von GFP-CTD, welcher jedoch geringer war als der von K799N R800G.

## Ergebnisse

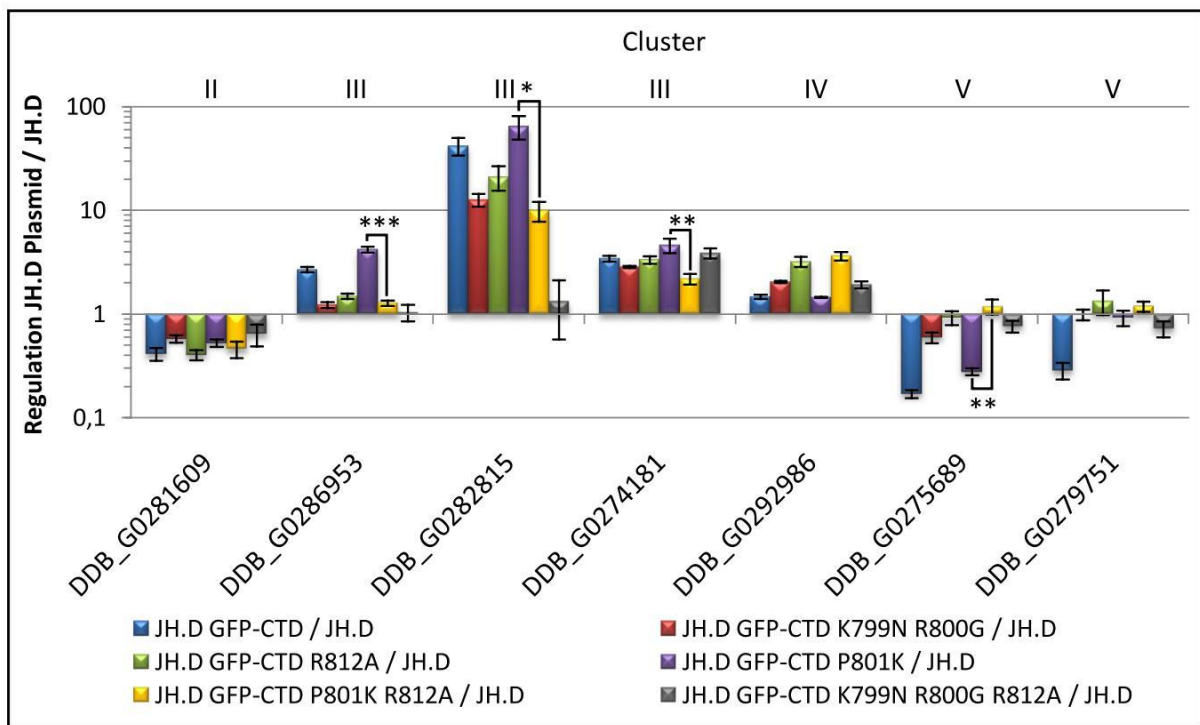


**Abbildung 21: Kernlokalisierung der GFP-Cbfa<sup>724-998</sup> in JH.D;** von links nach rechts: Durchlichtbild, Farbauszug der GFP-Fluoreszenz, Überlagerung beider; Größenstandart ist 10µm; Sequenz entspricht der AS 798-820 von Cbfa; Teile der NLS fett, AT-Haken in rot, geänderte Aminosäuren in blau

Diese Veränderung ist offenbar bestimmend für die Funktionsweise des AT-Hakens. Deshalb wurden sowohl cytoplasmatische (K799N R800G R812A) als auch vollkommen kernlokalisierte (P801K R812A) CTD-Versionen getestet.

### 3.10 Genregulation durch GFP-CbfA<sup>724-998</sup> und AT-Haken-Mutanten

Zur Ermittlung der regulatorischen Aktivität der unter 3.9 beschriebenen Mutanten sind qRT-PCR-Messungen der für die Cluster repräsentativen Gene durchgeführt wurden. Jede Messung wurde mit mindestens drei unabhängigen RNA-Isolationen eines Klonen durchgeführt und mit mindestens einem weiteren Klon verifiziert. Alle Berechnungen beziehen sich auf JH.D als Kontrolle. In Abbildung 22 sind zu Gunsten der Übersicht nur die Werte der CTD-abhängigen Cluster II-V dargestellt. Die CTD-unabhängigen Cluster I und VI zeigen in allen untersuchten Mutationen ähnliche Expression wie in Abhängigkeit von GFP-CbfA<sup>724-998</sup>.



**Abbildung 22: Einfluss von Mutationen im Bereich des AT-Hakens auf die Regulation durch CbfA<sup>724-998</sup>;** dargestellt sind Mittelwert und Standardabweichung von mindestens 3 Messungen; Signifikanz im Vergleich P801K zu P801K R812A: \*p < 0,05 \*\*p < 0,005 \*\*\*p < 0,0005 (t-Test)

Eine Reduzierung der Regulation durch die Mutationen K799N R800G ( $\pm$ R812A) ist durch die verminderte Translokation in den Kern (vgl. Abbildung 21) zu erklären. Die Mutante R812A

zeigte zumindest für die Cluster III und V eine stark verringerte Fähigkeit zur Regulation. Für P801K wurde eine dem Wildtyp entsprechende Funktionsweise erwartet, welche sich im Cluster II, III und für ein Cluster V-Gen (DDB\_G0275689) erkennen lässt. Das zweite Cluster V-Gen wird in beiden untersuchten Klonen nicht reguliert.

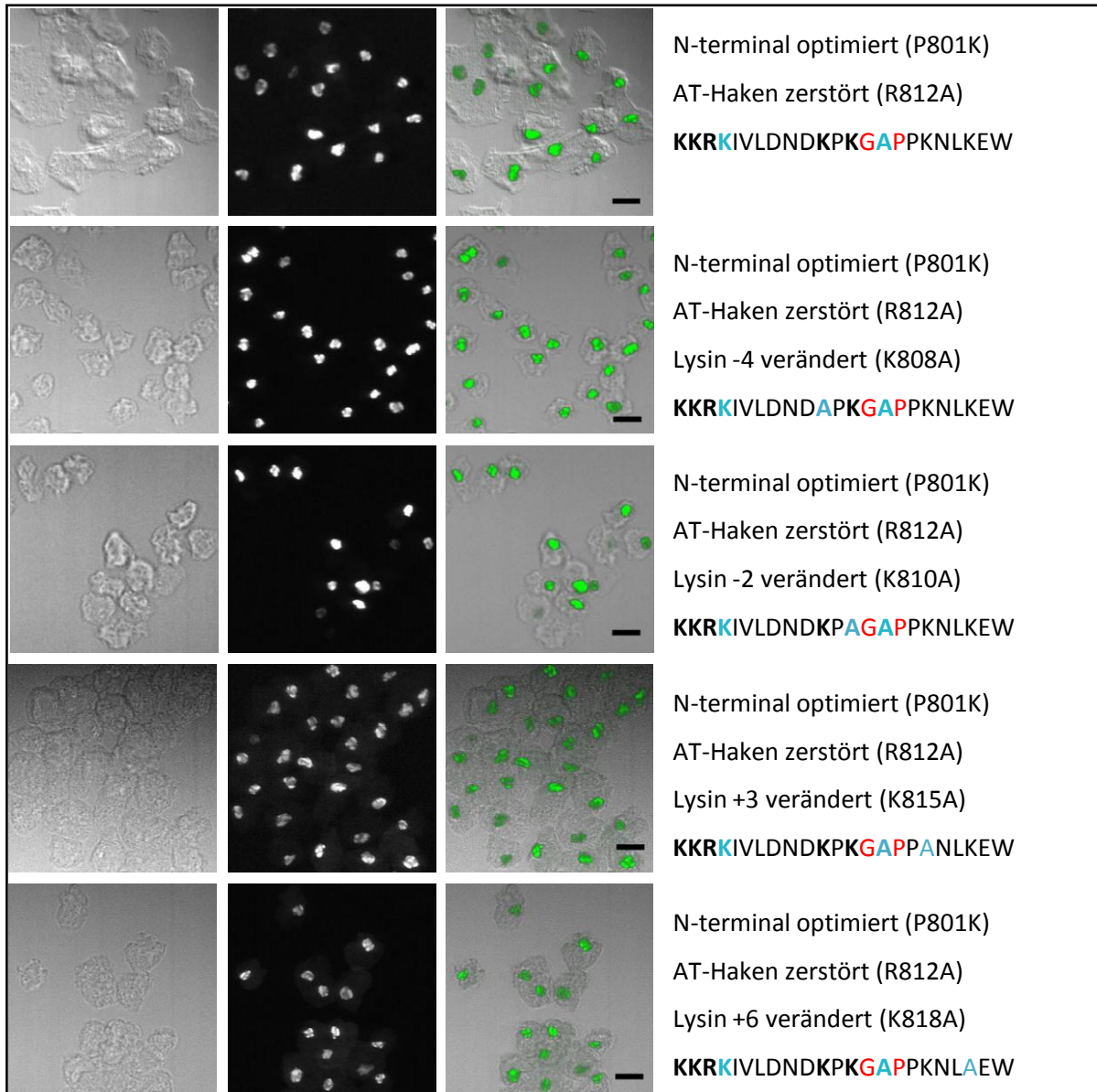
Die für die Beteiligung des AT-Hakens aussagekräftigste Mutante ist das kerngängige P801K R812A-Protein. Der Faktor Kernlokalisierung ist hier im Vergleich zur Mutante P801K vernachlässigbar. Für alle Cluster III-Gene kann mittels eines t-Testes ein statistisch signifikanter Unterschied festgestellt werden. Dies gilt ebenfalls für das durch die P801K-Mutante regulierte Cluster V-Gen. Die Stärke der Signifikanz ist in Abbildung 22 durch Sterne dargestellt. Interessant ist die Tatsache, dass die Regulation zweier Cluster III-Gene nicht auf Werte um eins sinkt. Eine Restaktivität der Mutante scheint noch vorhanden zu sein.

Das getestete Cluster II-Gen zeigt keine signifikanten regulatorische Effekte auf die durchgeführten Manipulationen. Das Cluster IV-Gen verhält sich bis auf P801K entgegen der Erwartungen. Die Mutation des zentralen Arginin 812 resultiert in einer erhöhten Regulation des untersuchten Gens. Zu erwarten wäre hier eine Annäherung an eine differentielle Expression von eins. Allgemein ist aber die Funktion des nicht mutierten GFP-CbfA<sup>724-998</sup> bereits sehr gering, was eine Diskussion dieses Ergebnisses erschwert.

### **3.11 Regulatorischer Einfluss von Lysinen im Bereich des AT-Hakens**

Eine Sequenzanalyse bekannter Typ-III AT-Haken zeigt die hohe Konservierung von Argininen und Lysinen im direkten Umfeld des AT-Hakens (Arvind und Landsman 1998). Der Einfluss dieser basischen Aminosäuren wurde in den strukturell etwas anders aufgebauten Typ-II AT-Hakenproteinen der HMGA-Familie bereits etwas näher untersucht (Huth *et al.* 1997, Reeves *et al.* 2001 a, Reeves 2001 b) und lässt vermuten, dass sie an der DNA-Bindung beteiligt sind. Um dies für die Carboxy-terminale Domäne von CbfA zu überprüfen, wurden jeweils die zwei dem Kernmotiv Glycin Arginin Prolin des AT-Hakens nächst gelegenen Lysine ausgetauscht. Dies erfolgte an dem Konstrukt, welches bereits die Mutationen P801K R812A in [GFP-CbfA<sup>724-998</sup>] aufwies. Die Lokalisation der entstandenen GFP-CbfA<sup>724-998</sup>-Mutanten in den Transformanten von JH.D ist in Abbildung 23 erkenntlich. Unterschiede in der Zellmorphologie sind auf eine

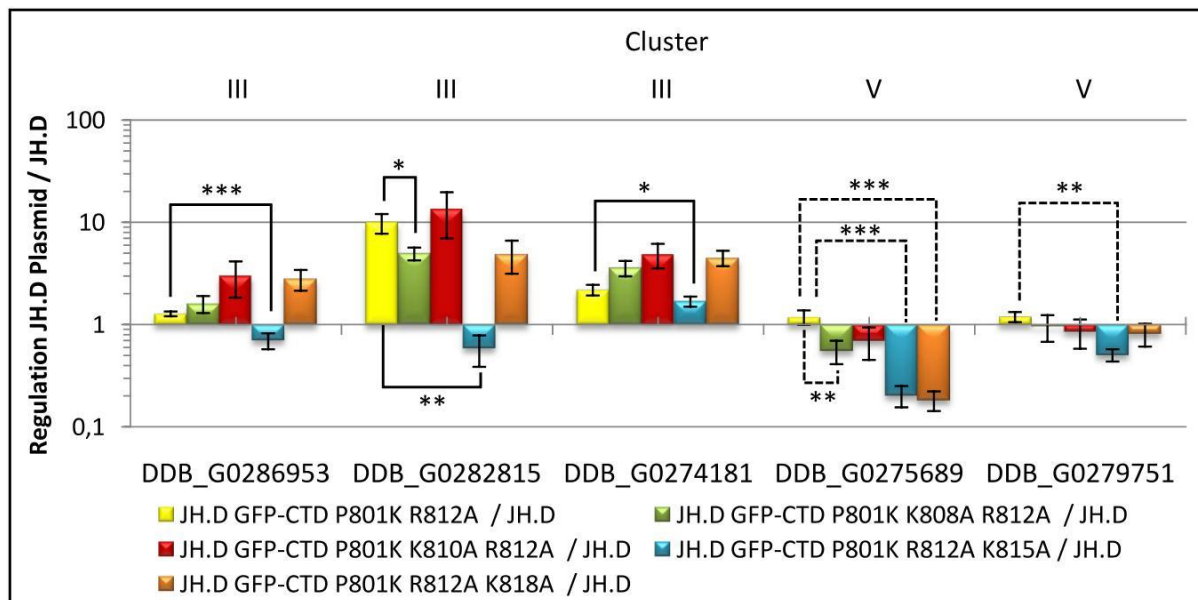
unzureichende Fixierung der später erstellten Stämme zurückzuführen, in Kultur erscheinen alle Stämme gleich. Die Fluoreszenzverteilung unterscheidet sich für die N-terminalen Lysine nicht von der kerngängigen AT-Haken-Mutante. Die beiden C-terminalen Veränderungen erhöhen den cytoplasmatischen Anteil der Fluoreszenz minimal.



**Abbildung 23: Kernlokalisierung von AT-Haken-Mutanten in JH.D;** von links nach rechts: Durchlichtbild, Farbauszug der GFP-Fluoreszenz, Überlagerung beider; Größenstandard ist 10µm; Sequenz entspricht der AS 798-820 von CbfA; **Teile der NLS fett**, **AT-Haken in rot**, **geänderte Aminosäuren in blau**; Position der Lysine bezieht sich auf das ursprüngliche Arginin 812 des AT-Hakens

## Ergebnisse

In Abbildung 24 ist die regulatorische Kapazität der untersuchten Transformanten dargestellt. Experimente wurden wie in 3.10 beschrieben durchgeführt und auf den Ausgangsstamm JH.D bezogen. Da für Cluster II und IV aufgrund des vorherigen Experimentes keine schlüssigen Aussagen mehr möglich sind, werden nachfolgend nur die Cluster III und V betrachtet. Für Cluster III ist bei allen drei untersuchten repräsentativen Genen eine signifikante Reduzierung der Aktivität durch den Austausch des Lysins 815 (+3) erkennbar. Für das Gen DDB\_G0282815 ist ferner die Mutation des Lysins 808 (-4) statistisch signifikant von Einfluss, die Mutation Lysin 818 (+6) mit  $p = 0,06$  zeigt eine deutliche Tendenz. Entgegen der Erwartung steigen für die beiden anderen Cluster III-Gene die Regulationen der Mutante P801K R812A K818A im Vergleich zu P801K R812A wieder an. Die Cluster V-Gene verhalten sich völlig anders. DDB\_G0279751 zeigt ebenfalls eine Reaktion auf die Mutation des Lysins 815 (+3), welche zu einer leichten Verringerung der Regulation führt. Da aber bereits die Mutation P801K R812A die reprimierende Wirkung von GFP-CTD aufhebt, sind Rückschlüsse auf die Funktion der Lysine in diesem Fall schwer zu treffen. Das Gen DDB\_G0275689 zeigt diesen Effekt für alle Lysine außer K810 (-2), wobei die reprimierende Wirkung von GFP-CbfA<sup>724-998</sup> durch die Veränderung der Lysine 815 (+3) und 818 (+6) fast vollständig wieder erreicht wird.

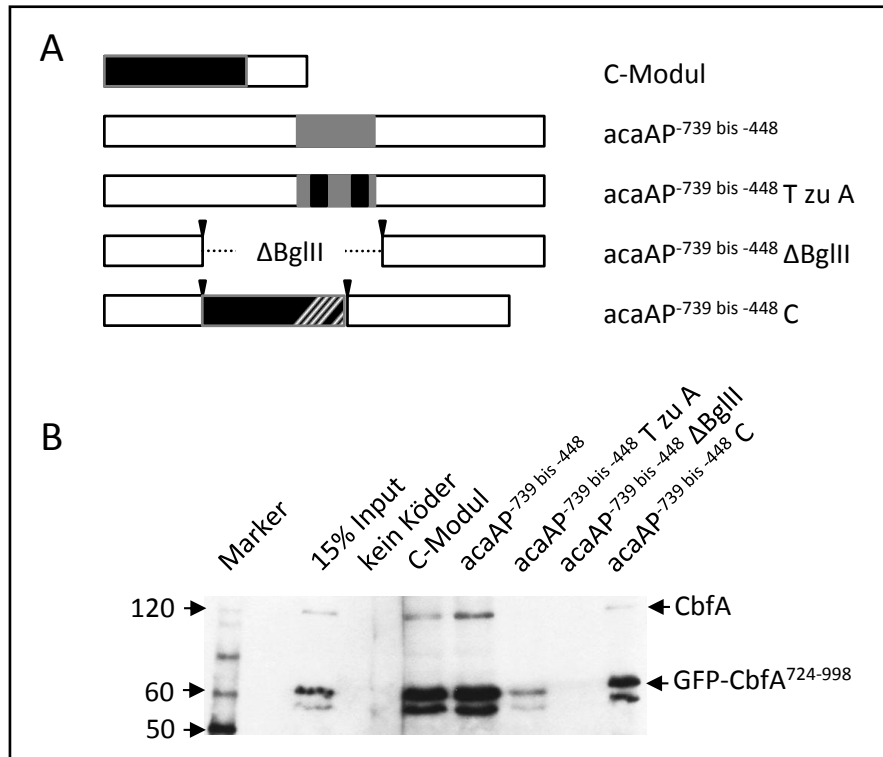


**Abbildung 24: regulatorischer Einfluss von Lysinen im Bereich des AT-Hakens;** dargestellt sind Mittelwert und Standardabweichung von mindestens 3 Messungen; Signifikanz im Vergleich P801K R812A zu entsprechenden Lysinmutanten: \* $p < 0,05$  \*\* $p < 0,005$  \*\*\* $p < 0,0005$  (t-Test)

### 3.12 Bindung der AT-Hakenmutanten an den Modellpromotor von *acaA*

Der beobachtete Einfluss des Lysins 815 sollte in einem weiteren Experiment gezeigt werden. Dazu wurden DNA-Protein-Bindungstests durchgeführt. Eine Interaktion von CbfA mit dem Bereich -739 bis -448 des Genes der Adenylylcyclase A (*acaAP*<sup>-739 bis -448</sup>) ist in der Literatur beschrieben (Siol *et al.* 2006a). Obwohl die CbfA-CTD nicht ausreichend ist, um *acaA* zu aktivieren (vgl. Kapitel 3.8 insbesondere Abbildung 20), zeigen Versuche eine spezifische Bindung der Proteindomäne an ein biotinyliertes Promotorfragment. Diese Bindung wurde mit bekannten Modifikationen des Promotors von *acaA* verifiziert. Des Weiteren wurde die Bindung des C-Moduls von TRE5-A untersucht, da hier zumindest eine Beteiligung des AT-Hakens der CbfA-CTD an der Bindung dieses Elementes bekannt ist (Horn *et al.* 1999).

Die verwendeten Promotorfragmente und ihre Modifikationen sind schematisch in der Abbildung 25 A skizziert. Ein repräsentativer *Western-Blot* eines durchgeführten Bindungsnachweises ist in Abbildung 25 B dargestellt. Aufgrund der in JH.D gebildeten Menge an CbfA kann das Bindungsverhalten der CTD direkt mit dem gesamten Protein verglichen werden. Das Muster der Bindung an die biotinylierten PCR-Produkte ist ähnlich, die CTD bindet sowohl an das C-Modul als auch an *acaAP*<sup>-739 bis -448</sup>. Der mit dem Interaktionsbereich des C-Moduls modifizierte *acaAP*<sup>-739 bis -448</sup> wird ebenfalls gebunden. Entfernt man den Bindebereich ersatzlos, kann keine Interaktion mehr nachgewiesen werden. Der Unterschied von CbfA und seiner Carboxy-terminalen Domäne in der Interaktion mit dem *acaAP*<sup>-739 bis -448</sup>, bei dem im identifizierten Bindebereich jeweils zweimal zehn Thyminidine durch Adenosine ersetzt werden, ist auf ein allgemein stärkeres Signal der GFP-CTD im Vergleich zum Gesamt-CbfA zurückzuführen. Die Bindung an dieses Konstrukt ist zwar gering, wiederholt sich aber in allen Experimenten. Es ist anzumerken, dass die gezeigten Intensitäten nicht absolut vergleichbar sind, da sich das Protein und seine Domäne in der Masse unterscheiden und GFP-CTD erheblich mehr als CbfA in JH.D gebildet wird.

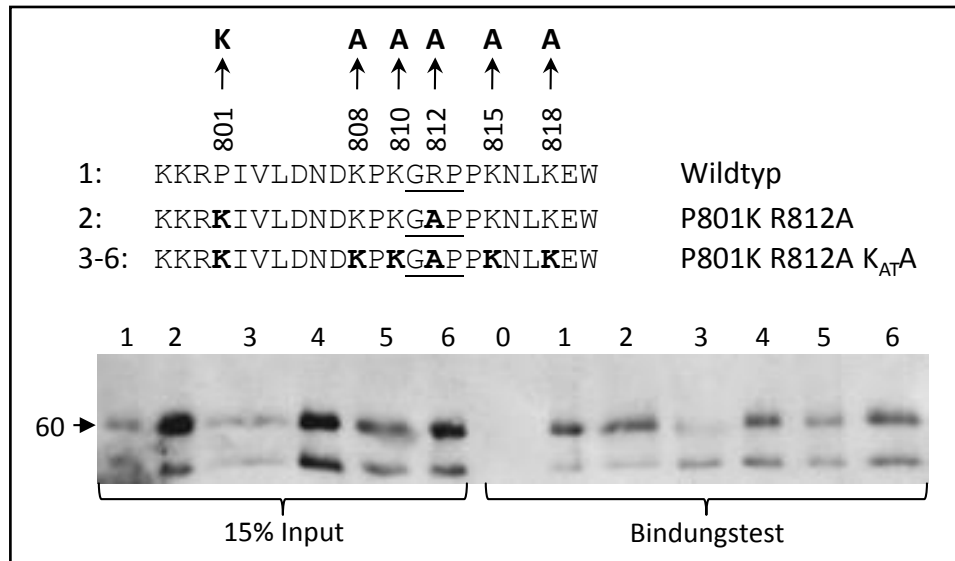


**Abbildung 25: Aufreinigung von CbfA-Fragmenten mit Hilfe spezifischer DNA-Proben; (A)** schematische Darstellung der benutzten DNA-Fragmente: C-Modul ist 134 bp-langer Bereich mit Promotoraktivität in der 3'-untranslatierten Region von TRE5-A, schraffierter Bereich enthält die bekannten CbfA-Bindemotive CMBS-1 und 2 (Geier *et al.* 1996, Horn *et al.* 1999); *acaAP*<sup>-739 bis -448</sup> mit den zwei homopolymeren Thymidinbereichen in grau; *acaAP*<sup>-739 bis -448</sup> T zu A mit Unterbrechung der Thymidinbereiche durch jeweils zehn Adenosine in schwarz; *acaAP*<sup>-739 bis -448</sup> ΔBglIII mit entfernten putativen CbfA-Bindebereich; *acaAP*<sup>-739 bis -448</sup> C Bindebereich mit dem entsprechenden Bereich des C-Moduls ersetzt (Siol *et al.* 2006a) **(B)** Verifizierung der Bindung von CbfA und GFP-CbfA<sup>724-998</sup> an bekannte DNA-Proben; Molekülmasse in kDa

Um die Fähigkeit zur Interaktion der AT-Hakenvariationen mit der DNA zu testen, sind Kernextrakte präpariert worden. Dazu wurden mit GFP-CbfA<sup>724-998</sup> und dessen Mutationen versehene Transformanten in der JH.D genutzt. Die erhaltenen Proteingemische sind auf gleiche Signalstärken der GFP-CTD mit Hilfe eines *Western-blots* mit dem Antikörper 7F3 eingestellt worden. Anschließend wurde mit der biotinylierten *acaAP*<sup>-739 bis -448</sup>-Probe inkubiert und die gebildeten Komplexe über ihre Affinität zu Streptavidin gereinigt. Der *Western-Blot* gegen die CTD von CbfA einer dieser Reinigungen ist in Abbildung 26 zu sehen. Die schwache Bande in der Spur 3 des Inputs ist auf das mit Abstand größte aufgetragene Volumen zurückzuführen. Dadurch kam es zu einer Verbreiterung der Spur dieser Probe im Gel, was die optisch geringe Bandenintensität erklärt. Bis auf die Mutante P801K K808A R812A ist kein Unterschied in der



Bindung an *acaAP*<sup>-739 bis -448</sup> zu erkennen. Eine Wiederholung des Versuches mit neuen Kernextrakten bestätigte dieses Ergebnis nicht, hierbei war die mit intakten Lysinen versehene P801K R812A sogar etwas schwächer. Durch Änderungen der Versuchsparameter, wie zum Beispiel der Salzkonzentration, konnte keine graduelle Bindung der Mutanten festgestellt werden. Ein Verlust der Bindung ging immer auch mit einem Verlust der als positiv identifizierten Interaktion CTD mit *acaAP*<sup>-739 bis -448</sup> einher (vgl. Abbildung 25 B).



**Abbildung 26: Aufreinigung von GFP-CbfA<sup>724-998</sup> und dessen Mutanten mit *acaAP*<sup>-739 bis -448</sup>**; dargestellt sind die AS 798-820 von CbfA, Spur 0 = keine biotinylierte DNA eingesetzt, K<sub>AT</sub>A = jeweils ein Lysin im Bereich des AT-Hakens durch Alanin ersetzt, zentrales AT-Haken-Motiv GRP unterstrichen, **geänderte AS fett**, Molekülmasse in kDa

### 3.13 Konstruktion von Hefe-2-Hybrid-Banken

Die in dieser Arbeit präsentierten Ergebnisse lassen eine Reihe von Möglichkeiten für die Funktion der Carboxy-terminalen Domäne vermuten. In einem biologischen System ist dabei die Betrachtung eines Faktors allein meist nicht sehr aussagekräftig. Oftmals wird die Funktion durch Protein-Protein-Interaktionen entscheidend bestimmt. Zur Identifizierung und weiteren Untersuchung dieser Interaktionen wird häufig das Hefe-2-Hybrid-System genutzt, wobei für die Suche gute Bibliotheken (Banken) geeigneter Fusionsproteine vorhanden sein müssen. Eine Recherche nach brauchbaren Banken für *Dictyostelium discoideum* war erfolglos. Aus diesem Grund wurden mit Hilfe des **Matchmaker™ Library Construction & Screening System** (Clontech)

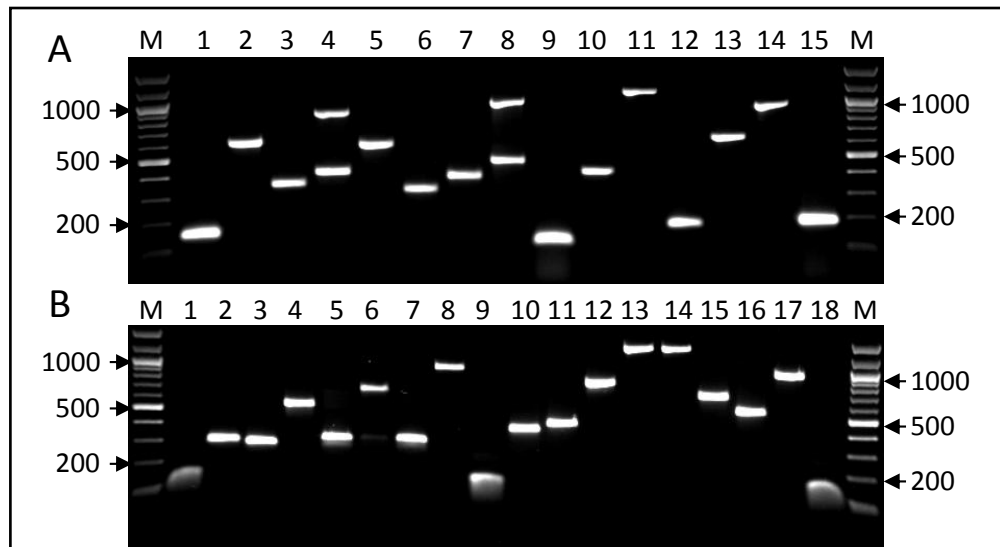
zwei komplexe Banken konstruiert. Vorteil dabei ist das Umgehen des aufgrund des hohen AT-Gehaltes oft unzureichend funktionierenden, bakterienabhängigen Klonierens. Beide Banken sind aus der mRNA vegetativ wachsender AX2 erstellt worden und unterschieden sich in der Art des Primers der Erststrangsynthese. Ein am 3'-Ende mit 30 dT versehener Primer – oligo (dT) – bindet an das polyadenylierte 3'-Ende einer mRNA, während ein mit sechs zufälligen Nukleotiden endender Primer – (dN)<sub>6</sub> – an jeder Stelle einer mRNA binden kann. Die durch die Erststrangsynthese entstandene, einzelsträngige cDNA wird mit einer speziellen Form der PCR in eine doppelsträngige cDNA umgewandelt und amplifiziert. Die dabei entstandene DNA ist aufgrund der benutzten Primer durch bekannte Sequenzen flankiert. Diese Sequenzen ermöglichen eine homologe Rekombination des PCR-Produktes mit einem entsprechenden Vektor in der Hefezelle. Da dieser Vektor linearisiert ist, kann es nur durch erfolgreiche Rekombination zu einer Expression des vektorcodierten Auxotrophiegenes aus dem rekombinierten Plasmid kommen. Beide Banken wurden abweichend von der Anleitung des Herstellers in den Hefestamm Y187 konstruiert und vermehrt. Dieser Stamm ist aufgrund seines Matingtyps (*MATα*) für ein breiteres Spektrum an für die Interaktionssuche benötigten, im Labor vorhandenen Matingstämmen des Typs *MATα* geeignet. Ferner zeigt Y187 eine höhere Transformationseffizienz als AH109, was die Komplexität der Banken erhöht.

**Tabelle 7: Eigenschaften der erstellten Hefe-2-Hybridbanken;** \*ergibt sich aus den sequenzierten Klonen einer Verdünnungsreihe (Liste im Anhang F) abzüglich der flankierenden Primersequenzen

Bank	oligo (dT)	(dN) <sub>6</sub>
Ausgangsmaterial	mRNA vegetativ wachsender AX2	mRNA vegetativ wachsender AX2
Erststrangsynthese	CDS III-Primer mit oligo (dT)	CDS III/6-Primer mit (dN) <sub>6</sub>
Konstruktionsstamm	<i>Saccharomyces cerevisiae</i> Y187	<i>Saccharomyces cerevisiae</i> Y187
Anzahl Transformanten	2,1*10 <sup>6</sup> cfu	3*10 <sup>6</sup> cfu
Lebendzellzahl Aliquots	1*10 <sup>8</sup> cfu / ml	8*10 <sup>7</sup> cfu / ml
Klone mit Integration	11 von 15 (73%)	15 von 18 (83%)
∅ cDNA-Integration*	482bp (min. 232bp / max. 904bp)	416bp (min. 100bp / max. 1127bp)
∅ Entfernung vom 3'-Ende*	80bp (min. 0bp / max. 504bp)	680bp (min. 0bp / max. 1755bp)

Nach der Rekombination wurden alle entstandenen Klone vereinigt und zu 1-1,5 ml aliquotiert. Für die durchgeführte Suche nach Interaktionen ist jeweils eins dieser Aliquots genutzt wurden. Eine Zusammenfassung über wichtige Daten der Hefe-2-Hybrid-Banken ist in Tabelle 7 gegeben.

Das für die Konstruktion von Banken sehr wichtige Verhältnis von Ereignissen mit spezifischer Integration in den Zielvektor zu Ereignissen ohne („leere“ Vektoren) wurde durch PCR mit anschließender Sequenzierung von aus Einzelklonen gewonnener Vektor-DNA abgeschätzt. Klone ohne Integrationsereignis resultieren in einem 180bp großen PCR-Fragment. Die jeweiligen Bilder der Gelelektrophoresen dieser PCRs sind in Abbildung 27 zu sehen. Der Anteil spezifischer Integrationen liegt bei 73 bzw. 83 % Eine Liste der sequenzierten Klone befindet sich im Anhang F.



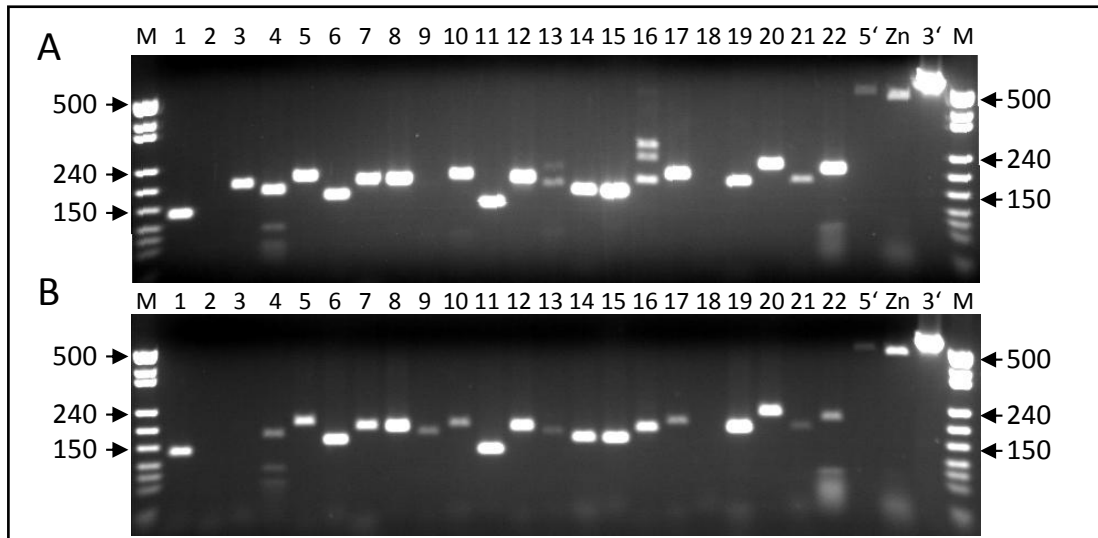
**Abbildung 27: Gelelektrophorese von PCRs an Einzelklonen der Banken; (A) oligo (dT)-Bank (B) (dN)<sub>6</sub>-Bank; „leerer“ Vektor ergibt 180bp; M = 100bp-Leiter von NEB; alle Größenangaben in bp**

Zwei untersuchte Klone der oligo (dT)-Bank enthalten jeweils zwei Bankplasmide, was aber nur einen leichten Mehraufwand bei einer etwaigen Interaktion genau dieser Klone verursacht. In der (dN)<sub>6</sub>-Bank sind, bedingt durch den verwendeten Primer, Restmengen ribosomaler RNA in der cDNA-Erstellung integriert. Des Weiteren enthält Klon 8 zwei Gene in einen Bankplasmid. Dies ist auf eine 12bp lange Mikrohomologie der beiden Gene zurückzuführen.

### 3.14 Charakterisierung der Hefe-2-Hybrid-Banken mittels PCR

Um die Komplexität der erstellten Banken bewerten zu können, wurde jeweils aus einem unter 3.13 beschriebenen Aliquot die Plasmid-DNA isoliert. Diese diente dann als Ausgangspunkt für eine Serie von PCR-Ansätzen. Zur Amplifizierung wurden im Labor vorhandene Kombinationen von qRT-PCR-Primern genutzt. Diese sind meist gegen das 3'-Ende eines Genes gerichtet. Des Weiteren wurden drei Primerkombinationen gegen verschiedene Bereiche des *cbfA*-Gens getestet. Dieses Gen eignet sich aus mehreren Gründen gut für die Abschätzung der Qualität der Banken. Das *cbfA*-Gen hat gemessen an der Anzahl der erfassten *Reads* im Stamm AX2 im Rahmen des Solexaexperimentes eine eher unterdurchschnittliche Expression (11.863 reads zu 35.020 im Durchschnitt). Ferner besitzt es mit einer cDNA-Länge von 3.003 bp eine vergleichsweise lange codierende Sequenz, der Durchschnitt beträgt 1.557 bp (Eichinger *et al.* 2005). Ein Nachweis des 5'-Endes von *cbfA* in den Banken deutet auf das Vorhandensein auch anderer, in der Methode der Konstruktion der cDNA-Bank bedingter, unterrepräsentierter Bereiche am 5'-Ende von Genen hin. In Abbildung 28 sind die Bilder der elektrophoretisch aufgetrennten PCR-Produkte zu sehen.

In der oligo (dT)-Bank konnten 15 von 22 getesteten Gene mit ihrer erwarteten Bandengröße nachgewiesen werden. Vier Gene zeigten mehrere Banden, wobei die stärkste jeweils der erwarteten Größe entspricht. Für die Gene DDB\_G0274289, DDB\_G0275689 und DDB\_G0282075 konnte auch durch Wiederholung der PCR kein sichtbares Fragment amplifiziert werden. Alle getesteten *cbfA*-Fragmente ließen sich in beiden Banken nachweisen. Die Intensität der Bande nahm mit steigender Entfernung zum 3'-Ende ab. Die (dN)<sub>6</sub>-Bank umfasste nachweislich 17 von 22 getestete Gene. Zwei weitere Gene ergaben mehrere Banden. Von den drei nicht nachweisbaren Genen unterschied sich diese Bank nur durch DDB\_G0272783 von der anderen Bank. Das in der oligo (dT)-Bank fehlende Gen DDB\_G0275689 war in der (dN)<sub>6</sub>-Bank vorhanden. Die nicht in den Hefe-Banken vertretenen Gene DDB\_G0274289 und DDB\_G0282075 zeigten in den Transkriptomanalysen (Kapitel 3.6) eine kleine Anzahl von *reads* sowohl in AX2 als auch in JH.D, was auf ein sehr geringes Maß an Expression hindeutet.



**Abbildung 28: Nachweis bekannter Fragmente in den Hefebanken; (A) oligo (dT)-Bank (B) (dN)<sub>6</sub>-Bank;** 1: DDB\_G0274181 (erwartetes Fragment 133 bp), 2: DDB\_G0274289 (128), 3: DDB\_G0272783 (191), 4: DDB\_G0291281 (173), 5: DDB\_G0277503 (206), 6: DDB\_G0275153 (156), 7: DDB\_G0267528 (193), 8: DDB\_G0292378 (189), 9: DDB\_G0275689 (175), 10: DDB\_G0287461 (197), 11: DDB\_G0277403 (130), 12: DDB\_G0284861 (189), 13: DDB\_G0279751 (175), 14: DDB\_G0292986 (160), 15: DDB\_G0282815 (157), 16: DDB\_G0279187 (185), 17: DDB\_G0291646 (204), 18: DDB\_G0282075 (170), 19: DDB\_G0289599 (184), 20: DDB\_G0272867 (232), 21: DDB\_G0287363 (187), 22: DDB\_G0284549 (213), 5': *cbfA* Nukleotid 130-617 (Teil der JmjC-codierenden Sequenz über Intron, 487 bp), Zn: *cbfA* Nukleotid 604-1071 (Bereich vor ersten Zinkfinger, 487 bp), 3': *cbfA* Nukleotid 2424-2974 (Teil der CTD-codierenden Sequenz am 3'-Ende, 550 bp); M = pUC19 DNA geschnitten mit *MspI* (Fermentas); alle Werte in der Abbildung in bp

### 3.15 Suche nach Protein-Protein-Interaktionen mittels Hefe-2-Hybrid

Zum Auffinden von Interaktionen wurden sogenannte Interaktionsmatings durchgeführt. Im Hefestamm Y187, welcher den Matingtyp *MAT $\alpha$*  besitzt, befindet sich eine Bibliothek von Fusionsproteinen aus *Dictyostelium*-cDNA codiertem Anteil und der aktivierenden Domäne des Transkriptionsfaktors GAL4. Dieser Stamm wird mit einem kompatiblen *MAT $\alpha$* -Stamm (AH109 oder KFY1), welcher mit dem zu untersuchenden Köderplasmid transformiert ist, zusammen inkubiert. Dabei entstehende Zygoten besitzen sowohl das Köderplasmid als auch mindestens ein Bankplasmid. Das Köderplasmid codiert für eine Fusion der DNA-bindenden Domäne des GAL4 und dem zu untersuchenden Proteinfragmentes. Kommt es nun zu einer Interaktion des Köder- und des Beuteproteins, wird die Funktion von GAL4 durch räumliche Nähe wieder hergestellt. Durch die Selektion auf die durch den *GAL4*-Promotor gesteuerten Reportergene können nur Zygoten mit einer durch die Plasmide vermittelten Protein-Protein-Interaktion wachsen.

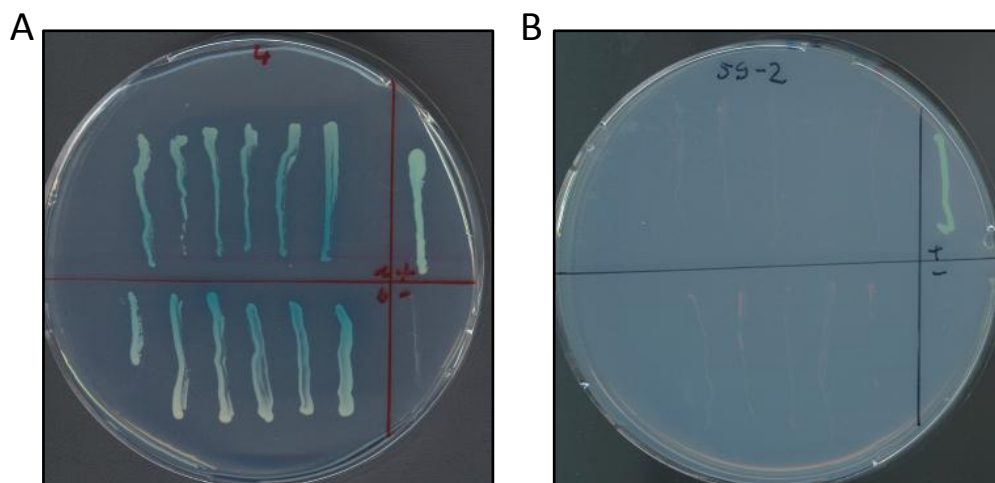
## Ergebnisse

Für die durchgeführten Untersuchungen ist CbfA<sup>793-1000</sup> als Köder benutzt worden. Die für andere Experimente benutzte CbfA<sup>724-998</sup> eignet sich für diese Methode weniger, da der enthaltene asparaginreiche Bereich 732-769 zu falsch positiven Ereignissen führen kann. Nach dem obligatorischen Test auf Selbstaktivierung, das heißt Ausschluss der Interaktion des Köders mit aktivierender Domäne des Beuteplasmides allein, sind sieben Interaktionsmatings durchgeführt worden. Dabei wurde dreimal die oligo (dT)-Bank genutzt, zweimal die (dN)<sub>6</sub> und zweimal eine Kombination aus beiden. Die dabei erreichte Effizienz der Zygotenbildung lag dabei immer deutlich über den geforderten 2%. Die Matings unterschieden sich durch Art und Vorbereitung des Köderstammes, Dauer und Durchführung des sexuellen Schrittes sowie in der Zusammensetzung der Selektionsplatten. Tabelle 8 gibt einen Überblick über die getesteten Parameter sowie die erzielten Ergebnisse.

**Tabelle 8: Übersicht über die durchgeführten Interaktionsmatings;** \*Effizienz bestimmt sich aus dem Verhältnis der Zygoten und den limitierenden Elternstamm (hier immer Y187); n.b. nicht bestimmt; <sup>§</sup>Selektion bezieht sich auf nicht im Medium enthaltene Auxotrophiemarker (H = Histidin, A = Adenin, L = Leucin = Selektion auf Bankplasmid, T = Tryptophan = Selektion auf Köderplasmid), 3AT = 3-Amino-1,2,4-triazol

Mating	Bank	Köderstamm	Zygoten	Effizienz*	Selektion <sup>§</sup>	Bemerkungen und Ergebnis
1	oligo (dT)	AH109	2,13*10 <sup>7</sup>	25%	HALT	5 Klone (4 Selbstaktivierer, eine Interaktion nicht wiederholbar)
2	oligo (dT)	KFY1	1,22*10 <sup>7</sup>	14%	HLT + 3AT	ein Klon (Selbstaktivierend)
3	oligo (dT)	KFY1	n.b.	n.b.	HLT + 3AT	sexueller Schritt auf Platten, 3 Klone (2 Selbstaktivierer, eine Interaktion nicht wiederholbar)
4	(dN) <sub>6</sub>	AH109	1,62*10 <sup>7</sup>	14%	HALT	9 Klone (8 Selbstaktivierer, eine Interaktion nicht wiederholbar)
5	(dN) <sub>6</sub>	AH109	3,59*10 <sup>7</sup>	24%	ALT, dann HALT	84 Klone auf ALT, 72 Klone auf HALT-Masterplatte (71 Selbstaktivierer, eine Interaktion nicht wiederholbar)
6	oligo (dT) + (dN) <sub>6</sub>	AH109	9,90*10 <sup>6</sup>	8%	HALT	sexueller Schritt auf Platten, keine MEL1-positiven Klone
7	oligo (dT) + (dN) <sub>6</sub>	AH109	1,56*10 <sup>7</sup>	11%	HALT	6 Klone (6 Selbstaktivierer)

Alle im Mating wachsenden Kolonien mit AH109-Hintergrund wurden einem Enzymtest auf  $\alpha$ -Galaktosidaseaktivität des ebenfalls unter dem *GAL4*-Promotor stehenden *MEL1*-Genproduktes unterzogen. Die aus Mating 2 und 3 hervorgegangenen Klone wurden auf  $\beta$ -Galaktosidaseaktivität durch *LacZ* überprüft. Ereignisse, welche zur Umwandlung des entsprechenden Galaktosides in der Lage waren, sind weiteren Tests unterzogen worden. Einerseits wurde nach Isolierung der Plasmid-DNA versucht, die Interaktion durch Transformation beider Plasmide in einen Stamm zu wiederholen, andererseits ist die Selbstaktivierung des Beuteplasmides mit der DNA-bindenden Domäne überprüft worden. Für alle Untersuchungen sind mindestens sechs Klone einer Transformation überprüft worden, bei nicht einheitlichen Ergebnissen wurden sechs weitere Klone einer neuen Transformation betrachtet. In Abbildung 29 sind typische Resultate der Retransformationen dargestellt.



**Abbildung 29: exemplarische Beispiele der Überprüfung der im Mating gefundenen Klone; (A)** Mating 4 Klon 4 (Selbstaktivierer) **(B)** Mating 5 Klon 59 (nicht wiederholbare Interaktion); oberer Teil Transformation von Köderplasmid (pGBKT7-CTD) mit isolierten Beuteplasmid (pGADT7-REC-M4#4 bzw. M5#59.2) in AH109; unterer Teil Transformation des leeren Köderplasmides (pGBKT7) mit isolierten Beuteplasmid; + entspricht positiver Kontrolle: Interaktion aus Mating 1 Klon 14 (starker Selbstaktivierer DDB\_G0286355); - ist negative Kontrolle: leeres Köderplasmid (pGBKT7) mit leerem Beuteplasmid (pGADT7) transformiert in AH109; alles auf vierfach selektiven Medium (minus HALT) plus X- $\alpha$ -Gal

Die Hefen im Teil A der Abbildung zeigen ein Wachstum auf vierfach selektiven Agarplatten sowie  $\alpha$ -Galaktosidaseaktivität nach der Transformation mit dem Beuteplasmid und dem leeren Köderplasmid. Die durch das Beuteplasmid codierte Proteinfusion interagiert mit der Bindedomäne des Köderplasmides. Es handelt sich um einen Selbstaktivierer. Für die sich wie in

Teil B der Abbildung verhaltenden potentiellen Interaktionen konnte kein Wachstum auf vierfach selektiven Platten durch Retransformation der Elternstämme AH109 und Y187 mit den isolierten Beute- und Köderplasmiden erzielt werden.

Wie aus Tabelle 8 hervorgeht, konnte keine spezifische Protein-Protein-Interaktion für CbfA<sup>793-1000</sup> beobachtet werden. Es sind ausgehend von der Zygotenanzahl über  $1 \cdot 10^8$  Ereignisse überprüft worden. Bis auf Mating 5 fällt die für dieses System geringe Anzahl von Klonen auf, was die hohe Stringenz der Ansätze belegt. Zwei mit gleich hohem Selektionsdruck von Heike Dölz durchgeführte Interaktionsmatings mit CbfA<sup>2-367</sup> als Köder ergaben jeweils über 140 deutlich positive Klone. Davon sind einige zumindest im Hefe-2-Hybrid-System nach allen geforderten Kontrollen als wahr zu betrachten. Durch die Verringerung der Stringenz im Mating 5 steigt lediglich die Anzahl der falsch positiven an. Der Test einer alternativen Selektionsmethode auf Histidin in Verbindung mit 3AT in dem adeninautotrophen Elternstamm KFY1 zeigte ähnliche Ergebnisse wie das Standardprotokoll. Die beschriebenen Selbstaktivierer wurden teilweise weiter analysiert. Dabei fiel eine Anhäufung von drei Konstrukten auf, welche in Tabelle 9 zusammengefasst ist.

**Tabelle 9: häufig auftretende Selbstaktivierer im Interaktionsmating mit CbfA<sup>793-1000</sup>; Größe der durch Verdau entstehenden Fragmente und der PCR-Produkte aus Sequenzierung abgeleitet**

Fusion mit	PCR Y2h5'/3'	Verdau <i>EcoRI/BamHI</i>	Identifizierbar durch	gefunden (bei Mating)
DDB_G0288123	974 bp	978 bp	Verdau mit <i>BglII</i> 8388 bp + 550 bp Verdau mit <i>NsiI</i> 6083 bp + 2855 bp	min. 8x (1, 4, 5, 7)
DDB_G0286355	589 bp	593 bp	Verdau mit <i>SpeI/NsiI</i> 5951 bp + 2987 bp	min. 12x (1, 2, 3, 4, 5)
DDB_G0275227	346 bp	350 bp	PCR mit pGAD426seq/Y2h-1#1 430bp	min. 28x (1, 3, 4, 5, 7)



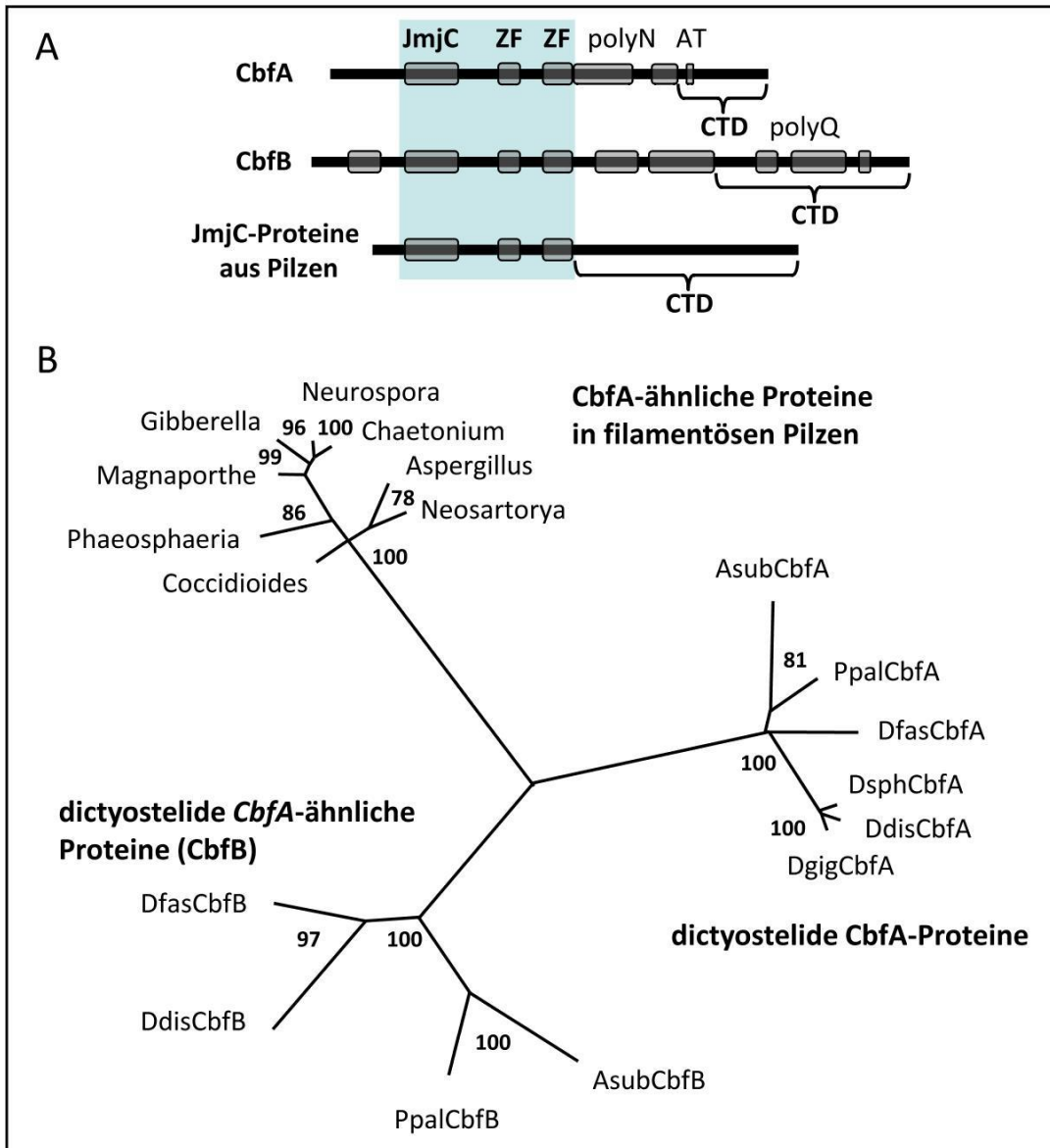
---

### 3.16 Betrachtungen zur Phylogenie von CbfA

Wie die in Abschnitt 3.6 dargestellten Ergebnisse zeigen, sind große Mengen an Sequenzinformationen mit Hilfe der neuen Technologien wesentlich schneller und vor allem preiswerter zu erlangen als noch vor wenigen Jahren. Dies eröffnet die Möglichkeit, nach potentiell orthologen Sequenzen zu *cbfA* in anderen Organismen zu suchen. Die phylogenetische Abstammung eines Genes ermöglicht oft Einblicke in seine vergangene und auch aktuelle Funktion. Um dieser Fragestellung Rechnung zu tragen, wurden verschiedene Ansätze mit dem Ziel der Identifizierung von CbfA-ähnlichen Proteinen durchgeführt.

Erster Kandidat ist das durch *cbfB* (DDB\_G0293470) codierte Genprodukt aus *D. discoideum*. Es handelt sich bei *cbfB* um ein paraloges Gen zu *cbfA*. CbfB besitzt wie CbfA eine JmjC-Domäne und zwei Zinkfinger. Die Identität auf Aminosäureebene beträgt gemessen an der Gesamtlänge von CbfA 30%. Die beiden Genprodukte unterscheiden sich im Amino-terminalen Teil nur wenig, der Bereich nach der asparaginreichen Domäne zeigt aber keinerlei Identität. Die Carboxy-terminalen Domänen sind vollkommen unterschiedlich. Die CTD von CbfA hat eine Länge von 230 AS und besitzt eine DNA-bindende Funktion durch den AT-Haken. Die CTD von CbfB ist nach der Asparagindomäne 485 AS lang und besitzt Strukturelemente wie weitere asparagin- und glutaminreiche Bereiche. Ein AT-Haken ist nicht detektierbar.

Weitere CbfA-ähnliche Proteine sind in der Familie der *Dictyosteliales*, zu den *D. discoideum* gehört, gefunden worden. Dazu sind zum Teil unveröffentlichte Daten von Genomprojekten dieser Organismen genutzt worden. Des Weiteren wurden im Rahmen der Diplomarbeiten von Markus Bohnert und Michael Sauer einige Genome sozialer Amöben mit degenerierten Primern nach *cbfA*-Orthologen Genen untersucht. Die dritte Ebene war die Suche nach ähnlichen Proteinen in Organismen außerhalb der *Dictyosteliales*. Die Abfragen der Datenbanken anhand der Primärsequenz der Aminosäuren erbrachten keine Ergebnisse. Aufgrund des Wissens über CbfB ist deshalb eine Suche nach den Strukturelementen JmjC-Domäne in Verbindung mit dem CbfA-typischen Zinkfinger durchgeführt worden. Es wurden im JmjC-Zinkfinger-Bereich orthologe Proteine in filamentösen Pilzen gefunden, welche nachfolgend als CbfA-ähnlich bezeichnet werden.



**Abbildung 30: phylogenetisches Verhältnis von CbfA-ähnlichen Proteinen anhand ihrer Aminosäuresequenz;** (A) Übersicht über den zur Erstellung des Baumes genutzten homologen Sequenzbereich (hellblauer Kasten); ZF = Zinkfinger, polyN = Asparagin-reiche Domäne, polyQ = Glutamin-reiche Domäne, AT = AT-Haken (B) der Baum wurde mit dem Programm **TREE-PUZZLE 4.0** (Schmidt *et al.* 2002) erstellt, Zahlen an den Knotenpunkten sind *Bootstrap*-Werte; identifizierte Gene der filamentösen Pilze: *Coccidioides immitis* CIMG\_04230 (GenBank-Nr. 119188365), *Phaeosphaeria nodorum* SNOG\_05303 (169604578), *Magnaporthe grisea* MGG\_01068 (39973571), *Gibberella zeae* FG10931.1 (46138831), *Neurospora crassa* NCU01554 (85078924), *Chaetonium globosum* CHGG\_08009 (116199707), *Aspergillus nidulans* AN4306.2 (67528200), *Neosartorya fischeri* JmjC domain protein (119501326); *Dictyosteliales*: *Acytostelium subglobosum* (Phyletische Gruppe nach Schaap *et al.* (2006): 2), *Polysphonylium pallidum* (2), *Dictyostelium fasciculatum* (1), *D. sphaerocephalum* (4), *D. discoideum* (4), *D. giganteum* (4)

Der phylogenetische Baum in Abbildung 30 zeigt die Verwandtschaft der Proteine mit ähnlicher Struktur wie CbfA. Es kann dabei in drei Klassen unterschieden werden. Diese sind orthologe Proteine zu CbfA in anderen Vertretern der *Dictyosteliade*, wobei hierbei die allgemeine Phylogenie der Gruppen dieser Familie wiedergegeben wird (vergleiche Schaap *et al.* 2006). Gerade innerhalb der Gruppe 4, zu den auch *D. discoideum* zählt, ist CbfA hoch konserviert. Der zweite Bereich des Baumes sind CbfA-ähnliche Proteine, welche eine höhere Sequenzidentität zu CbfB aufweisen. Die Sequenz außerhalb der JmjC-Domäne und der Zinkfinger zeigt keinerlei Ähnlichkeit zu CbfA. Von beiden Gruppen phylogenetisch gleich weit entfernt befindet sich die Gruppe CbfA-ähnlicher Proteine aus dem Phylum der *Ascomycota*. Auch hier beschränken sich die Sequenzhomologien auf den Bereich der JmjC-Domäne und der Zinkfinger.

### **3.17 Aufbau der CTD in CbfA-orthologen Proteinen**

Die aus der Suche nach CbfA-verwandten Proteinen (Kapitel 3.17) hervorgegangenen Sequenzen erlauben zusätzlich einige Aussagen über die Verbreitung und Struktur der Carboxy-terminalen Domäne in den untersuchten Organismen. Sequenzhomologien zu der CTD von CbfA aus *D. discoideum* konnten nur in potentiell orthologen Genen anderer *Dictyosteliales* gefunden werden. Dabei zeigten Vertreter der Gruppe 4 auf Grund ihrer engeren phylogenetischen Verwandtschaft zu *D. discoideum* die höchste Ähnlichkeit (Abbildung 31). Es lässt sich in dieser Gruppe eine Konservierung eines Bereiches, welcher reich an Serin und Threonin ist, erkennen. Diese Aminosäuren sind oftmals Ziel von Phosphorylierungen, was eine Möglichkeit von post-transkriptionellen Modifikationen mit Auswirkungen auf die Funktion des Faktors bietet, wie z.B. für die Serin-Threonin-reiche Domäne der RAP55-Proteinfamilie (Marnef *et al.* 2009). Außerhalb der Gruppe 4 ist die Struktur der CTD variabler. Der Bereich, welcher der *D. discoideum*-CTD zugeordnet werden kann, zeigt mehrere hoch konservierte Abschnitte, die sich mit variablen Bereichen abwechseln. Zu den erhaltenen Anschnitten zählt der AT-Haken in Verbindung mit der Kernlokalisierungssequenz. Zwei weitere, räumlich getrennte, konservierte Bereiche könnten für die reprimierende und aktivierende Funktion der CTD verantwortlich sein, was für eine einheitliche Funktionsweise in den unterschiedlichen Taxa spricht. Interessant ist, dass am Ende der Gruppe 4 CbfAs sich in den anderen Organismen noch weitere Aminosäuren anschließen, welche artspezifisch sind.

DdisCbfa	1	TTSSLSSLSTSLSSSS-TSTQKPMDVHS	KKRPIVLDNDKP	KGRPPKNLKEWTS	-THKFII
DpurCbfa	1	TTSTLTTLSTLSSSAPTSTKPMDIHS	KKRPIVLDNDKP	KGRPPKNLKEWTS	-THKFII
DfasCbfa	1	VHKPLGRPRKVVTAEAISSDTNIK	--EGKKRPIQMNNDKP	KGRPPKNLKEWTS	-THKFFI
PpalCbfa	1	GINHHYHSKHQHGGKSSIESDVQLES	KKRPIQY-NEKP	KGRPAKNLKDWD	-SHKFIL
AsubCbfa	1	PGRPKGSLNHSNPNNAATTPPNS	-TEKKRPLTESTTSN	KGRPPKILKET	SDTHKSVC
DdisCbfa	59	SLIELFRSSNNAILGKPNPHYKPIENL	PPLVQLYLSQRKAF	GGVLWAKTNSCPLL	PCIWV
DpurCbfa	60	SLVELNRS--NSVLGKNNPNYKRFEDL	PTLVQLYLTKRRAF	GGVLWAKTNSCPLL	PCLWV
DfasCbfa	58	NLIESSRK----GTPYHSLHINTHEVLP	SLISLYIDRKEKF	GGVMWAKTINHPLL	PCLWV
PpalCbfa	59	GLIEIYRN---GSKYSSTKALPVDTAV	PELVQLYLENRECF	GGLLWAKTINFPLL	PCLWV
AsubCbfa	60	AILDALRS----HIPYPHTAVTAPDPL	PTAVRLYVNHKDKL	GGLLWAKTVSCPW	MPCVWV
DdisCbfa	119	KDLSVIPNTKLLPSLIQGKKIVVLF	FFG-DQDQEEYVGI	VGKKSIFS	FDEVNQTL
DpurCbfa	118	KDLSIIPPNTKLLPSLVQGKKIVVLF	FFGCNIEDEESIGM	VGKKSIFS	FDEVNQTL
DfasCbfa	114	KDLTVIIPATKLLPSLCNGKKIVVLF	FFG---DQEEYVGI	VGKKSIFS	FDEVNQTL
PpalCbfa	116	RDLTVIIPSTKLLPSLRVGKKIVVLF	FFG---NQDEYVGI	VGKKSIFAF	FDEVNQTL
AsubCbfa	116	KDLHVVPSSLKLLPSLA-GKKIVVLF	FFG---QPDDFRI	VGKKS	TMTFVEATE
DdisCbfa	178	EVPLAQLEDLNFNTEPEIAMKKDIAAF	NYKNQIEEKEEGLYVKQELYN	KKII	
DpurCbfa	178	EVPLKQLEDLNFNTEPEIATKKDIST	FNKYTOMEKEEGLYVKQEL	-----	
DfasCbfa	171	EGPLKQLELLFNMTPEVIAASKLEIAT	FNKYQISEEGLTIINKPD	-----	+ 255 AS
PpalCbfa	173	SEPLQLEILFSVTDPLIANKKEIQTF	DYKSQIDETEGLAVKASL	-----	+ 107 AS
AsubCbfa	171	EEVLQQLDLNFNICTPTVATKVEFNE	FNYQSOLVETEGLAVAGTPG	NGRMM	+ 223 AS

**Abbildung 31: Sequenzvergleich der Carboxy-terminalen Domänen von CbfA und orthologen Proteinen;** die Anordnung wurde mit **ClustalX 2.0.11** (Larkin *et al.* 2007) erstellt und mit BoxShade 3.21 ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)) eingefärbt; *Dictyostelium discoideum* (Phyletische Gruppe nach Schaap *et al.* (2006): 4), *D. purpureum* (4), *D. fasciculatum* (1), *Polysphonylium pallidum* (2), *Acytostelium subglobosum* (2), Position 1 entspricht AS 771 von *D. discoideum* CbfA, schwarze AS identisch, graue AS mit ähnlichen physiko-chemischen Eigenschaften

Die Carboxy-terminalen Domänen der CbfB-Proteine zeigen untereinander nur geringe Ähnlichkeit. Eine gewisse Sequenzhomologie ist bei den phylogenetisch eng verwandten Organismen *Polysphonylium pallidum* und *Acytostelium subglobosum* in der CTD der CbfB-Proteine zu erkennen.

### 3.18 Bakterielle Expression der Carboxy-terminalen Domäne zur kristallographischen Strukturaufklärung

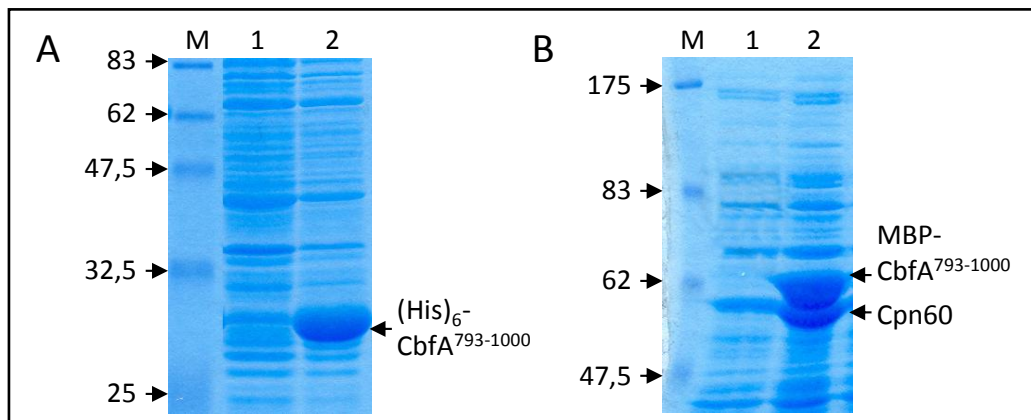
Oft lassen sich bestimmte funktionelle Eigenschaften eines Proteins aus strukturellen Untersuchungen ableiten. Dazu benötigt man immer größere Mengen des zu untersuchenden Proteins. Es existieren für Proteine und Proteinfragmente eine ganze Reihe von Systemen zur Proteinaufreinigung, dessen Vor- und Nachteile sich je nach gewünschter Anwendung unterscheiden. Für die als Ziel angestrebte Aufklärung der dreidimensionalen Struktur der CTD mittels Röntgenkristallographie ist die Expression in Bakterien mit anschließender Aufreinigung über eine Affinitätsmarkierung angewandt worden. Diese Methode produziert das höchste Maß

an einheitlichem Protein und das Risiko eines regulatorischen Effektes ist aufgrund der phylogenetischen Distanz sehr gering. Als Startpunkt diente die für Aminosäuren 793-1.000 codierende Sequenz von *cbfA* aus AX2. Diese wurde mit den Primern CMBF-124/125 amplifiziert und nach einer Zwischenklonierung in die Expressionsvektoren pMAL-c2x (NotI) und pET33b(+) kloniert. Die entstehenden Vektoren codieren für Fusionsproteine mit einer Größe von 66,4 kDa (pMAL codierte Fusion mit Maltose-bindendem Protein) bzw. 29 kDa (pET codierte Fusion mit Histidinhexamer). Die beiden Vektoren unterscheiden sich in der Art des zur Aufreinigung verwendeten Affinitätspeptides und vermitteln dadurch bedingt unterschiedliche Eigenschaften des Fusionsproteins (Brizzard 2008). Die Fusion mit dem maltosebindenden Protein (MBP) addiert ein gesamtes funktionelles Polypeptid zu der Zielsequenz. Die Aufreinigung erfolgt über die Bindung des Fusionsproteins an immobilisierte Maltose, die Elution durch Erhöhung der Maltosekonzentration. Das mit dem Histidinhexamer markierte Protein kann über eine Metal-Chelat-Chromatographie mit Nickel gereinigt werden. Die Elution erfolgt durch eine Verdrängungsreaktion mit Imidazol.

Beide Konstrukte wurden in fünf verschiedenen Expressionsstämmen getestet. Dies sind eine Serie von *Escherichia coli* BL21 (DE3)-Stämmen von Novagen ohne und mit zusätzlichen, plasmidcodierten tRNAs („RIL“ und „ROSETTA“). Ein bei der Aufreinigung von rekombinanten Proteinen sehr oft beobachtetes Problem ist dem unterschiedlichen Codongebrauch des Donor- und des Expressionsorganismus geschuldet. Ein und dieselbe Aminosäure besitzt in jedem Organismus eine unterschiedliche Reihenfolge in der Gebrauchshäufigkeit eines für sie codierenden Basentriplets. Diese Reihenfolge wird direkt durch den GC-Gehalt bestimmt. Dieser unterscheidet sich zwischen *E. coli*, etwa 51%, und *D. discoideum* mit 22% (Eichinger *et al.* 2005) sehr stark. Die auftretenden Probleme bei der Expression können durch zusätzliche Expressierung von selten in *E. coli* genutzten tRNAs überwunden werden. Ein weiteres sehr häufig beobachtetes Phänomen ist die Temperaturabhängigkeit des löslichen Anteiles des Fusionsproteins. Lösliches Fusionsprotein ist unlöslichem vorzuziehen, da Techniken zur Aufreinigung unlöslicher Polypeptide eine weitere sehr komplexe Ebene in die Fragestellung einbringen. Deshalb wurden in die Expressionsversuche die Stämme *E. coli* ArcticExpress™ (DE3) und (DE3) RIL von Stratagene mit einbezogen. Diese exprimieren zusätzlich zu dem

gewünschten Fusionsprotein die Chaperonine Cpn10 und Cpn60 aus *Oleispira antarctica* (Ferrer *et al.* 2003), was ein Wachstum des Expressionsstammes bei 12°C erlaubt. Dies erhöht in einigen Fällen die Faltung und damit die Löslichkeit des Proteins.

Alle fünf benutzten Stämme wurden mit beiden Expressionsvektoren transformiert. Anschließend sind jeweils vier Klone konserviert worden, welche dann als Ausgang für einen Induktionsversuch genutzt worden sind. Dabei wurden die BL21-Stämme bei 37°C für 2 Stunden induziert, die ArcticExpress™ bei 12°C für 22-24 Stunden. Die erreichten optischen Dichten waren vergleichbar. Eine Verlängerung der Induktion ging bei den bei 37°C wachsenden Bakterien mit einer weiteren Zunahme der optischen Dichte verbunden mit mehr gebildeten Fusionsprotein einher, die ArcticExpress™-Stämme wuchsen nicht weiter. Der Anteil an induziertem Fusionsprotein wurde mittels einer SDS-PAGE für alle erhaltenen Klone abgeschätzt. In Abbildung 32 sind exemplarisch zwei Klone unter unterschiedlichen Expressionsbedingungen dargestellt. Die Menge des gebildeten Fusionsproteins war in allen getesteten Kombinationen ähnlich. Für weitere Analysen wurde jeweils ein Klon gewählt.



**Abbildung 32: Vergleich der Induktion von Fusionsproteinen mit CbfA<sup>793-1000</sup>;** (A) Fusion mit Hisdidinhexamer exprimiert bei 37°C in *E. coli* BL21 (DE3) (B) Fusion mit Maltose-bindenden Protein exprimiert bei 12°C in *E. coli* ArcticExpress™ (DE3) RIL; M = prestained protein marker, broad range (NEB), 1 = Probe vor Induktion, 2 = Probe 2 Stunden nach Induktion; alle Molekülmassen in kDa

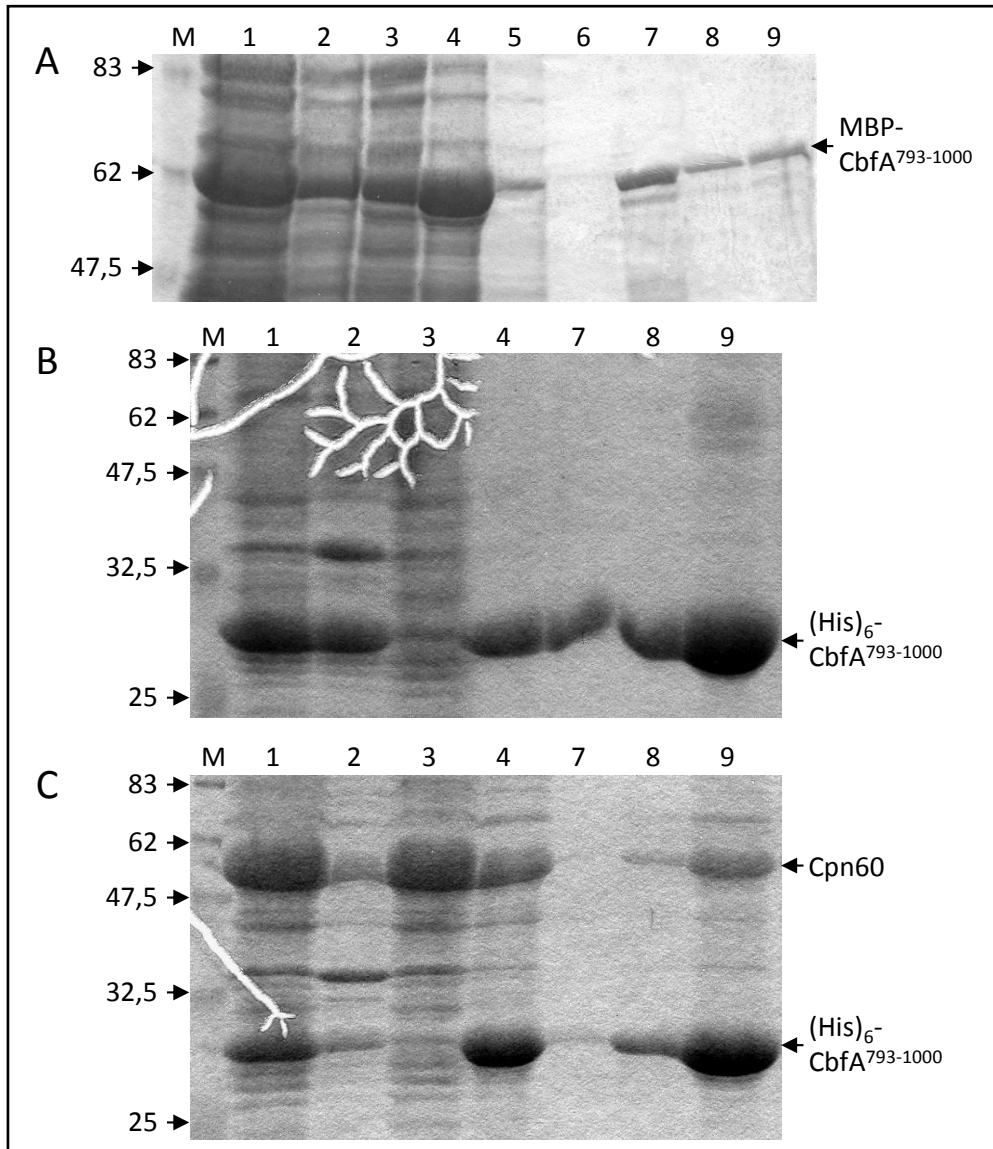
Die Löslichkeit von markiertem CbfA<sup>793-1000</sup> ist in Abhängigkeit von der Temperatur und der Induktionsdauer getestet worden. ArcticExpress™ (DE3)-Zellen wurden zusätzlich zu dem 12°C-Versuch bei 16°C und Raumtemperatur induziert, BL21 (DE3) bei Raumtemperatur, 30°C und 37°C. Die Menge an gebildeten, löslichen Fusionsprotein war für alle untersuchten Konditionen

standardisiert auf die optische Dichte bei 600 nm gleich und lag immer bei über 50% (siehe Spuren 1 und 2 der Abbildung 33). Auch die Induktionsdauer hatte keinen sichtbaren Einfluss. Beobachtete Variationen sind durch unterschiedlich effiziente Zelllyse, welche in späteren Versuchen durch mikroskopische Betrachtung überprüft wurde, zu Stande gekommen.

Für die weitere Aufreinigung des Zielproteins sind Versuche in Verbindung mit der Affinitätschromatographie durchgeführt worden. In Abbildung 33 sind Bilder von einigen repräsentativen Gelelektrophoresen kleiner Aufreinigungen im so genannten *Batch*-Verfahren dargestellt. In den mit 1 und 2 gekennzeichneten Spuren ist die Löslichkeit des jeweiligen Fusionsproteins zu erkennen. Hier dargestellte Unterschiede weisen auf eine unzureichende Zelllyse hin. Die Spuren 3 und 4 zeigen die Bindung an das Säulenmaterial, Ni-NTA Agarose (Qiagen) für die Histidinmarkierung und Amylose (NEB) für das maltosebindende Protein. Der Anteil an gebundenen Protein ist für die (His)<sub>6</sub>-Fusion etwas höher, die Nebenbanden von ebenfalls bindenden Proteinen kann durch eine Erhöhung der Imidazolkonzentration des Bindepuffers gesenkt werden (vergleiche Teil B und C der Abbildung 33). Nach dem Waschen des Säulen-Protein-Komplexes erfolgte die Elution mit Maltose bzw. Imidazol. Dabei sind die Volumenverhältnisse Elutionspuffer zu Säulenmaterial von großer Bedeutung und mussten optimiert werden. Bei initialen Versuchen, wie in Abbildung 33 dargestellt, wurde durch zu geringes Volumen nur wenig Protein eluiert. Bei der Fusion mit dem MBP konnte die Effizienz der Elution nicht mehr gesteigert werden ohne gleichzeitig den Grad der Verunreinigungen zu steigern. Es wurde sich deshalb für Versuche im größeren Maßstab für das Histidinhexamer als Affinitätsmarkierung entschlossen. Im Bezug auf die Expressionsstämme ist die Menge des aufgereinigten Proteins in allen späteren Versuchen vergleichbar geblieben. Wie aber schon in der Abbildung 33 C in den Spuren 7 bis 9 zu erkennen ist, wurde eine große Menge Cpn60 in den ArcticExpress™-Stämmen mit gereinigt. Es ist davon auszugehen, dass dies durch die Interaktion des Fusionsproteins mit dem Chaperon vermittelt wird. Eine direkte Bindung von Cpn60 an das Säulenmaterial in beiden Systemen ist sehr unwahrscheinlich. Ferner kann sie durch Erhöhung des Imidazolgehaltes des Waschpuffers nicht beeinflusst werden. Eine Trennung von Fusionsproteinen und Cpn60 mit anderen Maßnahmen ist zwar möglich (Joseph und Andreotti 2008), resultiert aber in einer weiteren Unwägbarkeit. Die Expression des Fusionsproteins in

## Ergebnisse

*E. coli* BL21 (DE3) bei 37°C mit Hilfe des Plasmides pET33b(+)-CTD wurde anhand der dargelegten Versuche als Ausgangspunkt für die Optimierung eines Ansatzes im großem Maßstab ausgewählt.



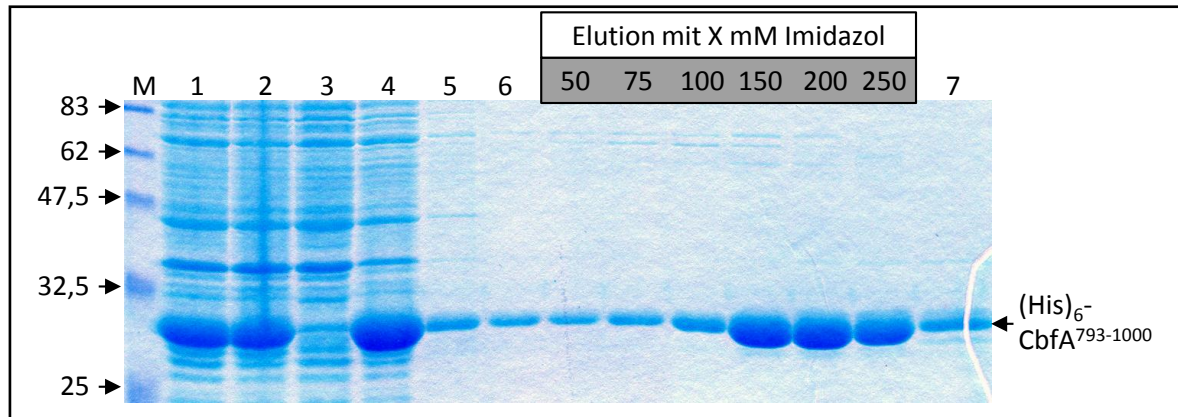
**Abbildung 33: Induktions- und Aufreinigungsversuche im kleinem Maßstab; (A)** pMAL-c2x (NotI)-CTD in *E. coli* BL21 (DE3) bei 37°C induziert **(B)** pET33b(+)-CTD in *E. coli* BL21 (DE3) bei 37°C induziert **(C)** pET33b(+)-CTD in *E. coli* ArcticExpress™ (DE3) bei 16°C induziert; M = prestained protein marker, broad range (NEB), 1 = lösliche Fraktion nach Zellaufschluss, 2 = unlösliche Fraktion nach Zellaufschluss, 3 = Durchfluss nach Bindung an Säulenmaterial, 4 = an Säulenmaterial gebundenes Protein, 5 und 6 = Waschschritte, 7 = erste Elution, 8 = zweite Elution, 9 = an Säule verbleibendes Protein; alle Molekülmassen in kDa, Gelbilder durch Verringerung der Farbsättigung bearbeitet



### 3.19 Herstellung von rekombinanter CTD im großen Maßstab

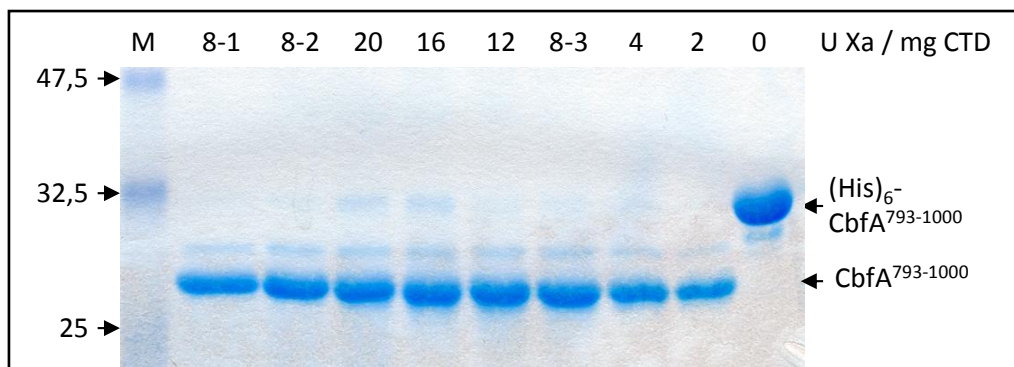
Nach Festlegung der Wachstumsparameter im kleinem Maßstab wurde das genutzte Expressionsplasmid pET33b(+)-CTD durch Mutagenese modifiziert. Das Ausgangsplasmid codiert für eine 50 Aminosäuren (AS) lange Fusion, welche neben den (His)<sub>6</sub> eine Thrombinschnittstelle enthält. Ein späterer Schnitt an dieser Stelle lässt ein 33 AS langes Peptid am N-Terminus übrig. Dies ist für viele Anwendungen von Nachteil. Deshalb wurde direkt vor dem Triplet für die erste CTD-Aminosäure eine Faktor Xa-Schnittstelle eingeführt. Ein Schnitt mit diesem Enzym führt zu einer von Fremdsequenzen freien CbfA<sup>793-1000</sup>. Das modifizierte Plasmid wird als pET33b(+)-CTD (Xa) bezeichnet und diente nach einer testweisen Aufreinigung im kleinem Maßstab als Ausgang für alle Optimierungsschritte der Aufreinigung im größeren Maßstab. Dabei wurden alle Arbeitsschritte im Einzelnen auf ihre effizienteste Methode überprüft. Die Vorversuche zur Festlegung des Expressionsstammes und des zu nutzenden Systems zeigten bereits eine Unabhängigkeit der Löslichkeit und die Fähigkeit der Bindung zum Säulenmaterial von der Temperatur und Induktionsdauer. Deshalb wurden alle Versuche bei 37°C mit einer 2,5- bis 4-stündigen Induktion durchgeführt. Die Induktion wurde nach dem Erreichen des Maximums der optischen Dichte abgebrochen und die Zellen geerntet. Die Vorversuche zeigten ferner oft eine nicht komplette Lyse der Zellen. Die Effektivität der Lyse konnte durch das Benutzen einer *French-Press* erhöht werden. Zum Binden an das Säulenmaterial, zum Waschen und für die Elution sind Versuche im oben genannten *Batch*-Verfahren sowie mit durch Sedimentation selbstgepackten Säulen durchgeführt worden. Dabei stellte sich die selbstgepackte Säule mit steigendem Volumen als effizienter dar. Gerade die Elution ist hier vollständiger bei gleichbleibend geringem Anteil von Verunreinigungen. Des Weiteren wurde die Elution von (His)<sub>6</sub>-CbfA<sup>793-1000</sup> in Abhängigkeit von der Imidazolkonzentration bestimmt (Abbildung 34). Bis 100 mM Imidazol bleibt die dem technischen Vorgang des Waschens geschuldete eluierte Menge gleich, ab 150 mM ist eine aktive Elution zu erkennen. Aus dieser Abbildung ist ebenfalls ersichtlich, dass immer ein geringer Rest von Fusionsprotein an dem Säulenmaterial verbleibt.

## Ergebnisse



**Abbildung 34: Abhängigkeit der Elution von  $(\text{His})_6\text{-CbfA}^{793-1000}$  von der Imidazolkonzentration;** M = prestained protein marker, broad range (NEB), 1 = lösliche Fraktion nach Zellaufschluss, 2 = unlösliche Fraktion nach Zellaufschluss, 3 = Durchfluss nach Bindung an Säulenmaterial, 4 = an Säulenmaterial gebundenes Protein, 5 und 6 = Waschschritte mit 20 mM Imidazol, 7 = an Säule verbleibendes Protein; alle Molekülmassen in kDa

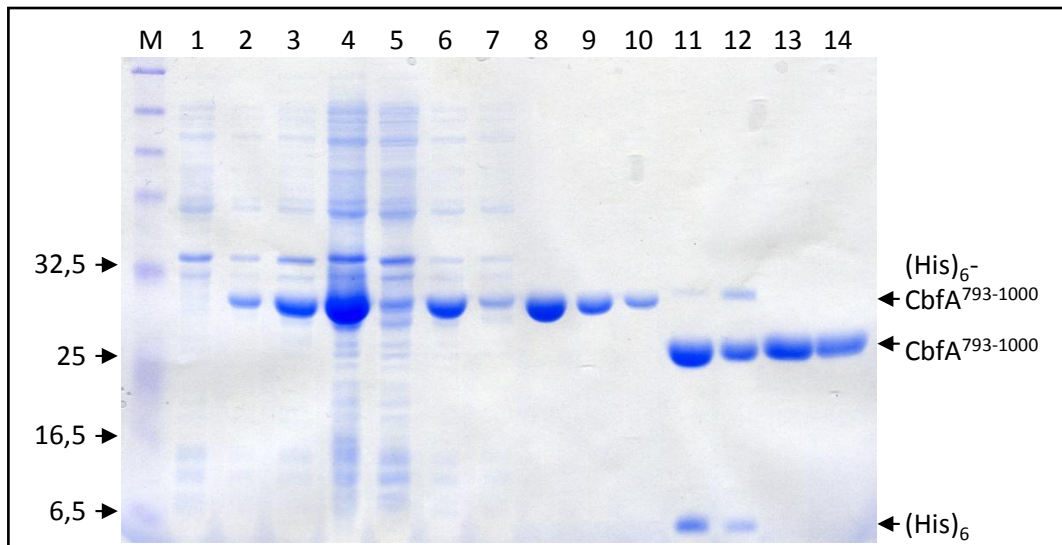
Das gereinigte Protein wurde nach der Elution von seiner für die Aufreinigung genutzter Markierung getrennt. Dazu erfolgt zuerst eine Dialyse der Aufreinigung gegen einen für die zum Schneiden genutzte Faktor Xa-Protease geeigneten Puffer. Anschließend wurde experimentell das Verhältnis der Protease zur Proteinmenge bestimmt. Es konnte für das Fusionsprotein gezeigt werden, dass pro mg zwei Units Enzym ausreichend sind (Abbildung 35). Die mit steigender Faktor Xa-Konzentration zunehmende Bande bei 30 kDa entspricht dem Molekulargewicht einer der Ketten des Faktors Xa. Die Bande oberhalb der geschnittenen CTD wurde nicht weiter analysiert, da sie in folgenden Reinigungsschritten entfernt wurde.



**Abbildung 35: Bestimmung des optimalen Verhältnisses Faktor Xa zum Fusionsprotein;** in jedem Versuch wurden 250 µg gereinigtes Protein mit der angezeigten Menge Faktor Xa 24 Stunden bei 4°C geschnitten; 8-1 1 mg / ml Proteinkonzentration, 8-2 2,5 mg / ml, Rest 5 mg / ml; alle Molekülmassen in kDa

Die letzten Reinigungsschritte dienen der Entfernung des Faktor Xa sowie des geschnittenen Affinitätsmarkers über entsprechendes Säulenmaterial. Ferner können auch so etwaige nicht geschnittene Mengen an CTD entfernt werden. Besondere Aufmerksamkeit wurde dabei auf die Optimierung des verwendeten Waschpuffers der Ni-NTA gelegt. Dieser enthält das gereinigte Protein.

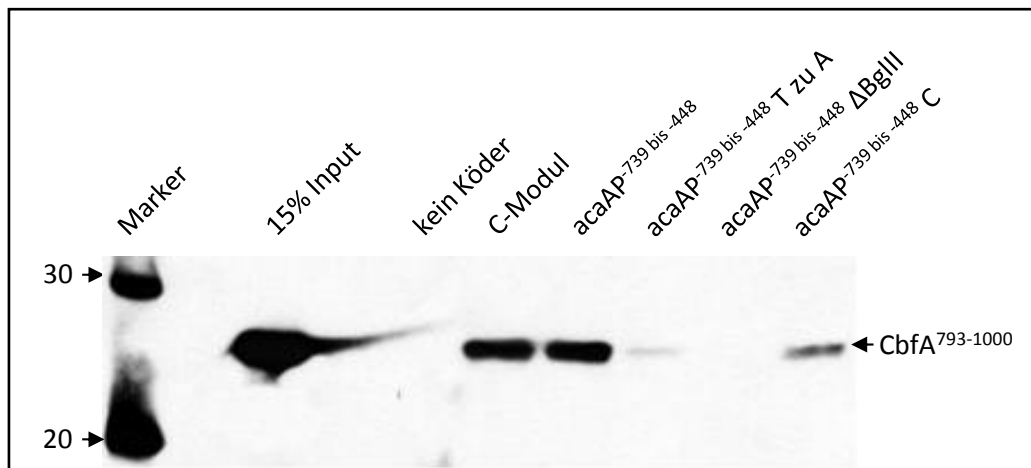
Das in Material und Methoden dargestellte Verfahren ergibt sich aus dem praktikabelsten Varianten der genannten Optimierungen unter Beachtung einer möglichst hohen Ausbeute. Die elektrophoretische Auftrennung von Proben (*SDS-PAGE*) einer Aufreinigung basierend auf einer zwei Liter umfassenden *E. coli*-Kultur ist in Abbildung 36 zu sehen. In der dargestellten Aufreinigung konnten aus 4 g Feuchtgewicht Bakterien 33 mg CbfA<sup>793-1000</sup> isoliert werden.



**Abbildung 36: Aufreinigung von rekombinatem CbfA<sup>793-1000</sup> mittels Affinitätschromatographie;** *SDS-PAGE* von Proben der Proteinaufreinigung im großen Maßstab; M = prestained protein marker, broad range (NEB); 1: nicht induzierte Probe (5µg Gesamtprotein), 2: induzierte Probe (5µg), 3: unlösliches Fraktion nach Lyse (gleiches Volumen wie 4), 4: lösliche Fraktion (15µg), 5: nicht gebundener Durchfluss (gleiches gewichtetes Volumen wie 6), 6: an Ni-NTA gebundenes Protein (15µg), 7: Waschschrift 1 (gleiches Volumen wie 5), 8: Elution 1 (15µg), 9: Elution 2 (gleiches Volumen wie 8), 10: an Ni-NTA verbleibendes Protein (gleiches gewichtetes Volumen wie 8), 11 und 12: mit Faktor Xa geschnittenes Protein (15µg), 13 und 14 aufgereinigtes Protein (15µg)

### 3.20 Nachweis der Funktionalität des rekombinant gereinigten Proteins

Für weiterführende Analysen ist es von Bedeutung, dass das gereinigte Protein in seiner natürlichen Konformation vorliegt. Einen guten Anhaltspunkt bietet dabei ein Test auf die biologische Funktion. Da für das rekombinant in *D. discoideum* exprimierte GFP-CbfA<sup>724-998</sup> eine Bindung an das C-Modul sowie an den Promotor von *acaA* gezeigt werden konnte (Abbildung 25 B), wurde dieser Test auch auf diese Fragestellung angewendet. Der einzige methodische Unterschied bestand in der Zugabe von 1 mg/ml Bovinen Serumalbumines zum Interaktionsansatz. Dies entsprach den gemessenen Gesamtproteinkonzentrationen der unter Kapitel 3.12 benutzten Kernextrakte. In Abbildung 37 ist der Nachweis der CTD durch einen *Western-Blot* mit dem Antikörper 7F3 nach dem Interaktionstest zu erkennen.



**Abbildung 37: Funktionstest des rekombinanten Proteins durch Bindung an bekannte Promotorfragmente;** Bindung von gereinigter CTD an biotinylierte *acaAP*<sup>-739 bis -448</sup>, C-Modul von TRE5-A, *acaAP*<sup>-739 bis -448</sup> T zu A: homopolymere dT-Bereiche durch 2 mal 10 dA unterbrochen; *acaAP*<sup>-739 bis -448</sup> ΔBgIII: CbfA-bindender Bereich -676 bis -555 entfernt; *acaAP*<sup>-739 bis -448</sup> C: CbfA-bindender Bereich -676 bis -555 mit bindenden Bereich des C-Moduls ersetzt (Vergleiche Abbildung 25); Marker = Magic Mark (Invitrogen)

Für den gezeigten Test wurden 10 pmol rekombinante CTD und 5 pmol von der DNA-Probe verwendet. Ein Viertel des Bindungsversuches wurde aufgetragen. Die Proben mit dem C-Modul, *acaAP*<sup>-739 bis -448</sup> und die DNA mit den ausgetauschten Bindebereichen werden spezifisch gebunden. Interessanterweise wird auch in diesem Test die Promotorregion mit den unterbrochenen Thymidinbereichen im geringen Umfang gebunden. Eine Erhöhung des Anteiles

der eingesetzten rekombinanten CTD auf 125 pmol intensiviert diese Bindung. Der Promotorbereich ohne Bindemotiv wurde auch unter diesen Umständen nicht gebunden.

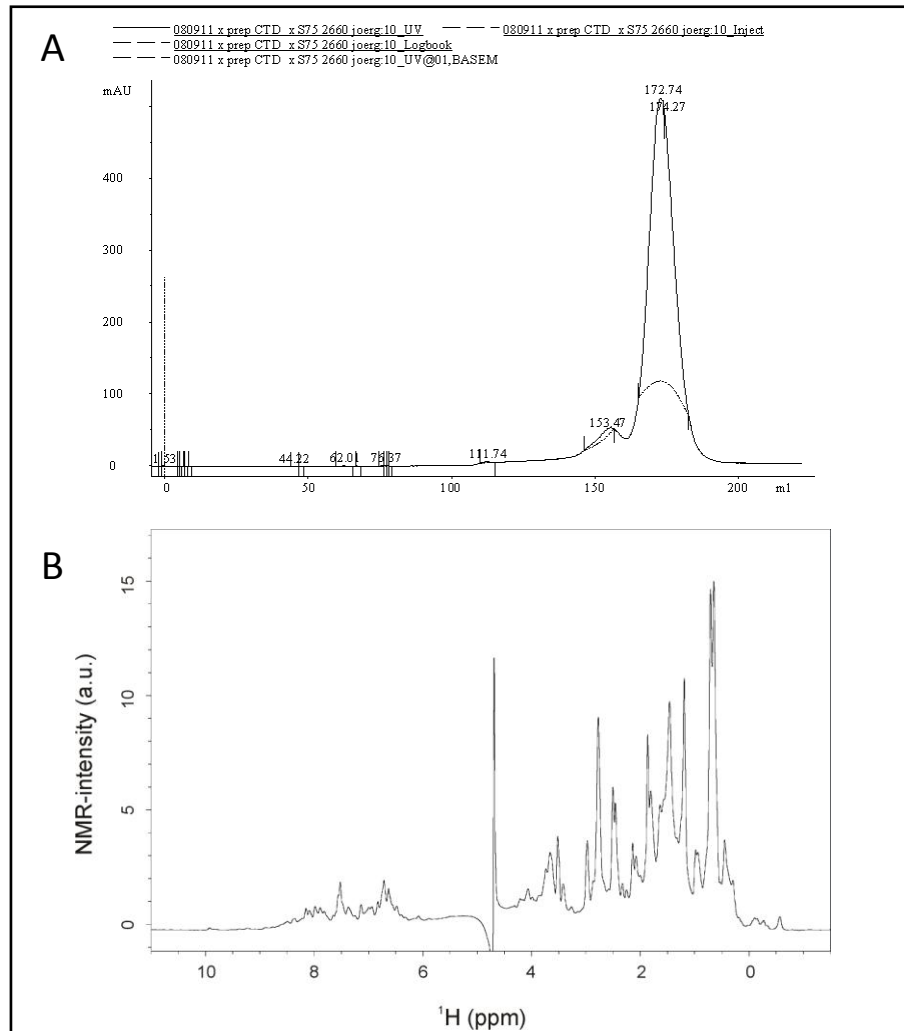
Es ist davon auszugehen, dass zumindest ein Teil der Präparation von CbfA<sup>793-1000</sup> aus *E. coli* BL21 (DE3) spezifische DNA-bindende Eigenschaften hat und somit die ursprüngliche Konfirmation aus *D. discoideum* besitzt.

### **3.21 Weiterführende Arbeiten zur kristallographischen Strukturaufklärung**

Die Primärstruktur der Carboxy-terminalen Domäne ist nach aktuellem Wissensstand nur in CbfA und seinen homologen Proteinen im näheren Umfeld der *Dictyosteliales* vorhanden (Vergleiche Kapitel 3.17). Rückschlüsse auf die Funktionsweise der CTD sind deshalb aufgrund mangelnder Homologien mit bekannten Proteinstrukturen nicht zu ziehen. Die höhere Struktur der CTD kann dabei Homologien offenlegen, welche sich aufgrund der Sequenz nicht erschließen. Deshalb wurden zusammen mit Abteilung Physikalische Biotechnologie der Martin-Luther-Universität Halle-Wittenberg unter der Anleitung von Dr. Than-Breithaupt Versuche zur Bildung von Proteinkristallen durchgeführt.

Dazu wurde die gereinigte, auf ihre Funktionalität überprüfte CTD über eine Sephadex™ G75-Säule gelfiltriert. In Abbildung 38 A ist das Diagramm der Extinktion bei 280 nm gegen das Elutionsvolumen einer solchen präparativen Größenausschlusschromatographie zu erkennen. Dabei ist die Verweildauer in der Säule von der Masse des Proteins sowie von seiner Konfirmation abhängig. Kleinere Proteine benötigen länger für die Elution. Neben dem Hauptpeak bei 175 ml existiert ein Vorpeak bei 150 ml. Dieser enthält neben dem Protein auch einen Anteil von DNA, welche durch ein zweites Maximum des Spektrums im Bereich um 260 nm identifizierbar ist. Eine Inkubation mit DNase während und kurz nach dem Zellaufschluss minimiert diese Verunreinigung in späteren Aufreinigungen. In dem gezeigten Beispiel konnten beide Peaks aufgrund des großen Volumens getrennt werden. Das Elutionsvolumen der CTD entspricht der geeichten Säule nach circa 20-30 kDa. Es konnte in keiner präparativen oder analytischen Gelfiltration die Bildung von Oligomeren der CTD beobachtet werden.

## Ergebnisse



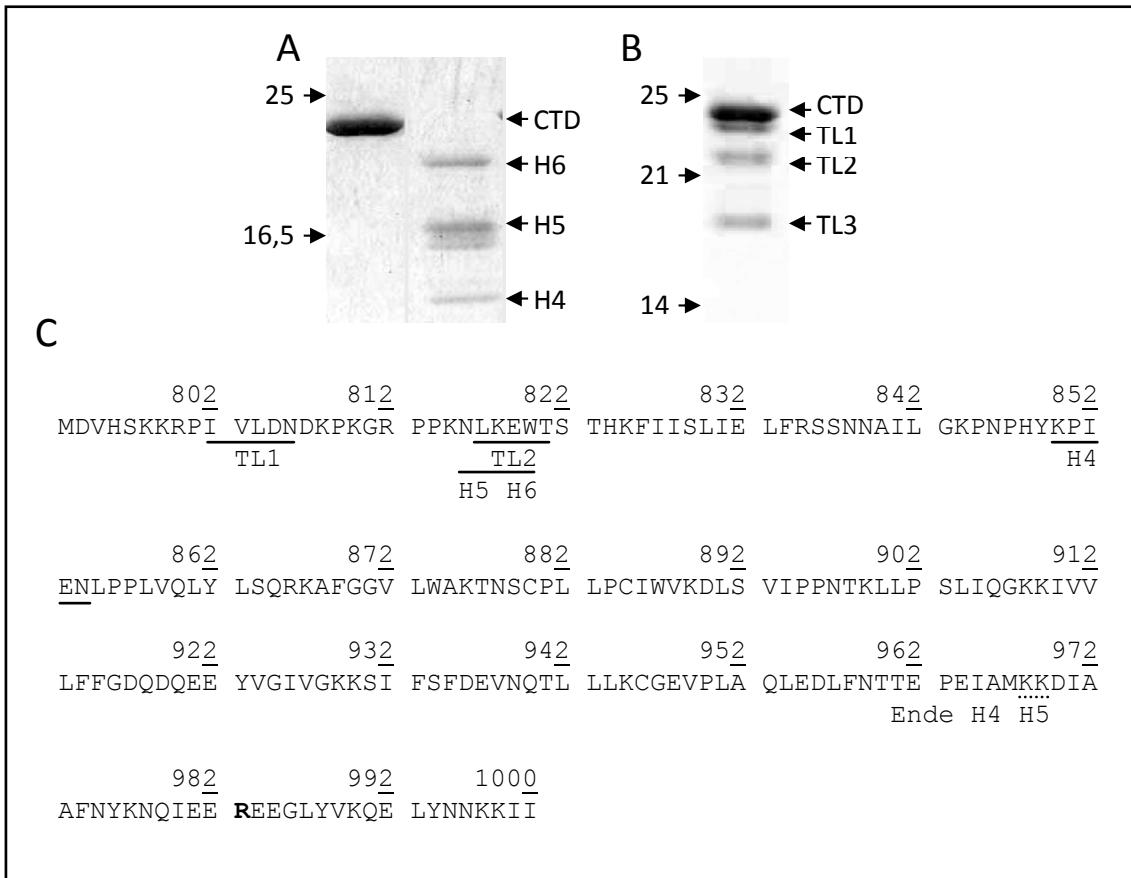
**Abbildung 38: Überprüfung der Strukturierung und Qualität des rekombinant gereinigten Proteins; (A)** Präparative Gelfiltration von 22 mg CbfA<sup>793-1000</sup>; gezeigt ist die Extinktion bei 280 nm in mAU gegen das Elutionsvolumen in ml, 1 AU entspricht einer Extinktion von 2 bei einer Schichtdicke von 1 cm **(B)** 1D-NMR einer Probe des Hauptpeaks von A; dargestellt ist die Intensität gegen die chemische Verschiebung  $\delta$  in ppm

Die durch Elektrospray-Ionisation-Massenspektroskopie (ESI-MS) bestimmte Masse des Hauptpeaks liegt bei 23772 Dalton und hat eine Differenz von 28 Dalton zur berechneten Masse. Nähere Untersuchungen dieser Differenz decken die Änderung von Lysin<sub>983</sub> zu Arginin auf. Aufgrund nicht eindeutiger Datenlage an dieser Stelle in der Sequenz des GenBank®-Eintrages AF052006.2 sowie der ähnlichen physiko-chemischen Eigenschaften von Arginin und Lysin ist mit dem R<sub>983</sub>-haltigen Protein weiter gearbeitet worden. Der Nachweis der Strukturierung der gereinigten CTD erfolgte durch eine eindimensionale Kernspinresonanzspektroskopie (1D-NMR). Das entsprechende Protonenspektrum wurde von

Dr. Ulrich Weininger, Institut für Physik, Biophysik, Martin-Luther-Universität Halle-Wittenberg aufgenommen und ist in Abbildung 38 B dargestellt. Von großer Aussagekraft über die Faltung eines Proteins sind dabei die Bereiche über 8,5 ppm und zwischen 1 und -1 ppm sowie das Fehlen eines deutlichen Peaks bei 8,3 ppm (Rehm *et al.* 2002). Das dargestellte Spektrum deutet auf eine Strukturierung der rekombinant gereinigten CTD hin. Anschließend sind mehrere hundert Kristallisationsansätze im kleinen Maßstab präpariert worden. Nach mehrmonatiger, ergebnisloser Inkubation fiel dabei die komplette Proteolyse des in den Ansätzen enthaltenen Proteins auf. Entgegen der extensiv auf Proteolyse getesteten, Thrombin-geschnittenen CTD ohne Faktor Xa-Schnittstelle zeigte das neue Konstrukt Abbauprodukte bei Raumtemperatur oder bei langer Inkubation bei 4°C (Abbildung 39 A). Dies deutet auf eine Restaktivität des Faktor Xa selbst oder auf eine pufferbedingte Aktivierung anderer Proteasen durch das eingetragene Calcium hin. Die Zugabe von EDTA nach Beendigung der Aufreinigung reduzierte zwar die proteolytische Aktivität, Abbaufragmente waren trotzdem nach einer Woche Inkubation bei 25°C erkennbar.

Die Art der Proteolyse ist oftmals ein Hinweis auf flexible Bereiche innerhalb eines Proteins, welche eine Kristallbildung verhindern können. Um dies weiter zu spezifizieren, wurde gereinigte, durch Einfrieren vor Degradierung geschützte CTD mit verschiedenen Proteasen inkubiert. Dabei ist das durch Thermolysin entstehende Spaltmuster (Abbildung 39 B) am interessantesten, da es mehrere Produkte mit ähnlicher Größe wie die spontane Proteolyse ergibt. Andere getestete Proteasen ergaben jeweils ein oder kein Abbaufragment. Die N-terminale Sequenzierung der spontanen Abbauprodukte sowie der Thermolysin-Fragmente sind in der Arbeitsgruppe von Prof. Dr. Grosse am Leibniz-Institut für Altersforschung – Fritz-Lipmann-Institut e.V. (Jena) durchgeführt worden. Die Massen wurden in Halle bestimmt. Einen Überblick darüber gibt Tabelle 10, die bestimmten Fragmente sind in der Sequenz von Abbildung 39 C markiert. Die N-terminale Sequenzierung des Fragmentes TL3 schlug leider fehl und wird momentan wiederholt. Dies ist aufgrund der terminlichen Koordination der Partnerlabore erst nach Erstellung dieser Arbeit möglich.

## Ergebnisse



**Abbildung 39: Aufklärung der Abbaufragmente der Kristallisationsansätze; (A) SDS-PAGE** spontaner Proteolyse und **(B) Thermolysinverdau**, jeweils 5µg; **(C) Sequenz Cbfa<sup>793-1000</sup>** mit Zuordnung der N-terminalen Sequenzierung (Linie) und des C-terminalen Abbruchs über die Gesamtmasse (gestrichelte Linie); R<sub>983</sub> **fett** (Erläuterung im Text); TL3 nicht auswertbar; alle Molekülmassen in kDa

**Tabelle 10: Masse und N-terminale Sequenzierung der CTD-Abbauprodukte; \*Massenbestimmung** ergibt zwei Werte entsprechend der nicht getrennten Doppelbanden in Abbildung 39, N-Terminus beider Banden identisch; geklammerte Aminosäuren des N-Terminus sind Nebensequenzen

Fragment	Masse	N-Terminus	Abdeckung CTD
CTD (K983R)	23.772 Da	M(q)DVHS	793-1.000
Thermolysin 1 (TL1)	22.692 Da (22.466 Da)*	IVLDN	802-1.000 (998)
Thermolysin 2 (TL2)	21.020 Da	L(q)KEW(q,g)T	817-1.000
Thermolysin 3 (TL3)	18.202 Da (17.975 Da)*	Sequenzierung mehrfach fehlgeschlagen	
spontaner Abbau (H4)	12.981 Da	K(q)PIEN	850-968
spontaner Abbau (H5)	16.954 Da	N(q)LKEW	818-968
spontaner Abbau (H6)	20.907 Da	N(q)LKEW	818-1.000



Es lassen sich zwei Hauptbereiche der Spaltung identifizieren. Einerseits die Region um die Aminosäuren 815 bis 817, welcher in den Fragmenten TL2, H5 und H6 den N-Terminus des Abbauproduktes darstellt. Dies ist insofern problematisch, da hier die Interaktionsfläche mit der DNA, der AT-Haken, vermutet wird (Horn *et al.* 1999 und Kapitel 3.10 dieser Arbeit). Dieser Bereich ist ferner in allen orthologen CbfAs konserviert (Abbildung 31). Der zweite Bereich des Abbaus betrifft das C-terminale Ende von CbfA<sup>793-1000</sup>. Um die Aminosäure 968 kommt es zum Abbruch der Fragmente H4 und H5. Auch hier ist eine gewisse phylogenetische Konservierung zu erkennen, wobei eine etwaige Funktion im Unklaren liegt.

## 4 Diskussion

### 4.1 Einsichten in die regulatorische Funktion von CbfA

Durch Studien mit der *cbfA*-Suppressionmutante JH.D im Vergleich zum Laborstamm AX2 sind bereits einige durch CbfA-Mangel ausgelöste, phänotypische Veränderungen beschrieben worden. Neben dem Fehlen eines vollständigen Entwicklungszyklus in JH.D sind dies veränderte Pino- und Phagozytoseraten, eine etwas längere Generationsdauer und veränderte Transkriptmengen des Retrotransposons TRE5-A (Winckler *et al.* 2001 und 2004 sowie Siol *et al.* 2006a). Durch die Anwendung verschiedener vergleichender Transkriptomanalysen konnten diesen phänotypischen Änderungen zwischen AX2 und JH.D differentiell exprimierte Gene zugeordnet werden. Des Weiteren konnten phänotypisch noch nicht aufgefallene Änderungen identifiziert werden. Die differentiell exprimierten Gene sind nach biologischen Prozessen und Funktionen gruppiert worden.

Zu den bekannten phänotypischen Veränderungen von JH.D im Vergleich zu AX2 gehört die Erhöhung des Zellvolumens durch gesteigerte Pinozytose (Winckler *et al.* 2001). Diese Beobachtung kann durch die vergleichende Transkriptomanalyse zum Teil durch eine Anreicherung von in den Prozess der Pinozytose und Endozytose involvierten Proteinen spezifiziert werden (siehe Abbildung 16 A). Bemerkenswert ist dabei, die in JH.D im Vergleich zum Wildtyp geringere Expression der Gene *clc* (DDB\_G0277403) und *proA* (DDB\_G0287125), welche für eine in Bezug auf axenisch wachsende Stämme normale Pino- und Endozytoserate mitverantwortlich sind (Wang *et al.* 2003). ProA ist zusätzlich ein Regulator der Phagozytose (Cvrckova *et al.* 2004). Dies könnte im Zusammenhang mit der beobachteten verringerten Aufnahme von Bakterien in JH.D stehen (Winckler *et al.* 2001). Sillo und Kollegen untersuchten jüngst die transkriptionellen Unterschiede zwischen axenisch und auf Bakterien wachsenden Zellen (Sillo *et al.* 2008). Aus dieser Arbeit geht hervor, dass die Anpassung von der unter axenischen Bedingungen dominierenden Makropinozytose auf die Phagozytose von Bakterien mit der Änderung der Expression einer Reihe von Genen einhergeht. Der Lipidmetabolismus verändert sich, die Zusammensetzung der Zellmembran wird zugunsten von Sterolen verschoben. Eine Änderung des Lipidmetabolismus kann durch eine Funktionsanreicherung

innerhalb der *Microarray*-Daten ebenfalls in Abhängigkeit von CbfA belegt werden (Abbildung 16 A). Das Genprodukt von *plbE* (DDB\_G0277455) ist dabei am Abbau von Phospholipiden beteiligt (Morgan *et al.* 2004), DDB\_G0277473 katalysiert die Spaltung von Triacylglycerol zu Diacylglycerol und einem Fettsäureanion (DDB, FB, MGI, GOA, ZFIN curators 2001). Die Anpassung auf die Nahrungsquelle führt des Weiteren zu Änderungen im Kohlenwasserstoffmetabolismus. Der Abbau von Glukose, die Energiequelle axenischer *Dictyostelium*-Medien (Watts und Ashworth 1970, Franke und Kessin 1977), wird zugunsten der in den Nahrungspartikeln vorhandenen Inhaltsstoffe reduziert. Als verwertbare Alternative werden hier meist der Katabolismus von Aminosäuren genannt (Anjard *et al.* 2006). Der Abbau zyklischer Aminosäuren im CbfA-Suppressionsstamm JH.D ist stärker als in AX2. Ferner kommt es zu einem Anstieg der Glukoneogenese in JH.D-Zellen. Die Änderungen im Zellwand- und Peptidoglykanstoffwechsel sowie bei Komponenten des Zytoskelettes und des Drogentransportes sind durch die Aufnahme und Verdauung der Bakterien bedingt (Sillo *et al.* 2008). Für alle diese Prozesse konnte eine statistisch signifikante Anreicherung der gleichen Art im Vergleich von AX2 zu JH.D gefunden werden (Abbildung 16 und Anhang D). Zu erwähnen ist hierbei, dass zwar die bestimmten *GO-Terms* identisch sind, sich aber die geänderten Gene kaum überschneiden. Nur 17 der deutlich CbfA-abhängigen Gene sind auch in der Liste der 443 phagozytoseabhängigen Gene vertreten (Sillo *et al.* 2008). Die Daten deuten darauf hin, dass CbfA eine regulatorische Funktion auf einige an der Phagozytose beteiligten Gene besitzt, was die besseren phagozytotischen Eigenschaften von AX2 im Vergleich zu JH.D erklärt. Die Suppression von CbfA in JH.D führt zu einer gesteigerten Pinozytose. Ob dies als Reaktion auf eine geringere Phagozytose bedingt ist oder direkt durch CbfA reguliert wird, bleibt zu klären.

Als weiterer großer Komplex der CbfA-regulierten Gene fällt die Nukleotidbiosynthese auf. Durch die Suppression von CbfA in JH.D wird die Expression von fünf an der Neusynthese von Purinnukleosid-Monophosphaten beteiligten Genprodukten, bezogen auf das *Microarray*-Experiment, erhöht. Lässt man die gesetzte Mindestregulation von 1,5-fach außer Betracht oder bezieht man sich auf die durch RNA-Sequenzierungsmethoden entstandenen Daten, so sind mit Ausnahme des Gens *purA* alle am genannten Stoffwechselweg beteiligten Gene differentiell exprimiert (Abbildung 17). Dies gilt auch für die an der Bildung von Pyrimidinnukleosid-

Monophosphaten beteiligten Gene *pyr1-3*, *pyr4*, *pyr 5-6* und *thyA* (Wales *et al.* 1989, Myllykallio *et al.* 2002). Auch diese werden in JH.D um den Faktor 1,3 bis 1,6-fach mehr exprimiert als in AX2. Die fehlregulierte Aufnahme von Nahrungspartikeln könnte eine erhöhte *De-novo*-Synthese von Nukleosid-Monophosphaten zur Folge haben, da diese nicht mehr aus der Nahrung gewonnen werden können. Interessanterweise werden die an der Bildung von Nukleosid-Diphosphaten beteiligten Enzyme, wie zum Beispiel *gmkA* oder *adkB*, nur noch maximal 1,2-fach durch die CbfA-Suppression differentiell exprimiert. Die an der Bildung von Nukleosid-Triphosphaten, die für die Synthese der DNA und RNA benötigte Form der Nukleotide, beteiligten Enzyme *ndkM* und *ndkC* werden in JH.D 1,5-fach schwächer exprimiert als in AX2. Dies deutet darauf hin, dass die durch die CbfA-Suppression gesteigerte Syntheserate auf den Mangel an verwertbaren Nukleosid-Monophosphaten zurückzuführen ist, die DNA-Syntheserate gemessen an der Bildung von Triphosphaten aber in JH.D geringer als in AX2 ist. Dies könnte durch die verlängerten Generationszeit der Mutante (Winckler *et al.* 2001) erklärt werden.

Die genannten Eingriffe in basale Prozesse des Stoffwechsels vegetativ wachsender Zellen könnten den Umstand erklären, warum wiederholte Versuche zur Erstellung eines Laborstammes ohne CbfA-Expression, z.B. durch *Knock-out*-Techniken (Winckler *et al.* 2001), nicht zum Erfolg führten.

#### **4.2 Beteiligung von CbfA<sup>724-998</sup> an der CbfA-abhängigen Genregulation in JH.D**

Für den Einfluss einer Proteindomäne, in diesem Fall der Carboxy-terminalen Domäne von CbfA, auf die genregulatorische Funktion eines Transkriptionsfaktors sind vier Szenarien denkbar. Das erste Szenario ist eine Nichtbeteiligung der CTD an der Regulation eines Teiles der CbfA-abhängigen Gene. Dies würde zu Grunde legen, dass der N-terminale Teil entweder durch Elemente wie die beiden Zinkfinger an die DNA rekrutiert wird (wie u.a. in Elrod-Erickson *et al.* 1996 beschrieben) oder durch Protein-Protein-Interaktionen Teil eines Transkriptionskomplexes wird. Dieses Interaktionsverhalten ist in Brayer und Segal (2008) für die Zinkfinger und Trewick *et al.* (2007) für die Jumonji-Domäne anderer Proteine beschrieben worden. Die Funktion der Amino-terminalen Domäne könnte dabei in der Modifikation des Methylierungsstatus der

Histone und damit direkt in der Transkriptionsregulation liegen. Dies ist für JmjC-Domänen anderer Proteine bekannt (u. a. Tsukada *et al.* 2006, Klose *et al.* 2006, Chang *et al.* 2007, Seward *et al.* 2007). Ferner ist auch eine Interaktion abseits der DNA denkbar, welche zu einer Modifikation eines anderen Transkriptionsfaktors führt (z.B. Sumoylierung wie in Tronnersjö *et al.* 2006).

Die zweite Möglichkeit setzt die Beteiligung mindestens einer der N-terminalen Strukturen von CbfA an dem Regulationsmechanismus der CTD voraus. Denkbar sind auch hier die unter der ersten Möglichkeit genannten Wirkungsweisen.

Für diese beiden Möglichkeiten konnten in dieser Arbeit zwei Cluster von CbfA-abhängigen Genen identifiziert werden, welche nicht durch die CTD allein in ihrer Expression beeinflusst werden. In Abbildung 18 A sind einige durch Cluster I regulierte Prozesse zu erkennen. Dazu gehören neben der Pinozytose noch Prozesse und Funktionen, welche mit dem Abbau von Futterpartikeln in Verbindung gebracht werden können. Eine Beteiligung der CTD an der Regulation der Cluster I und VI-Gene kann aber im Rahmen der Datenlage dieser Arbeit nicht ausgeschlossen werden. Dazu sollte in weiterführenden Experimenten eine Komplementation des JH.D-Phänotypes mit dem N-terminalen Teil von CbfA Aufschluss geben.

Die wohl offensichtlichste Änderung des Phänotypes von AX2 zu JH.D ist die fehlende Fähigkeit zur Bildung von Fruchtkörpern. In dem Stamm JH.D werden eine Reihe von Entwicklungsmarkern nicht oder nur wenig durch das Hungern der Zellen induziert (Winckler *et al.* 2004). Dabei kann dieser Effekt jeweils durch die Zugabe von cAMP, rekombinantem CbfA (DDB\_G0279409), rekombinanter *pkaC* (DDB\_G0283907) oder rekombinanter *acaA* (DDB\_G0281545) zumindest phänotypisch geheilt werden (Winckler *et al.* 2004, Siol *et al.* 2006a). Die Ergebnisse des Kapitels 3.8 zeigen, dass die Komplementation des Entwicklungsphänotypes nicht durch die ektopische Expression der CTD von CbfA in JH.D vermittelt werden kann. Sowohl makroskopisch (Abbildung 19), als auch auf genregulatorischer Ebene (Abbildung 20) kann keine Revertierung des Phänotypes von JH.D durch die Expression von GFP-CbfA<sup>724-998</sup> festgestellt werden.

Es gilt als gesichert, dass das Ausbleiben der genregulatorischen Kaskade zur Bildung von Fruchtkörpern in JH.D auf die fehlende Regulation des Genes *acaA* zurückzuführen ist (Siol *et al.* 2006a). Es ist jedoch zu erwähnen, dass die CTD *in vitro* in der Lage ist, den Promotor der Adenylylcyase A (*acaAP*) zu binden. Dies gilt sowohl für das in *D. discoideum* exprimierte Proteinfragment (Abbildung 25 B), als auch für das in Bakterien exprimierte CbfA<sup>793-1000</sup> (Abbildung 37). Dabei ist die Bindungsspezifität ähnlich dem intrinsischen CbfA (Vergleiche Siol *et al.* 2006a und Abbildung 25 B). Diese Beobachtung kann dadurch erklärt werden, dass die CTD zumindest teilweise die DNA-Bindung vermittelt, zur Transkriptionsregulation aber nicht ausreichend ist. Dementsprechend muss die vom *acaA*-Genprodukt abhängige Entwicklung in Cluster I eingeordnet werden. Da die Expression von *acaA* seinen Höhepunkt nach acht bis zwölf Stunden Entwicklung erreicht und in vegetativen Zellen kaum Transkript vorhanden ist (Winckler *et al.* 2004, Transkriptomsequenzierung in dieser Arbeit), wird *acaA* in der *Microarray*-Analyse nicht als differentiell exprimiert erkannt.

Die dritte Möglichkeit für den Beitrag von CbfA<sup>724-998</sup> an der Transkription von Genen ist eine Bindung von CbfA an die DNA durch vermutlich den AT-Haken der CTD. Denkbar ist auch die Interaktion von einer noch nicht spezifizierten Protein-Protein-Interaktionsfläche innerhalb der CTD mit anderen Faktoren. In diesem Szenario vermittelt die Carboxy-terminale Domäne nur die räumliche Nähe der NTD zu dessen potentiellen Ziel, z.B. für die Demethylierung von Histonen (Chang *et al.* 2007). Diese Funktionsweise drängt sich für die zwölf in Cluster II und IV identifizierten Gene auf. Eine Überexpression von CbfA<sup>724-998</sup> könnte zu einer kompetitiven Verdrängung von Gesamt-CbfA aus seinen Bindungsstellen an Cluster II- und IV-Genen führen. Ein durch die Suppression von CbfA erzielter Effekt wird zusätzlich verstärkt. Dies ist für alle Gene der Cluster II und IV auch in der Überexpressionssituation in AX2 zu erkennen. Positiv von CbfA regulierte Gene werden in ihrer Transkription durch die zusätzliche Expression der CTD reprimiert, negativ regulierte aktiviert. Das ebenfalls JmjC-haltige, aus zwei Domänen bestehende Protein JcdA (DDB\_G0290765) zeigt für alle durch es regulierte Gene diese Art der Regulationsweise. Exprimiert man die in diesen Fall am Amino-Terminus gelegene, DNA-bindende Domäne von JcdA im Wildtyp über, kommt es hier zu ähnlichen Effekten wie bei der Überexpression von CbfA-CTD. Eine zusätzliche Expression von der hier am C-Terminus

gelegenen, JmjC-haltigen Domäne zeigt eher verstärkende Wirkung auf die Regulation der Gene (Accari und Fisher 2007).

Die vierte Möglichkeit für den Einfluss der CTD auf die Regulation durch CbfA ist die Eigenständigkeit der Domäne als regulatorische Einheit. Für die Erfüllung der Aufgaben ist einzig der Carboxy-terminale Teil notwendig. Dies ist für die in Cluster III und V zusammengefassten Gene zu vermuten. Ihre durch CbfA-Suppression bedingte Transkriptionsänderung wird durch die Expression der CTD umgekehrt (siehe Abbildung 14). Dabei sind sowohl aktivierende als auch reprimierende Regulationen zu beobachten. Unter den aktivierten Genen des Clusters III befinden sich Vertreter des Zellwand- und des Peptidoglykanstoffwechsels sowie Proteine mit Lysozymaktivität (Abbildung 18 A). Dies sind alles Prozesse, die in den Abbau von Bakterien involviert sein könnten. Die reprimierten Gene des Clusters V sind über einige grundlegende Prozesse vegetativ wachsender Zellen, wie Drogentransport, Purin- und Aminosäurestoffwechsel, verteilt (Abbildung 18 B).

Die Expression der TRE5-A-Elemente zeigt eine Abhängigkeit in Bezug auf die Plusstrang als auch der Minusstrang-RNA von CbfA (siehe Kapitel 1.9 und Abbildung 10). Dabei kann die Bildung des Minusstranges direkt durch die Interaktion von CbfA mit dem als Promotor agierenden C-Modul in Zusammenhang gebracht werden (Geier *et al.* 1996, Horn *et al.* 1999). Es kann mit Hilfe eines strangunspezifischen *Northern-blots* (Abbildung 10) die Abhängigkeit dieser Expression alleine von der CTD gezeigt werden. Sonstige Proteinbestandteile sind für diese Regulation nicht notwendig. Diese CTD-abhängige Regulation wird durch das *Microarray*-Experiment zusätzlich belegt. Das Gen DDB\_G0293690, welches für ein TRE5-A-Protein codiert, kann den Cluster III zugeordnet werden (Kapitel 3.3 und Anhang B). Seine durch CbfA-Suppression verringerte Expression wird durch die CTD wieder hergestellt. Ähnlich wie im *Northern-blot* ist die CTD bedingte Expression höher, was entweder durch die Überexpression bedingt ist oder auf eine Einschränkung der Funktion durch den N-Terminus verweist.

Es existieren folglich einige Gene, deren Expression von der Carboxy-terminale Domäne von CbfA ohne Beteiligung des restlichen Teiles des Faktors reguliert werden. Die CTD ist eine genregulatorische Einheit.

### 4.3 Einfluss des Kernmotives des AT-Hakens und seiner umgebenden Region auf die CTD-vermittelte Regulation

CbfA kann aufgrund seiner bekannten regulatorischen Eigenschaften sowie der Fähigkeit zur DNA-Bindung als Transkriptionsfaktor angesehen werden (Horn *et al.* 1999, Winckler *et al.* 2004, Siol *et al.* 2006a). Dies wird durch die Daten des *Microarray*-Experimentes deutlich belegt (siehe Kapitel 3.2). Die in Abbildung 14 dargestellte und in Abschnitt 4.2 diskutierte Clusteranalyse deutet auf die Eigenständigkeit der CTD als genregulatorische Einheit für die Cluster III und V hin.

Ein wichtiges Merkmal vieler Transkriptionsfaktoren ist die Translokation in den Zellkern. Dafür wird eine Kernlokalisationssequenz (NLS) auf Aminosäureebene oder strukturelle Eigenschaften wie Sekundärstruktur oder Aminosäuremodifizierungen benötigt (Imamoto 2000). Die genutzten Peptidsequenzen können dabei in mehrere Klassen eingeordnet werden. Neben dem „klassischen“ KKRK-Motiv ist vor allem die bipartite (zweigeteilte) NLS von Bedeutung (u.a. Cristophe *et al.* 2000). Diese besitzt das Sequenzmotiv  $bb-X_{10}-(b_{3-4} X_{1-2})$ , wobei b für die basischen Aminosäuren Arginin und Lysin und X für jede beliebige Aminosäure steht. Dieses Strukturelement ist in CbfA zweimal vorhanden. K<sub>488</sub>KIKCHRCEKRFKKF<sub>502</sub> liegt im Bereich des zweiten Zinkfingers, K<sub>798</sub>KRPVLDNDKPKGR<sub>812</sub> in der Carboxy-terminalen Domäne. Beide potentiellen NLS besitzen einen sehr kurzen variablen Bereich. Die in Winckler *et al.* (2001) diskutierte Kernlokalisierungssequenz liegt nach Revision des GenBank®-Eintrages AF052006.2 nicht mehr in der codierenden Sequenz von CbfA. Um die Funktionalität der CTD-NLS zu testen, wurden mehrere Aminosäureaustausche in die codierende Sequenz des GFP-CTD-exprimierenden Konstrukts pPT132 Cmbf1-2 eingeführt. Wie in Kapitel 3.9 behandelt und in Abbildung 21 zu erkennen, sind sowohl der vorderere (gezeigt durch K799N R800G) als auch der hintere Teil (R812A) an der Kernlokalisierung der CTD beteiligt, wobei der Austausch von K799 und R800 eine größere Wirkung auf die Kernlokalisierung besitzt. K<sub>798</sub>KRPVLDNDKPKGR<sub>812</sub> kann daher zumindest für die CTD als Lokalisation bestimmendes Strukturelement identifiziert werden. Die Untersuchung von der im Bereich des Zinkfingers liegenden NLS sowie der Einfluss der Carboxy-terminalen auf die Lokalisation von Gesamt-CbfA sind Gegenstand aktueller Arbeiten.



Durch den Aminosäureaustausch P801K und das so entstehende „klassische“ K<sub>798</sub>KRK-Motiv (Cristophe *et al.* 2000) kann die Funktionalität des vorderen Teiles der NLS so gesteigert werden, dass trotz gleichzeitiger Änderung von R812 keine cytoplasmatische Lokalisation mehr beobachtet wird. Dies erlaubt die Untersuchung der für den AT-Haken funktionsvermittelnden Aminosäure Arginin 812 (Arvind und Landsman 1998).

Für die regulatorische Aktivität eines Transkriptionsfaktors sind diverse Funktionsweisen denkbar. Aufgrund des in CbfA<sup>724-998</sup> enthaltenen AT-Hakens ist eine direkte Bindung an die DNA zu vermuten. Neben den in dieser Arbeit erzielten Ergebnissen weist folgende Argumentationskette darauf hin. Es ist bekannt, dass AT-Haken-haltige Proteindomänen die kleine Furche AT-reicher DNA binden (Arvind und Landsman 1998). Diese Interaktion kann durch Distamycin, welches die kleine Furche AT-reicher DNA durch Bindung blockiert (Wemmer 2001), verhindert werden. Dies gilt auch für die *in vitro*-Bindung von CbfA an das C-Modul von TRE5-A (Horn *et al.* 1999).

Um die Beteiligung des AT-Hakens an den genregulatorischen Eigenschaften von CbfA-CTD *in vivo* zu untersuchen, wurde eine kernlokalisierte Mutante (P801K) mit einer AT-Haken-Mutante (P801K R812A) auf die Fähigkeit zur Regulation der repräsentativen Gene aller Cluster (Kapitel 3.3) hin überprüft. Das Ergebnis der qRT-PCR (Abbildung 22) zeigt deutlich eine Abhängigkeit des Einflusses der CTD auf alle Cluster III-Gene sowie auf das Cluster V-Gen DDB\_G0275689 von dem AT-Haken. Für Cluster I, II und VI ist kein Einfluss dieser Mutation erkennbar, was in den CTD-unabhängigen Clustern I und VI zu erwarten ist. Das Verhalten des Cluster II-Gens deutet nicht auf eine Beteiligung des AT-Hakens an den Regulationsmodus zumindest für dieses Gen hin. Protein-Protein-Interaktionen sind hier ein denkbarer Mechanismus. Das Cluster IV-Gen verhält sich unerwartet. Nimmt man die durch die Expression von CTD in JH.D erzielten Effekte als Grundlage, ist hier von einer Konkurrenz zwischen CbfA und dem CTD-Fragment um eine Bindungsstelle auszugehen (siehe Kapitel 4.2). Handelt es sich dabei um eine DNA-Interaktion, sollte die AT-Haken-Mutante der CTD diese Konkurrenz zugunsten von Wildtyp-CbfA verschieben, da diese unmutiert besser an die DNA bindet. Der Effekt der Überexpression der CTD sollte geringer werden, sich also einen Wert von eins annähern. Vermittelt die CTD durch Protein-Protein-Interaktionen seine Funktion, sollte sich

ähnlich wie im Cluster II durch R812A nichts verändern, sieht man von dem sehr unwahrscheinlichen Fall ab, dass Arginin 812 an der Protein-Protein-Interaktion beteiligt ist. Da dies beides nicht der Fall ist, wäre die einfachste Erklärung eine Abhängigkeit des Gens von nicht erfassbaren Faktoren, welche durch leichte Unterschiede in den Klonen und den Kulturen zu begründen sind. Für das Cluster V-Gen DDB\_G0279751 ist deutlich eine Wirkung von P801K R812A im Vergleich zum nicht mutierten Protein zu sehen ( $p < 0,002$ ). Da aber die zum Vergleich genutzte Mutante P801K so gut wie keine Regulation zeigt, bleibt hier die Wirkung des AT-Hakens auf dieses Gen offen.

Es ist auffällig, dass die Wirkung der Mutation des zentralen Arginins für die Cluster III-Gene DDB\_G0282815 und DDB\_G274181 nicht ausreichend ist, um die Fähigkeit der CTD zur Genregulation ganz zu eliminieren. Nimmt man dazu das Wissen über die Beteiligung weiterer basischer Aminosäuren an der DNA-Bindung von Typ II-AT-Haken (Huth *et al.* 1997, Reeves *et al.* 2001 a, Reeves 2001 b) hinzu, stellt sich die Frage, ob dies beim Typ III-AT-Haken von CbfA ähnlich ist. Dazu wurden die dem zentralen GRP-Motiv benachbarten Lysine zu Alaninen mutiert. Das Lysin an der Stelle -2 vom zentralen Arginin 812 gerechnet, ist in der in Arvind und Landsman 1998 veröffentlichten Sequenzanalyse der Typ III-AT-Haken fast durchweg ein Arginin. Dies gilt auch für den Typ II-AT-Haken, welcher sich nur im etwas weiter entfernten Umfeld vom Typ III unterscheidet. Es ist eine verlockende Hypothese, dass der Gebrauch von Lysinen anstelle von Argininen an diesen konservierten Positionen eine Anpassung an den hohen AT-Gehalt von *D. discoideum* darstellt. In Abbildung 24 ist der Einfluss der basischen Aminosäuren auf die AT-Haken-vermittelte Regulation von GFP-CbfA<sup>724-998</sup> zu erkennen. Betrachtet man nur die Cluster III-Gene, so ist eine statistisch signifikante Verringerung der regulatorischen Aktivität für das Lysin 815 (+3) zu erkennen, welche sogar für das durch die Mutationen P801K R812A bereits stark in seiner genregulatorischen Funktion eingeschränkte Gen DDB\_G0286953 gilt. Der Austausch der für den im HMG-I-Protein vorhandenen AT-Haken wichtigen basische Aminosäure an der Position -2 (Huth *et al.* 1997), welche bei CbfA der Aminosäure Lysin 810 entspricht, zeigt keinen Einfluss auf die genregulatorische Funktion der CbfA-CTD. Eine leichte Reduktion der Regulation ist für das Gen DDB\_G0282815 auch für das Lysin 808 (-4) zu erkennen.

An der Regulation der Cluster III-Gene ist sehr wahrscheinlich das Lysin 815 beteiligt. Ein diesen Effekt erklärender, verstärkter Abbau des mutierten Proteins scheint nicht zu erfolgen, die Intensität der Fluoreszenz und der Proteinmenge im *Western-Blot* blieb im gesamten Experiment gleich. Die Beteiligung des Lysin 815 kann nicht durch die Bindungsversuche der mutierten CTD-Varianten an den Modelpromotor von *acaA* belegt werden (Abbildung 26). Das Gen *acaA* ist aufgrund der in Kapitel 3.8 gezeigten Ergebnisse nicht ausschließlich von der Carboxy-terminalen Domäne abhängig. Es ist per Definition kein Cluster III-Gen. Die Promotoren vorhandener Cluster III-Gene sind aber im Bezug auf ihre mögliche Bindung an CTD noch nicht hinreichend untersucht.

Ein Einfluss der Mutationen der Lysine in der Umgebung von Arginin 812 für die Cluster V-Gene wird aufgrund der bereits geringen Regulationsfähigkeit der Mutante P801K R812A nicht erwartet. Entgegen dieser Vorhersage zeigen beide Gene eine gesteigerte reprimierende Wirkung durch die Mutation des Lysins 815. Dies würde dem regulatorischem Level des unmutierten GFP-CbfA<sup>724-998</sup> entsprechen (vergleiche Abbildung 15 A und Abbildung 22). Dieser Effekt zeigt sich auch für die Lysine 808 (-4) und 818 (+6) auf die Regulation des Gens DDB\_G275689. Die dargestellten Regulationen sind in ihrer Art und Weise durch weitere Klone in ähnlichem Maß bestätigt. Eine simple Bindung der CTD durch den AT-Haken an die DNA liefert keine augenscheinliche Möglichkeit zur Erklärung dieses Phänomens. Auch die Regulation durch Protein-Protein-Interaktionen scheint hier nicht ausreichend dieses Problem zu erklären.

Es ist zu vermuten, dass sich die CTD-abhängige Genregulation in Bezug auf aktivierte und reprimierte Gene unterscheidet.

---

#### 4.4 Suche nach Interaktionspartnern der Carboxy-terminalen Domäne

Unterschiedliche Funktionen ein und desselben Proteins werden in der Regel durch die interagierenden Proteine vermittelt. In der Annahme, dass die Carboxy-terminale Domäne von CbfA als Transkriptionsfaktor wirkt, sind hier vor allem andere die Transkription beeinflussende Proteine zu nennen. Für kernständige Proteine allgemein, so auch für andere JmjC-haltige Proteine, sind auch Interaktionen außerhalb der direkten Transkriptionsregulation denkbar. Die JmjC-Domäne von Gis1 aus *Saccharomyces cerevisiae* dient als Plattform für Protein-Protein-Interaktionen, die zur Sumoylierung von Komponenten dieses Komplexes führen. (Tronnersjö *et al.* 2006). Der beschriebene Interaktionspartner Nfi1/Siz2 (Johnson und Gupta 2001, Takahashi *et al.* 2001 und 2003) zeigt dabei in dem Bereich der Ligation von Sumo vermittelnden Aminosäuren knapp 50% Sequenzähnlichkeit (24% Identität) zu dem *D. discoideum* Protein DDB\_G0274743. Auch ein potentieller Vertreter der Sumo-Proteinfamilie ist in *D. discoideum* vorhanden (DDB\_G0286189).

Um die Suche nach Interaktionspartnern der CTD sowie anderer *D. discoideum*-codierter Proteinfragmente durchzuführen, wurde sich für das Hefe-2-Hybrid-System entschieden. Vorteile dabei sind die im Vergleich zu anderen Methoden hohe Sensitivität bei einem relativ niedrigen Preis, Einschluss posttranskriptioneller Modifikationen sowie die leichte Zugänglichkeit der für das interagierende Protein codierenden cDNA für weitere Analysen (Causier und Davies 2002). Einer der Nachteile ist die Notwendigkeit des Vorhandenseins einer komplexen cDNA-Bibliothek für den Zielorganismus. Für *D. discoideum* existierten zum Zeitpunkt dieser Arbeit zwei Banken, welche bereits Interaktionskomplexe offenlegten. Die Interaktion zwischen dem Calcium-bindenden Protein 3 (DDB\_G0282613) mit Actin 8 (DDB\_G0269234) wurde in einer pACT2-basierenden Entwicklungsbank der Zeitpunkte vier bis achtzehn Stunden gefunden (Chang *et al.* 2005). Da das Interesse aber auf vegetativen Regulationen und damit verbundenen Interaktionen aller zur Untersuchung vorgesehener Elemente liegt, wäre diese Bank nur bedingt geeignet gewesen. Die zweite existierende Bank basiert ebenfalls auf den Vektor pACT2 und zeigte die Interaktion zwischen Thioredoxin 1 (DDB\_G0294447) und Substratproteinen (Brodegger *et al.* 2004). Die Komplexität dieser Bank ist aber mit unter 300 primären Transformanten sehr gering.

Mit Hilfe des **Matchmaker™ Library Construction & Screening** System (Clontech) sind in dieser Arbeit zwei Banken aus der mRNA vegetativ wachsender Zellen erstellt worden. Beide Banken zeigen eine Komplexität von jeweils über  $2 \cdot 10^6$  Transformanten (vergleiche Tabelle 7). Dies ist mit der Komplexität kommerzieller, aus Säugetiergewebe erstellter Banken vergleichbar. Die Rate von Transformanten mit *Dictyostelium* cDNA-Integrationen liegt in einen annehmbaren Bereich von 70 bis 80%. Frühere Versuche mit *E. coli*-basierenden Systemen resultierten entweder in wenigen unabhängigen Transformanten oder in einer hohen Rate cDNA-freier Vektoren (u.a. Brodegger *et al.* 2004, Versuche im Rahmen dieser Arbeit). Die Suche nach Genfragmenten in Vektorpräparationen der Banken zeigt das Vorhandensein von jeweils 19 von 22 gesuchten Genen (Abbildung 28). Vergleicht man beide Datensätze, so werden insgesamt 20 Gene gefunden, wobei die zwei nicht repräsentierten Gene in AX2 schwach exprimiert sind. Auch verschiedene Fragmente von *cbfA* sind in den Banken nachweisbar. Die Verteilung der sequenzierten Klone in Bezug auf die Entfernung zum 3'-Ende unterscheidet sich aufgrund des verwendeten Primers der Erststrangsynthese (Anhang F). Die Möglichkeit, lange komplementäre DNAs zu klonieren, ist mit der oligo (dT)-Methode am größten. Das Vorhandensein von am 5'-Ende der cDNA codierten N-terminalen Interaktionsdomänen ist in der (dN)<sub>6</sub>-Bank wahrscheinlicher. Im Laufe der Arbeit sind deshalb Interaktionsmatings mit beiden Banken in Kombination durchgeführt worden.

In sieben Interaktionsmatings konnte keine initial gefundene Protein-Protein-Interaktion für CbfA<sup>793-1000</sup> bestätigt werden (siehe Tabelle 8). 93 Klone zeigen eine Selbstaktivierung der Beute (wie in Abbildung 29 A), wovon mindestens 48 drei näher untersuchten Konstrukten zugeordnet werden können (Tabelle 9). Angesichts über  $1 \cdot 10^8$  überprüften Ereignissen und der Anzahl der Matings ist dies eine annehmbare Zahl (Rutushny und Golemis 2008, Mitrofanova *et al.* 2009). Die meisten der Selbstaktivierer sind dem Mating 5 zuzuordnen, in dem eine geringere Stringenz angewandt wurde. Dies zeigt den Konflikt von niedriger Stringenz mit vielen Selbstaktivierern, aber auch potentiell schwachen Interaktionen, gegenüber einer hohen Stringenz mit wenig Selbstaktivierern und der Möglichkeit schwache, Interaktionen zu übersehen (diskutiert u.a. in Rutushny und Golemis 2008).

## 4.5 Phylogenie von CbfA

Die Wirkung von CbfA als Transkriptionsfaktor für grundlegende Prozesse des vegetativen Wachstums in *Dictyostelium discoideum* lässt vermuten, dass in nahe verwandten Organismen mit ähnlicher Lebensweise orthologe Proteine zu CbfA existieren. In allen untersuchten *Dictyosteliales* konnten diese Proteine nachgewiesen werden (Abbildung 30). Des Weiteren konnten auch orthologe Proteine zu dem CbfA-paralogen CbfB gefunden werden. Die *cbfB*-orthologen Gene zeigen untereinander eine wesentlich höhere Ähnlichkeit als zu den entsprechenden *cbfAs* aus dem jeweiligen Organismus. Ähnlichkeiten in der Sequenz zwischen CbfA und CbfB beschränken sich auf den Bereich der JmjC-Domäne und den Zinkfingern. CbfA-Proteine zeigen untereinander eine etwas höhere Sequenzhomologie als CbfB-Proteine. Eine Datenbankanalyse deckte eine dritte Klasse von CbfA-ähnlichen Proteinen auf. Diese kommen in den *Ascomycota*, jedoch nicht in den *Metazoa* (Tiere) oder den *Viridiplantae* (Pflanzen) vor. Die CbfA-ähnlichen Proteine aus den filamentösen Pilzen zeigen eine Sequenzähnlichkeit in ihren JmjC-Domänen und in den Zinkfingern. Dieser Cluster von Genen ist phylogenetisch in etwa gleich weit von den dictyosteliden CbfAs und CbfBs entfernt (Abbildung 30).

Der Carboxy-terminale Teil aller untersuchten Gruppen CbfA-ähnlicher Proteine unterscheidet sich sehr stark. Ein gewisses Maß an Sequenzähnlichkeit ist nur unter den CbfAs der *Dictyosteliales* zu erkennen (Abbildung 31). Dieser Bereich umfasst die Aminosäuren 793 bis 1.000 von *D. discoideum*-CbfA, also der in dieser Arbeit als „CTD“ charakterisierten Domäne. Für diese Domäne konnte in dieser Arbeit eine eigenständige genregulatorische Funktion nachgewiesen werden. Die Carboxy-terminalen Domänen evolutionär älterer *Dictyosteliales* (Schaap *et al.* 2006) zeigen Verlängerungen der CTD von unterschiedlichem Ausmaß. Die C-terminalen Abschnitte der CbfB-Proteine zeigen keine Sequenzhomologien, nur in den beiden Organismen *Polysphonylium pallidum* und *Acytostelium subglobosum* ist ein kurzer Bereich konservierter Sequenz erkennbar. Die CTDs der CbfA-ähnlichen Proteine aus Pilzen zeigen keine Ähnlichkeit aufgrund ihrer Primärsequenz.

Anhand dieser Daten kann folgendes Modell zur Entstehung der CbfA-ähnlichen Proteine abgeleitet werden. In den gemeinsamen Vorgänger von Pilzen und Dictyosteliden gab es ein

CbfA-ähnliches Protein. Es bestand mindestens aus einer JmjC-Domäne und zwei Zinkfingern. Über das Vorhandensein einer Carboxy-terminalen Domäne kann nur spekuliert werden. Vom Standpunkt der nachgewiesenen Funktion der *D. discoideum* CbfA-CTD kann dies vermutet werden. Nach der Trennung der Pilze und der *Amoebozoa* kam es zu einer Genduplikation in den *Dictyosteliales*. Der Amino-terminale der drei Gruppen der CbfA-ähnlichen Proteine blieb dabei bis auf eine zusätzliche N-terminale Verlängerung der CbfBs erhalten. Dies weist auf eine sequenzkonservierende Funktion dieses Bereiches hin, auf der Hand liegt hier die Transkriptionsregulation durch Demethylierung von Histonen (Tsukada *et al.* 2006, Klose *et al.* 2006, Chang *et al.* 2007, Seward *et al.* 2007).

Die Sequenz der CTD scheint sich diesbezüglich frei entwickelt zu haben. Filamentöse Pilze und CbfB-ähnliche Proteine zeigen keine konservierten Bereiche. CbfA-ähnliche Proteine aus Dictyosteliden haben eine „minimale“ Carboxy-terminale Domäne erhalten. Sie entspricht der CTD der Gruppe 4 Organismen wie *D. discoideum* und zeigt neben den AT-Haken zwei weitere Bereiche hoher Sequenzhomologien (Abbildung 31). Es ist zu vermuten, dass diese CTDs homologe Funktionen besitzen. Die beobachteten Erweiterungen der CTDs von weniger abgeleiteten *Dictyosteliales*-CbfAs könnten zusätzliche Funktionen vermitteln oder bereits vorhandene modulieren.

Die Variabilität der Carboxy-terminalen Domäne CbfA-ähnlicher Proteine könnte eine Anpassung an die unterschiedliche Regulation von Gennetzwerken unter dem Aspekt sich entwickelnder Zellfunktionen sein.

#### **4.6 Röntgenkristallographische Strukturaufklärung von CbfA<sup>793-1000</sup>**

Die Primärsequenz der Carboxy-terminalen Dömäne von CbfA aus *Dictyostelium discoideum* zeigt nur Ähnlichkeiten zu der Sequenz von CbfA-Orthologen anderer *Dictyosteliales* (Kapitel 3.17). Eine Funktionsvorhersage aufgrund von Sequenzverwandtschaft zu bekannten Strukturen lässt sich, bis auf die Ausnahme des AT-Hakens und der benachbarten Kernlokalisierungssequenz, nicht treffen. Der AT-Haken vermittelt die Interaktion zwischen Proteindomäne und DNA durch die Bindung der kleinen Furche der DNA durch ein prominentes Arginin (Arvind und Landsman 1998). An dieser Bindung sind weitere basische Aminosäuren

beteiligt (Huth *et al.* 1997). Für den AT-Haken von CbfA-CTD konnte neben dem zentralen Arginin an der Position 812 noch das Lysin 815 als an der Regulation beteiligte Aminosäure identifiziert werden (Kapitel 3.11). Um Einsichten in die höhere Struktur von CbfA<sup>793-1000</sup> zu bekommen, sollte das Faltungsmuster der CTD gelöst werden.

Dazu wurde eine Reihe von Optimierungsschritten durchgeführt (Kapitel 3.18 und 3.19). Der genutzte Ansatz erbrachte ausreichend gereinigte Mengen rekombinanter CTD. Dieses Protein zeigte in einem Funktionalitätstest DNA-bindene Eigenschaften vergleichbar mit CbfA und homolog exprimierter GFP-CbfA<sup>724-998</sup> (Kapitel 3.20). Weiterführende Arbeiten belegten die Homogenität der Präparation, gemessen am Laufverhalten in einer Gelfiltration, sowie die Strukturierung der Probe in einer eindimensionalen Kernspinresonanzspektroskopie (Kapitel 3.21). Die Proteinaufreinigung zeigte nach längerer Inkubation proteolytische Aktivität. Nach vier Monaten bei 18°C war das gesamte Protein in mindestens vier Fragmente gespalten (Abbildung 39 A). Tests mit diversen Proteasen ergaben mit Thermolysin, welches nach den Aminosäuren Leucin, Isoleucin, Valin und Methionin schneidet (Applegate und Reisler 1983), mehrere Fragmente. Die entstandenen Proteinteile sind in Abbildung 39 B dargestellt.

Die Identifizierung der proteolytischen Fragmente über ihren N-Terminus und ihre Masse zeigten zwei der Proteolyse zugängliche Bereiche innerhalb der CbfA-CTD. Dabei wird oft angenommen, dass diese Bereiche in räumlich wenig strukturierten *Loops* liegen (Goldshleger *et al.* 1995, Singh *et al.* 2001). Neuere Arbeiten z.B. am *Escherichia coli* Proteom zeigen aber auch eine Anhäufung dieser Schnitte in  $\alpha$ -Helices (Timmer *et al.* 2009). Die in Abbildung 39 C zugeordneten Fragmente wurden mit dem Ergebnis der Sekundärstrukturvorhersage **PSIPRED** (<http://bioinf.cs.ucl.ac.uk/psipred/> Bryson *et al.* 2005 nach der Methode von Jones 1999) verglichen. Dabei zeigte sich, dass der proteolysesensitive Bereich der CbfA-Aminosäuren 815 bis 817 in einer wenig strukturierten Region liegt. Dies wäre mit Beobachtungen an dem Typ II AT-Haken des HMG-I-Proteins identisch. Erst durch die Bindung der DNA kommt es zur Ausbildung einer definierten Struktur des AT-Hakens von HMG-I (Huth *et al.* 1997). Es ist zu vermuten, dass der Bereich um den AT-Haken zum Erkennen der Zielsequenzen eine gewisse Flexibilität benötigt, welche durch Bindung an diese stabilisiert wird. Aufgrund der DNA-bindenden Funktion ist der AT-Haken in allen CbfA-CTDs konserviert (Abbildung 31).



Der zweite proteolysesensitive Bereich befindet sich im Bereich der Aminosäure 968. Dieser liegt laut Vorhersage als  $\alpha$ -Helix vor. Die Homologie zu anderen CbfA-CTDs ist hier geringer als im Bereich des AT-Hakens.

## 4.7 Ausblick

Die vorliegende Arbeit über den C-Modul-bindenden Faktor A aus *Dictyostelium discoideum* im Allgemeinen und die Funktion und Struktur seiner Carboxy-terminalen Domäne im Speziellen beleuchtet einige wichtige Sachverhalte, welche in diesem Zusammenhang interessant erschienen. Erwartungsgemäß beantworten die durchgeführten Experimente nicht alle Fragen und werfen neue auf. Nachfolgend werden einige Fragestellungen skizziert, welche zum Teil bereits in der Arbeitsgruppe bearbeitet werden.

Welche Gene werden von CbfA neben den 162 im Rahmen des *Microarray*-Experimentes identifizierten Genen noch reguliert? Erste Abschätzungen anhand der Solexa-Daten weisen auf 250 bis 500 CbfA-abhängige Gene hin. Aktuell wird der Stoffwechselweg der *De-novo*-Synthese von Purinen als Modell mittels qRT-PCR untersucht. Des Weiteren wird der Versuch unternommen, Carboxy-terminale Domänen von CbfAs anderer *Dictyosteliales* in JH.D zu exprimieren. Wird dies toleriert und wenn ja gibt es Gene, die dadurch reguliert werden?

Welche der von der Expression der Carboxy-terminalen Domäne unabhängigen Gene der Cluster I und VI lassen sich durch die Expression des Amino-terminalen Teils von CbfA eigenständig regulieren? Kann dabei die Funktion der NTD auf ihre Bestandteile JmjC-Domäne, Zinkfinger-Bereich und asparaginreiche Domäne zugeordnet werden?

Zur Charakterisierung der für das Bindevverhalten der Carboxy-terminalen Domäne verantwortlichen Bereiche sollten Modellpromotoren innerhalb der Cluster III und V-Gene festgelegt werden. Der untersuchte Promotor von *acaA* eignet sich wegen der Einordnung in Cluster I nur begrenzt.

Da sowohl aktivierende als auch reprimierende Genregulationen durch die CbfA-CTD vermittelt werden, lassen sich Protein-Protein-Interaktionen mit unterschiedlichen Partnern vermuten. Da mit dem angewandten Hefe-2-Hybrid-System keine nachgewiesen werden konnten, werden alternativ biochemische Methoden genutzt. Denkbar ist dabei die Aufreinigung des interagierenden Komplexes durch Affinitätschromatographie des markierten CbfAs und Identifizierung der Interaktionspartner durch Massenspektroskopie von Peptidfragmenten.

Zur Aufklärung der Struktur der Carboxy-terminalen Domäne stehen mehrere Optionen zur Verfügung. Für weitere Versuche der röntgenkristallographischen Strukturbestimmung wird der flexible Bereich um den AT-Haken ausgeschlossen. Für diese Struktur existieren bereits mehrere Strukturen (u.a. Huth *et al.* 1997), so dass eine Modellierung dieses Bereiches an die restliche Struktur möglich ist. Ferner ist die Strukturlösung durch zweidimensionale Kernspinresonanzspektroskopie (2D-NMR) denkbar.

## 5 Zusammenfassung

Der in der vorliegenden Arbeit untersuchte C-Modul-bindende Faktor A (CbfA) aus *Dictyostelium discoideum* wurde durch seine Interaktion auf genomischer Ebene mit einem Teil des 3'-untranslatierten Bereiches des Retrotransposons TRE5-A erstmalig beschrieben. Bisherige Arbeiten über CbfA zeigten eine Beteiligung dieses Faktors an der Regulation der multizellulären Entwicklung von *Dictyostelium discoideum* und sein Wirken als Transkriptionsfaktor. In der vorliegenden Arbeit wurden mit Hilfe der *Microarray*-Technik 162 Gene identifiziert, die in der CbfA-depletierten Mutante JH.D im Vergleich zum Wildtyp differenziell exprimiert werden. Sowohl in den durch CbfA in ihrer Transkription aktivierten als auch reprimierten Genen konnten mehrere statistisch signifikante Anreicherungen innerhalb biologischer Funktionen und Prozesse identifiziert werden. Dazu zählen die positiv durch CbfA beeinflusste Aufnahme und Degradation von Bakterien sowie die Reprimierung von Aminosäure- und Nukleotidstoffwechsel durch CbfA. Die erlangten Daten sind mit zwei durch RNA-Sequenzierung erstellten Transkriptomanalysen sowie punktuell mit quantitativen RT-PCR-Messungen verglichen worden. Diese Vergleiche ergaben 52 Gene, welche über alle Experimente hinaus eine deutliche Abhängigkeit von CbfA zeigten.

Des Weiteren wurden anhand einer *Microarray*-Transkriptomanalyse die Funktionen der Carboxy-terminalen Domäne (CTD) von CbfA für die Regulation CbfA-abhängiger Gene untersucht. Die Experimente bestätigen eine CTD-abhängige Regulation der Expression des Retrotransposons TRE5-A, für die der Rest von CbfA nicht erforderlich ist. Es konnten sechs Gruppen von CbfA-abhängigen Genen identifiziert werden. Diese Cluster von Genen unterschieden sich neben ihrer Abhängigkeit von CbfA in ihrer Reaktion auf die Expression der Carboxy-terminalen Domäne von CbfA. Die Cluster III und V zeigten eine transkriptionelle Abhängigkeit ihrer Expression von der CTD in gleicher Art und Weise wie von dem Gesamt-CbfA. Demzufolge ist die Carboxy-terminale Domäne von CbfA eine vom Rest des Proteins unabhängige genregulatorische Einheit.

Zur weiteren Charakterisierung von CbfA wurde die Beteiligung des DNA-bindenden AT-Hakens an der genregulatorischen Aktivität der CTD untersucht. Es konnte neben dem bekannten zentralen Arginin des DNA-bindenden GRP-Motivs des AT-Hakens noch eine weitere Aminosäure identifiziert werden, die für die *In-vivo*-Funktion der CTD notwendig ist. Ferner wurde die Kernlokalisation der Carboxy-terminalen Domäne belegt.

Zur Strukturaufklärung der Carboxy-terminalen Domäne wurde ein Protokoll zur heterologen Expression und Aufreinigung großer Mengen CTD in Bakterien entwickelt. Die DNA-bindende Aktivität und Strukturierung der rekombinanten Proteindomäne konnte nachgewiesen werden. Die Identifizierung von spontan auftretenden, proteolytischen Fragmenten sowie der Verdau der rekombinanten Proteindomäne mit Proteasen legten die Flexibilität des AT-Hakens der Carboxy-terminalen Domäne offen, was möglicherweise die versuchte Kristallisierung der CTD verhindert hat.

Ferner wurde das Hefe Zwei-Hybrid-System zur Identifizierung von Interaktionspartnern der CTD verwendet. Dazu sind zwei komplexe cDNA-Bibliotheken erstellt worden, welche qualitativ als hochwertig charakterisiert werden konnten. In mehreren Interaktionsmatings mit insgesamt über  $1 \times 10^8$  Ereignissen konnten allerdings keine Interaktionspartner der Carboxy-terminalen Domäne von CbfA identifiziert werden.

Abschließend folgten Arbeiten hinsichtlich der Phylogenie der CbfA-Domänenstruktur und des paralogen Proteins CbfB innerhalb der *Dictyosteliales*. Im Rahmen dieser Untersuchungen wurden CbfA-ähnliche Proteine in filamentösen Pilzen identifiziert, welche sich durch ca. 25% Aminosäure-Identität im Bereich der JmjC- und Zinkfingerbereiche auszeichneten. Dem gegenüber waren die jeweiligen carboxy-terminalen Bereiche der CbfB-Proteine aus anderen Dictyosteliden und die CbfA-ähnlichen Proteine aus Pilzen nicht konserviert.

## 6 Summary

The *Dictyostelium discoideum* C-module-binding factor A (CbfA) was discovered by its interaction at the genomic level with a part of the 3'-untranslated region of the retrotransposon TRE5-A. Previous studies on CbfA showed an involvement of this protein in the regulation of multicellular development of *D. discoideum* and its function as a transcription factor. Experiments in this work using the microarray technology showed 162 genes to be differentially expressed between the wild type AX2 and the CbfA-depleted mutant JH.D. Categories representing certain biological functions and processes could be identified with statistically significant enrichment of both activated and repressed transcripts, as for example the CbfA-induced regulation of intake and degradation of bacteria as well as the negative regulation of amino acid and nucleotide metabolism. Microarray data were compared with results obtained from two RNA sequencing experiments and in part to quantitative RT-PCR measurements. Altogether 52 genes showed significant transcriptional dependence on CbfA throughout all experiments.

The function of the carboxy-terminal domain (CTD) of CbfA in the regulation of CbfA-dependent genes was investigated by microarray analysis. These experiments confirmed a CTD-dependent regulation to the expression of the retrotransposon TRE5-A, interestingly without requiring the rest of CbfA. Six groups of CbfA-dependent genes could be identified with varying response to CbfA and to the presence of CTD. The clusters III and V showed a transcriptional dependence on CTD in the same way they responded to CbfA. These results emphasize the function of carboxy-terminal Domain of CbfA as an independent gene-regulatory entity.

To further characterise CbfA, the involvement of the DNA-binding AT-hook in the gene regulatory activity of CTD was investigated. Apart from the known central arginine in the DNA-binding GRP-motive, one lysine residue within the AT-hook could be identified as critical for the *in vivo* activity of CTD. Furthermore, the nuclear localisation of the carboxy-terminal domain was shown and the nuclear localisation site was characterised.

To determine the structure of CbfA's CTD a protocol to express and purify large amounts of the protein in bacteria was developed. The functionality (i.e. DNA-binding activity) and the structuring of the recombinant domain were confirmed. The identification of spontaneously generated proteolytic fragments as well as the digestion of the recombinant CTD with certain proteases revealed the flexibility of the unbound AT-hook within the carboxy-terminal domain. This might have prevented the crystallisation of recombinant CTD.

The yeast two-hybrid system was used to search for proteins that interact with the CTD in vivo. Two high-quality cDNA libraries were generated from *D. discoideum* cells. However, although  $1 \times 10^8$  events were screened in several matings, no interacting partner of CbfA-CTD could be identified.

Finally the molecular phylogeny of the domain structure of CbfA and its paralogous protein CbfB was investigated within the dictyostelid clade. Proteins similar to CbfA with respect to the JmjC/zinc finger architecture were identified in filamentous fungi. These fungal CbfA-like proteins showed 25 % homology to CbfA and CbfB. In contrast, the carboxy-terminal regions of dictyostelid CbfB and *funagi* CbfA-like proteins were found to be highly diverged.

---

## Literaturverzeichnis

- Aasland R., Gibson T.J., Stewart A.F.** (1995) The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem. Sci.* **20**: 56-59.
- Accari S.L., Fisher P.R.** (2007) Characterisation of Kame, a novel JmjC domain-containing protein in *Dictyostelium discoideum*. *Tagungsband International Dictyostelium Conference 2007*: 96.
- Alvarez-Curto E., Rozen D.E., Ritchie A.V., Fouquet C., Baldauf S.L., Schaap P.** (2005). Evolutionary origin of cAMP-based chemoattraction in the social amoebae. *Proc. Natl. Acad. Sci.* **102**: 6385-90.
- Anjard C., Loomis W.F.** (2006) GABA induces terminal differentiation of *Dictyostelium* through a GABAB receptor. *Development* **133**: 2253-61.
- Applegate D., Reisler E.** (1983) Protease-sensitive regions in myosin subfragment 1. *Proc Natl Acad Sci.* **80(23)**: 7109-12.
- Aravind L., Landsman D.** (1998) AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nuc Acids Res* **26**: 4413–21.
- Aubry L., Maeda M., Insall R., Devreotes P. N., Firtel R. A.** (1997) The *Dictyostelium* mitogen-activated protein kinase ERK2 is regulated by ras and cAMP-dependent protein kinase (PKA) and mediates PKA function. *J. Biol. Chem.* **272**: 3883-6.
- Baldauf S.L., Roger A.J., Wenk-Siefert I., Doolittle W.F.** (2000). A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* **290**: 972-77.
- Beck P., Dingermann T., Winckler T.** (2002) Transfer RNA Gene-targeted Retrotransposition of *Dictyostelium* TRE5-A into a Chromosomal UMP Synthase Gene Trap. *J. Mol. Biol.* **318**: 273–285.
- Benjamini Y., Hochberg Y.** (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B.* **57**: 289-300.
- Bohnert M.** (2009) Evolution der Domänenstruktur des *Dictyostelium*-Transkriptionsfaktors CbfA. *Diplomschrift*.
- Borden K.L., Lally J.M., Martin S.R., O'Reilly N.J., Etkin L.D., Freemont P.S.** (1995) Novel topology of a zinc-binding domain from a protein involved in regulating early *Xenopus* development. *EMBO J.* **14**: 5947-56.
- Bozzaro S., Ponte E.** (1995) Cell adhesion in the life cycle of *Dictyostelium*. *Experientia* **51**: 1175-88.
- Brackenbury R.W., Schindler J., Alexander S. & Sussman M.** (1974) A choice of morphogenetic pathways in *Dictyostelium discoideum* induced by the adenosine analog formycin B. *J Mol Biol* **90**: 529-39.
- Bradford M.M.** (1976): A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-54.
- Brayer K.J., Segal D.J.** (2008) Keep your fingers off my DNA: protein-protein interactions mediated by C2H2 zinc finger domains. *Cell Biochem Biophys.* **50(3)**: 111-31.
- Brefeld O.** (1869) *Dictyostelium mucoroides*. Ein neuer Organismus aus der Verwandtschaft der Myxomyceten. *Abhandlungen der Senckenbergischen Naturforschenden Gesellschaft Frankfurt* **7**: 85-107.



**Brizzard B.** (2008) Epitope tagging. *Biotechniques* **44(5)**: 693-5.

**Brock D.A., Hatton R.D., Giurgiutiu D.V., Scott B., Ammann R., Gomer R.H.** (2002) The different components of a multisubunit cell number-counting factor have both unique and overlapping functions. *Development* **129(15)**: 3657-68.

**Brock D.A., Hatton R.D., Giurgiutiu D.V., Scott B., Jang W., Ammann R., Gomer R.H.** (2003) CF45-1, a secreted protein which participates in Dictyostelium group size regulation. *Eukaryot Cell* **2(4)**: 788-97.

**Brodegger T., Stockmann A., Oberstrass J., Nellen W., Follmann H.** (2004) Novel thioredoxin targets in Dictyostelium discoideum identified by two-hybrid analysis: interactions of thioredoxin with elongation factor 1alpha and yeast alcohol dehydrogenase. *Biol Chem.* **385(12)**: 1185-92.

**Bryson K., McGuffin L.J., Marsden R.L., Ward J.J., Sodhi J.S., Jones D.T.** (2005) Protein structure prediction servers at University College London. *Nucleic Acids Res.* **33(Web Server issue)**: W36-8.

**Caspi R., Foerster H., Fulcher C.A., Kaipa P., Krummenacker M., Latendresse M., Paley S., Rhee S.Y., Shearer A.G., Tissier C., Walk T.C., Zhang P., Karp P.D.** (2007) The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.* **36(Database issue)**: D623-31.

**Causier B., Davies B.** (2002) Analysing protein-protein interactions with the yeast two-hybrid system. *Plant Molecular Biology* **50**: 855-870.

**Chen Z., Zang J., Whetstone J., Hong X., Davrazou F., Kutateladze T.G., Simpson M., Mao Q., Pan C.H., Dai S., Hagman J., Hansen K., Shi Y., Zhang G.** (2006) Structural insights into histone demethylation by JMJD2 family members. *Cell* **125(4)**: 691-702.

**Chisholm R. L., Firtel, R. A.** (2004). Insights into morphogenesis from a simple developmental system. *Nature Reviews Molecular Cell Biology* **5**: 531-41.

**Chisholm R.L., Gaudet P., Just E.M., Pilcher K.E., Fey P., Merchant S.N., Kibbe W.A.** (2006) dictyBase, the model organism database for Dictyostelium discoideum. *Nuc Acids Res.* **34**: D423-7.

**Chung T., Siol O., Dingermann T., Winckler T.** (2007) Protein interactions involved in tRNA gene-specific integration of Dictyostelium discoideum non-long terminal repeat retrotransposon TRE5-A. *Mol Cell Biol.* **27(24)**: 8492-501.

**Clarke M., Gomer R.H.** (1995) PSF and CMF, autocrine factors that regulate gene expression during growth and early development of *Dictyostelium*. *Experientia* **51**: 1124-34.

**Clarke M., Maddera L.** (2006) Phagocyte meets prey: uptake, internalization, and killing of bacteria by Dictyostelium amoebae. *Eur J Cell Biol.* **85(9-10)**: 1001-10.

**Clissold P.M. und Ponting C.P.** (2001) JmjC: cupin metalloenzyme-like domains in jumonji, hairless and phospholipase A2 $\beta$ . *TRENDS in Biochemical Sciences* **26(1)**: 7-9.

**Christophe D., Christophe-Hobertus C., Pichon B.** (2000) Nuclear targeting of proteins: how many different signals? *Cell Signal.* **12(5)**: 337-41.

**Churchill G.A.** (2002) Fundamentals of experimental design for cDNA microarrays. *Nat Genet.* **32**: 490-5.

**Cotter D.A., Sands T.W., Viridy K.J., North M.J., Klein G., Satre M.** (1992). Patterning of development in *Dictyostelium discoideum* - Factors regulating growth, differentiation, spore dormancy, and germination. *Biochem. Cell Biol.* **70**: 892-919.

- Craig N.L., Craigie R., Gellert M., Lambowitz A.M. (2002) *Mobile DNA II*. Washington DC: ASM Press.
- Cvrckova F., Rivero F., Bavlínka B. (2004): Evolutionarily conserved modules in actin nucleation: lessons from *Dictyostelium discoideum* and plants. *Protoplasma* **224**: 15-31.
- DDB, FB, MGI, GOA, ZFIN curators (2001) Gene Ontology annotation through association of InterPro records with GO terms. Unpublished.
- De Lozanne A., Spudich J.A. (1987) Disruption of the *Dictyostelium* myosin heavy chain gene by homologous recombination. *Science* **236**: 1086-91.
- Devreotes P. N. (1994) G-protein-linked signaling pathways control the developmental program of *Dictyostelium*. *Neuron* **12**: 235-41.
- Dingermann T., Reindl N., Brechner T., Werner H., Nerke K. (1990) Nonsense suppression in *Dictyostelium discoideum*. *Dev Genet.* **11(5-6)**: 410-7.
- Dingermann T., Werner H., Schütz A., Zündorf I., Nerke K., Knecht D., Marschalek R. (1992) Establishment of a system for conditional gene expression using an inducible tRNA suppressor gene. *Mol Cell Biol.* **12(9)**: 4038-45.
- Dormann D., Vasiev B. and Weijer C.J. (1998) Propagating waves control *Dictyostelium discoideum* morphogenesis. *Biophys. Chem.* **72**: 21-35.
- Dutta S.K., Mandel M. (1972) Deoxyribonucleic acid base composition of some cellular slime molds. *J Protozool.* **19(3)**: 538-40.
- Early A.E., Gaskell M.J., Traynor D. and Williams, J.G. (1993) Two distinct populations of prestalk cells within the tip of the migratory *Dictyostelium* slug with differing fates at culmination. *Development* **118**: 353-62.
- Eichinger L., Lee S.S., Schleicher M. (1999) *Dictyostelium* as model system for studies of the actin cytoskeleton by molecular genetics. *Microsc Res Techniq.* **47**: 124-34.
- Eichinger L., Pachebat J.A., Glockner G., Rajandream M.A., Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q., Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F., Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P., Pilcher K., Chen G., Saunders D., Sodergren E., Davis P., Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N., Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C., Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I., Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R., Hauser H., James K., Quiles M., Madan Babu M., Saito T., Buchrieser C., Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A., Loulseged H., Mungall K., Oliver K., Price C., Quail M.A., Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M., Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A., Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y., Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C., Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R., Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A. (2005) The genome of the social amoeba *Dictyostelium discoideum*. *Nature* **435**: 43-57.
- Elrod-Erickson M., Rould M.A., Nekludova L. und Pabo C.O. (1996) Zif268 protein–DNA complex refined at 1.6 Å: a model system for understanding zinc finger–DNA interactions. *Structure* **4(10)**: 1171-80.
- Escalante R., Vicente J.J. (2000) *Dictyostelium discoideum*: a model system for differentiation and patterning. *Int J Dev Biol.* **44(8)**: 819-35.
- Farbrother P., Wagner C., Na J., Tunggal B., Morio T., Urushihara H., Tanaka Y., Schleicher M., Steinert M., Eichinger L. (2006) *Dictyostelium* transcriptional host cell response upon infection with Legionella. *Cell Microbiol.* **8(3)**: 438-56.

- Felsenstein J.** (1989) Mathematics vs. Evolution: Mathematical Evolutionary Theory. *Science* **246(4932)**: 941-2.
- Ferrer M., Chernikova T.N., Yakimov M.M., Golyshin P.N., Timmis K.N.** (2003) Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nat Biotechnol.* **21(11)**: 1266-7.
- Firtel, R. A.** (1995) Integration of signaling information in controlling cell fate decisions in *Dictyostelium*. *Genes Dev.* **9**: 1427-44.
- Franke J., Kessin R.** (1977) A defined minimal medium for axenic strains of *Dictyostelium discoideum*. *Proc Natl Acad Sci.* **74**: 2157-61.
- Friedman M.J., Shah A.G., Fang Z.H., Ward E.G., Warren S.T., Li S., Li X.J.** (2007) Polyglutamine domain modulates the TBP-TFIIB interaction: implications for its normal function and neurodegeneration. *Nat Neurosci.* **10(12)**: 1519-28.
- Gaudet P., Pilcher K.E., Fey P., Chisholm R.L.** (2007) Transformation of *Dictyostelium discoideum* with plasmid DNA. *Nat Protoc.* **2(6)**: 1317-24.
- Gerisch G., Malchow D., Roos W., Wick U.** (1979) Oscillations of cyclic nucleotide concentrations in relation to the excitability of *Dictyostelium* cells. *J. Exp. Biol.* **81**: 33-47.
- Gietz R.D., Woods R.A.** (2001): Genetic Transformation of Yeast. *BioTechniques* **30**: 816-31.
- Gilbert N., Lutz-Prigge S., Moran J.V.** (2002). Genomic deletions created upon LINE-1 retrotransposition. *Cell* **110(3)**: 315-25.
- Glöckner G., Szafranski K., Winckler T., Dingermann T., Quail M.A., Cox E., Eichinger L., Noegel A.A., Rosenthal A.** (2001) The complex repeats of *Dictyostelium discoideum*. *Genome Res.* **11(4)**: 585-94.
- Goldshleger R., Tal D.M., Karlisch S.J.** (1995) Topology of the alpha-subunit of Na,K-ATPase based on proteolysis. Lability of the topological organization. *Biochemistry* **34(27)**: 8668-79
- Hall T. M.** (2005) Multiple modes of RNA recognition by zinc finger proteins. *Current Opinion in Structural Biology* **15**: 367-73.
- Hedges D.J., Batzer M.A.** (2005) From the margins of the genome: mobile elements shape primate evolution. *Bioessays* **27(8)**: 785-94.
- Hofmann J., Schumann G., Borschet G., Gosseringer R., Bach M., Bertling W.M., Marschalek R., Dingermann T.** (1991) Transfer RNA genes from *Dictyostelium discoideum* are frequently associated with repetitive elements and contain consensus boxes in their 5'-flanking and 3'-flanking regions. *J Mol Biol.* **222**: 537-52.
- Huth J.R., Bewley C.A., Nissen M.S., Evans J.N., Reeves R., Gronenborn A.M., Clore G.M.** (1997) The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat Struct Biol.* **4(8)**: 657-65.
- Imamoto N.** (2000) Diversity in nucleocytoplasmic transport pathways. *Cell Struct Funct* **25(4)**: 207-16.
- Insall R.H., Nayler O., Kay R.R.** (1992) DIF-1 induces its own breakdown in *Dictyostelium*. *EMBO J.* **11(8)**: 2849-54.
- Insall R.H., Soede R.D.M., Schaap P., Devreotes P. N.** (1994) Two cAMP receptors activate common signaling pathways in *Dictyostelium*. *Mol. Biol. Cell* **5**: 703-11.

**Insal, R.H., Borleis J., Devreotes P.N.** (1996) The aimless RasGEF is required for processing of chemotactic signals through G-protein-coupled receptors in *Dictyostelium*. *Curr Biol* **6**: 719-29.

**International Human Genome Sequencing Consortium** (2004) Finishing the euchromatic sequence of the human genome. *Nature* **431(7011)**: 931-45.

**Jacobsen L., Calvin S., Lobenhofer E.** (2009) Transcriptional effects of transfection: the potential for misinterpretation of gene expression data generated from transiently transfected cells. *Biotechniques*. **47(1)**: 617-24.

**Jayakanthan M., Muthukumar J., Chandrasekar S., Chawla K., Punetha A., Sundar D.** (2009) ZifBASE: a database of zinc finger proteins and associated resources. *BMC Genomics* **10**: 421.

**Jennes I.** (2006) Differentielle Expressionsanalyse CbfA-regulierter Proteine während des Übergangs vom Wachstum zur Entwicklung in *Dictyostelium discoideum*. *Dissertationsschrift*.

**Jermyn K.A. and Williams J.G.** (1991). An analysis of culmination in *Dictyostelium* using prestalk and stalk-specific cell autonomous markers. *Development* **111**: 779-87.

**Jermyn K., Traynor D., Williams J.** (1996). The initiation of basal disc formation in *Dictyostelium discoideum* is an early event in culmination. *Development* **122**: 753-60.

**Johnson E.S., Gupta A.A.** (2001) An E3-like factor that promotes SUMO conjugation to the yeast septins. *Cell* **106**: 735-744.

**Johnston J.R.** (1994): Molecular genetics of yeast: a practical approach. Oxford: *IRL Press at Oxford Univ. Press*.

**Jones D.T.** (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**: 195-202.

**Joseph R.E., Andreotti A.H.** (2008) Bacterial expression and purification of interleukin-2 tyrosine kinase: single step separation of the chaperonin impurity. *Protein Expr Purif.* **60(2)**: 194-7.

**Kay R.R., Jermyn K.A.** (1983) A possible morphogen controlling differentiation in *Dictyostelium*. *Nature* **303**: 242-4.

**Kessin R.H.** (2001) *Dictyostelium*: Evolution, Cell Biology and the Development of Multicellularity. Cambridge: *Cambridge Univ. Press*.

**Kellerman K.A., McNally J.G.** (1999) Mound-cell movement and morphogenesis in *Dictyostelium*. *Dev. Biol.* **208**: 416-29.

**Kim L., Liu J., Kimmel A.R.** (1999) The novel tyrosine kinase ZAK1 activates GSK3 to direct cell fate specification. *Cell* **99(4)**: 399-408.

**Klein P.S., Sun T. J., Saxe 3rd C. L., Kimmel A. R., Johnson R. L., Devreotes P. N.** (1988) A chemoattractant receptor controls development in *Dictyostelium discoideum*. *Science* **241**: 1467-72.

**Klose R.J., Kallin E.M., Zhang Y.** (2006) JmjC-domain-containing proteins and histone demethylation. *Nat Rev Genet.* **7(9)**: 715-27.

**Knecht D.A., Cohen S.M., Loomis W.F., Lodish H.F.** (1986) Developmental regulation of *Dictyostelium discoideum* actin gene fusions carried on low-copy and high-copy transformation vectors. *Mol Cell Biol.* **6**: 3973-83.

- Konieczny A., Voytas D.F., Cummings M.P., Ausubel F.M.** (1991) A superfamily of *Arabidopsis thaliana* retrotransposons. *Genetics* **127**(4): 801-9.
- Kubohara Y.** (1999) Effects of differentiation-inducing factors of *Dictyostelium discoideum* on human leukemia K562 cells: DIF-3 is the most potent anti-leukemic agent. *Eur J Pharmacol* **381**: 57-62.
- Kumar A., Bennetzen J.L.** (1999). Plant retrotransposons. *Annu Rev Genet* **33**: 479-532.
- Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J., Higgins D.G.** (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947-48.
- Lee C.H., Jeong S.Y., Kim B.J., Choi C.H., Kim J.S., Koo B.M., Seok Y.J., Yim H.S., Kang S.O.** (2005) *Dictyostelium* CBP3 associates with actin cytoskeleton and is related to slug migration. *Biochim Biophys Acta*. **1743**(3): 281-90.
- Leung Y.F., Cavalieri D.** (2003) Fundamentals of cDNA microarray data analysis. *Trends Genet.* **19**(11): 649-59.
- Loomis W.F.** (1993). Lateral inhibition and pattern formation in *Dictyostelium*. *Curr. Topics Dev. Biol.* **28**: 1-46.
- Loomis W.F., Welker D., Hughes J., Maghakian D., Kuspa A.** (1995) Integrated maps of the chromosomes in *Dictyostelium discoideum*. *Genetics* **141**: 147-57.
- Loomis W.F.** (1998) Role of PKA in the timing of developmental events in *Dictyostelium* cells. *Microbiol Mol Biol Rev.* **62**(3): 684-94.
- Maeda M., Firtel R.A.** (1997) Activation of the mitogen-activated protein kinase ERK2 by the chemoattractant folic acid in *Dictyostelium*. *J Biol Chem* **272**: 23690-5.
- Malik H.S., Burke W.D., Eickbush T.H.** (1999) The age and evolution of non-LTR retrotransposable elements. *Mol. Biol. Evol.* **16**: 793-805.
- Mane S.P., Evans C., Cooper K.L., Crasta O.R., Folkerts O., Hutchison S.K., Harkins T.T., Thierry-Mieg D., Thierry-Mieg J., Jensen R.V.** (2009) Transcriptome sequencing of the Microarray Quality Control (MAQC) RNA reference samples using next generation sequencing. *BMC Genomics* **10**: 264.
- Mann S.K., Firtel R.A.** (1991) A developmentally regulated, putative serine/threonine protein kinase is essential for development in *Dictyostelium*. *Mech Dev* **35**: 89-101.
- Marnef A., Sommerville J., Ladomery M.R.** (2009) RAP55: insights into an evolutionarily conserved protein family. *Int J Biochem Cell Biol.* **41**(5): 977-81.
- Marschalek R., Hofmann J., Schumann G., Dinger T.** (1992a). Two distinct subforms of the retrotransposable DRE element in NC4 strains of *Dictyostelium discoideum*. *Nucl. Acids Res.* **20**: 6247-52.
- Marschalek R., Hofmann J., Schumann G., Gosseringer R., Dinger T.** (1992b). Structure of DRE, a retrotransposable element which integrates with position specificity upstream of *Dictyostelium discoideum* tRNA genes. *Mol. Cell. Biol.* **12**: 229-39.
- Martens H., Novotny J., Oberstrass J., Steck T.L., Postlethwait P., Nellen W.** (2002) RNAi in *Dictyostelium*: the role of RNA-directed RNA polymerases and double-stranded RNase. *Mol Biol Cell* **13**: 445-453.
- McClintock B.** (1948) Mutable loci in maize. *Carnegie Inst. Wash. Year Book* **47**: 155-169.

- Miller J., McLachlan A.D., Klug, A.** (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J* **4**: 1609-14.
- Mitrofanova A., Farach-Colton M., Mishra B.** (2009) Efficient and robust prediction algorithms for protein complexes using Gomory-Hu trees. *Pac Symp Biocomput.* **2009**: 215-26.
- Mehta A.D., Rief M., Spudich J.A., Smith D.A., Simmons R.M.** (1999) Single-molecule biomechanics with optical methods. *Science* **283**: 1689-95.
- Morgan C.P., Insall R., Haynes L., Cockcroft S.** (2004) Identification of phospholipase B from *Dictyostelium discoideum* reveals a new lipase family present in mammals, flies and nematodes, but not yeast. *Biochem J.* **382(Pt 2)**: 441-9.
- Mülhardt C.** (2009) Der Experimentator: Molekularbiologie / Genomics. Heidelberg: *Spektrum Akademischer Verlag*.
- Mutzel R.** (1991). Cellular slime molds: why and how to become pluricellular. *Bull. Inst. Paste* **89**: 51-58.
- Myllykallio H., Lipowski G., Leduc D., Filee J., Forterre P., Liebl U.** (2002) An alternative flavin-dependent mechanism for thymidylate synthesis. *Science* **297(5578)**: 105-7.
- Na J., Tunggal B., Eichinger L.** (2007) STATc is a key regulator of the transcriptional response to hyperosmotic shock. *BMC Genomics* **8**: 123.
- Noegel A., Welker D. L., Metz B. A., Williams K. L.** (1985). Presence of Nuclear Associated Plasmids in the Lower Eukaryote *Dictyostelium discoideum*. *Journal of Molecular Biology* **185**: 447-50.
- Olive L.S., Stoianovitch C.** (1975) The Mycetozoeans. New York: *Academic Press*.
- Otsuka H., Van Haastert P.J.** (1998) A novel Myb homolog initiates *Dictyostelium* development by induction of adenyl cyclase expression. *Genes Dev* **12**: 1738-48.
- Pace C.N., Vajdos F., Fee L., Grimsley G. und Gray T.** (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **11**: 2411-23.
- Papworth M., Kolasinska P., Minczuk M.** (2006) Designer zinc-finger proteins and their applications. *Gene* **366(1)**: 27-38.
- Pitt G.S., Milona N., Borleis J., Lin K.C., Reed R.R., Devreotes P.N.** (1992) Structurally distinct and stage-specific adenyl cyclase genes play different roles in *Dictyostelium* development. *Cell* **69(2)**: 305-15.
- Pfaffl M.W.** (2001) A new mathematical model for relative quantification in realtime RT-PCR. *Nuc Acids Res.* **29**: e45.
- Plyte S.E., O'Donovan E., Woodgett J.R., Harwood, A.J.** (1999) Glycogen synthase kinase-3 (GSK-3) is regulated during *Dictyostelium* development via the serpentine receptor cAR3. *Development* **126**: 325-33.
- Raper, K.B.** (1935) *Dictyostelium discoideum*, a new species of slime mold from decaying forest leaves. *J. Agr. Res.* **50**: 135-47.
- Raper K.B.** (1984) The Dictyostelids. Princeton: *Princeton Univ. Press*.

- Ratushny V, Golemis E.** (2008) Resolving the network of cell signaling pathways using the evolving yeast two-hybrid system. *Biotechniques* **44(5)**: 655-62.
- Reeves R., Beckerbauer L.** (2001a) HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim Biophys Acta* **1519**: 13-29.
- Reeves R.** (2001b) Molecular biology of HMGA proteins: hubs of nuclear function. *Gene* **277(1-2)**: 63-81.
- Rehm T., Huber R., Holak T.A.** (2002) Application of NMR in structural proteomics: screening for proteins amenable to structural analysis. *Structure* **10(12)**: 1613-8.
- Rot G., Parikh A., Curk T., Kuspa A., Shaulsky G., Zupan B.** (2009) dictyExpress: A *Dictyostelium discoideum* gene expression database with an explorative data analysis web-based interface. *BMC Bioinformatics* **10**: 265.
- Szafrański K., Glöckner G., Dingermann T., Dannat K., Noegel A.A., Eichinger L., Rosenthal A., Winckler T.** (1999) Non-LTR retrotransposons with unique integration preferences downstream of *Dictyostelium discoideum* transfer RNA genes. *Mol Gen Genet* **262**: 772-80.
- Sambrook J., Fritsch E.F., Maniatis T.** (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor: *Cold Spring Harbor Laboratory Press*.
- Schaap P., Winckler T., Nelson M., Alvarez-Curto E., Elgie B., Hagiwara H., Cavender J., Milano-Curto A., Rozen D.E., Dingermann T., Mutzel R., Baldauf S.L.** (2006) Molecular Phylogeny and Evolution of Morphology in the Social Amoebas. *Science* **314 (5799)**: 661-3.
- Schmidt H.A., Strimmer K., Vingron M., von Haeseler A.** (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing *Bioinformatics* **18(3)**: 502-4.
- Schumann G., Zündorf I., Hofmann J., Marschalek R., Dingermann T.** (1994) Internally located and oppositely oriented polymerase II promoters direct convergent transcription of a LINE-like retroelement, the *Dictyostelium* repetitive element, from *Dictyostelium discoideum*. *Mol Cell Biol.* **14(5)**: 3074-84.
- Seward D.J., Cubberley G., Kim S., Schonewald M., Zhang L., Tripet B., Bentley D.L.** (2007) Demethylation of trimethylated histone H3 Lys4 in vivo by JARID1 JmjC proteins. *Nat Struct Mol Biol.* **14(3)**: 240-2.
- Shaulsky G., Fuller D., Loomis W. F.** (1998) A cAMP-phosphodiesterase controls PKA-dependent differentiation. *Development* **125**: 691-9.
- Sillo A., Bloomfield G., Balest A., Balbo A., Pergolizzi B., Peracino B., Skelton J., Ivens A. und Bozzaro S.** (2008): Genome-wide transcriptional changes induced by phagocytosis or growth on bacteria in *Dictyostelium*. *BMC Genomics* **9**: 291-312.
- Singh S.K., Rozycki J., Ortega J., Ishikawa T., Lo J., Steven A.C., Maurizi M.R.** (2001) Functional domains of the ClpA and ClpX molecular chaperones identified by limited proteolysis and deletion analysis. *J Biol Chem.* **276(31)**: 29420-9.
- Siol O., Dingermann T., Winckler T.** (2006a) The C-Module DNA-Binding Factor Mediates Expression of the *Dictyostelium* Aggregation-Specific Adenylyl Cyclase ACA. *Eukaryotic Cell* **5(4)**: 658-64.
- Siol O., Boutliliss M., Chung T., Glöckner G., Dingermann T., Winckler T.** (2006b) Role of RNA polymerase III transcription factors in the selection of integration sites by the *Dictyostelium* non-long terminal repeat retrotransposon TRE5-A. *Mol Cell Biol.* **26(22)**: 8242-51.

- Siol O.** (2006c) Untersuchungen zum spezifischen Integrations-Mechanismus der TRE5-A Retrotransposons in *Dictyostelium discoideum*. *Dissertationsschrift*.
- Sucgang R., Chen G., Liu W., Lindsay R., Lu J., Muzny D., Shaulsky G., Loomis W., Gibbs R., Kuspa A.** (2003) Sequence and structure of the extrachromosomal palindrome encoding the ribosomal RNA genes in *Dictyostelium*. *Nucleic Acids Res.* **31(9)**: 2361-8.
- Sun T.J., Devreotes P.N.** (1991) Gene targeting of the aggregation stage cAMP receptor cAR1 in *Dictyostelium*. *Genes Dev* **5**: 572-82.
- Sussman R., Sussman M.** (1967) Cultivation of *Dictyostelium discoideum* in axenic culture. *Biochem. Biophys. Res. Commun.* **29**: 53-55.
- Takahashi Y., Toh-e A., Kikuchi Y.** (2001) A novel factor for the SUMO/Smt3 conjugation of yeast septins. *Gene* **275**: 223-31.
- Takahashi Y., Toh-e A., Kikuchi Y.** (2003) Comparative analysis of yeast PIAS-type SUMO ligases in vivo and in vitro. *J Biochem* **133**: 415-22.
- Takeuchi T., Yamazaki Y., Katoh-Fukui Y., Tsuchiya R., Kondo S., Motoyama J., Higashinakagawa T.** (1995) Gene trap capture of a novel mouse gene, jumonji, required for neural tube formation. *Genes Dev.* **9**: 1211-22.
- Thomason P.A., Traynor D., Cavet G., Chang W.T., Harwood A.J., Kay R.R.** (1998) An intersection of the cAMP/PKA and two-component signal transduction systems in *Dictyostelium*. *EMBO J.* **17**: 2838-45.
- Thompson J.D., Higgins D.G., Gibson T.J.** (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673-80.
- Timmer J.C., Zhu W., Pop C., Regan T., Snipas S.J., Eroshkin A.M., Riedl S.J., Salvesen G.S.** (2009) Structural and kinetic determinants of protease substrates. *Nat Struct Mol Biol.* **16(10)**: 1101-8.
- Trewick S.C., Minc E., Antonelli R., Urano T., Allshire R.C.** (2007) The JmjC domain protein Epe1 prevents unregulated assembly and disassembly of heterochromatin. *EMBO J.* **26(22)**: 4670-82.
- Tronnersjö S., Hanefalk C., Balciunas D., Hu G.Z., Nordberg N., Murén E., Ronne H.** (2006) The jmjN and jmjC domains of the yeast zinc finger protein Gis1 interact with 19 proteins involved in transcription, sumoylation and DNA repair. *Mol Genet Genomics* **277(1)**: 57-70.
- Tsukada Y., Fang J., Erdjument-Bromage H., Warren M.E., Borchers C.H., Tempst P., Zhang Y.** (2006) Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439(7078)**: 811-6.
- Tusher V.G., Tibshirani R., Chu G.** (2001) Significance analysis of microarrays applied to the ionizing radiation response. *PNAS.* **98(9)**: 5116-21.
- Urushihara H., Morio T., Saito T., Kohara Y., Koriki E., Ochiai H., Maeda M., Williams J.G., Takeuchi I., Tanaka Y.** (2004) Analyses of cDNAs from growth and slug stages of *Dictyostelium discoideum*. *Nucleic Acids Res.* **32(5)**: 1647-53.
- Wales M.E., Mann-Dean M.G., Wild J.R.** (1989) Characterization of pyrimidine metabolism in the cellular slime mold, *Dictyostelium discoideum*. *Can J Microbiol.* **35(4)**: 432-8.



- Wall P.K., Leebens-Mack J., Chanderbali A.S., Barakat A., Wolcott E., Liang H., Landherr L., Tomsho L.P., Hu Y., Carlson J.E., Ma H., Schuster S.C., Soltis D.E., Soltis P.S., Altman N., dePamphilis C.W.** (2009) Comparison of next generation sequencing technologies for transcriptome characterization. *BMC Genomics* **10**: 347.
- Wang N., Shaulsky G., Escalante R., Loomis W.F.** (1996) A two-component histidine kinase gene that functions in *Dictyostelium* development. *EMBO J* **15**: 3890-8.
- Wang J., Virta V.C., Riddelle-Spencer K., O'Halloran T.J.** (2003) Compromise of clathrin function and membrane association by clathrin light chain deletion. *Traffic* **4**: 891-901.
- Watts D.J., Ashworth J.M.** (1970) Growth of myxameobae of the cellular slime mould *Dictyostelium discoideum* in axenic culture. *Biochem J* **119**: 171-4.
- Wemmer D.E.** (2001) Ligands recognizing the minor groove of DNA: development and applications. *Biopolymers* **52**: 197-211.
- Wilkins M.R., Gasteiger E., Bairoch A., Sanchez J.C., Williams K.L., Appel R.D., Hochstrasser D.F.** (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol.* **112**: 531-52.
- Winckler T., Trautwein C., Tschepke C., Neuhäuser C., Zündorf I., Beck P., Dingermann T.** (2001). Gene function analysis by amber stop codon suppression: CMBF is a nuclear protein that supports growth and development of *Dictyostelium* amoebae. *J. Mol. Biol.* **305**: 703-14.
- Winckler T., Dingermann T., Glöckner G.** (2002). *Dictyostelium* mobile elements: strategies to amplify in a compact genome. *Cell. Mol. Life Sci.* **59**: 2097-2111.
- Winckler T., Iranfar N., Beck P., Jennes I., Siol O., Baik U., Loomis W.F., Dingermann T.** (2004) CbfA, the C-Module DNA-Binding Factor, Plays an Essential Role in the Initiation of *Dictyostelium discoideum* Development. *Eukaryotic Cell* **3(5)**: 1349-58.
- Winckler T., Szafranski K., Glöckner G.** (2005). Transfer RNA gene-targeted integration: an adaptation of retrotransposable elements to survive in the compact *Dictyostelium discoideum* genome. *Cytogenet Genome Res* **110**: 288-298.
- Xu Q, Shaulsky G.** (2005) GOAT: An R Tool for Analysing Gene Ontology Term Enrichment. *Appl Bioinformatics.* **4(4)**: 281-3.
- Yang Y. H., Dudoit S., Luu P., Lin D. M., Peng V., Ngai J., Speed T. P.** (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* **30(4)**: e15.
- Zingler N., Willhoeft U., Brose H.P., Schoder V., Jahns T., Hanschmann K.M., Morrish T.A., Löwer J., Schumann G.G.** (2005) Analysis of 5' junctions of human LINE-1 and Alu retrotransposons suggests an alternative model for 5'-end attachment requiring microhomology-mediated end-joining. *Genome Res.* **15(6)**: 780-9.

## Anhang

### A Liste der verwendeten Primer

Angegeben sind alle wesentlichen in dieser Arbeit benutzten Primer. Genbezeichnung und Primerbeszeichnung unterscheiden sich oft aufgrund geänderter Nomenklatur innerhalb der Referenzdatenbank **Dictybase**. Zur Klonierung verwendete Schnittstellen von Endonukleasen sind in kleinen Buchstaben dargestellt, zur Mutagenese vorgesehene Nuklotide sind unterstrichen. Die zur Sequenzierung der CbfA-ähnlichen Proteine verwendeten Primer sind in Bohnert (2009) aufgelistet.

Gen	Sequenz 5' → 3'	Produkt
<i>gpdA</i>	Q-GAPDH-01 GGTTGTCCCAATTGGTATTAATGG Q-GAPDH-02 CCGTGGGTTGAATCATATTTGAAC	156 bp / 247 bp
<i>acaA</i>	acaA-04 CATTCTAGAGGCGGTATTGGCAGTATCACC acaA-09 GGAGAAAATGTCTGATTTTCGCTTTGGATG	172 bp
<i>csaA</i>	csaA-05 GAAAGCTGGTATCTCAAATGTTGTAC csaA-06 GGAATCTGGAGCACAACTATATCAGTAG	182 bp
DDB_G0277473	DDB0169268-01 CACCTCGTGTGGTGATTCTGT DDB0169268-02 CGCATGTTTGGTCTTCACCACT	204 bp
DDB_G0277403 ( <i>clc</i> )	clc-01 CCACAACATTTGACAACACTCAAAC clc-02 GCTGGTGCAGTTTCTTTGG	130 bp
DDB_G0281609	DDB0204558-01 TCAACAGTTAGGTGGTGGCTGT DDB0204558-02 GGCGGGACTGTTACAAGCAAAT	205 bp
DDB_G0286953	DDB0187209-01 TGGTTGTAATTGTGCACCAGCC DDB0187209-02 TTGGGCAAAGGTTAGCACATGG	236 bp
DDB_G0282815 ( <i>orfSGP</i> )	Q-DDB0214941-01 CACACATGGAAACAAACTCTCACTG Q-DDB0214941-02 CACGATCTATTTGTTGGATATATCTC	157 bp
DDB_G0274181	geneDDB0167552-01 GGATGTCACAATCAGAAAATGTTGC geneDDB0167552-02 CTACCACCAGTGTATCCACC	231 bp
DDB_G0292986 ( <i>abcG10</i> )	abcG10-01 CACTTGGTCAAGTTGTATCTGC abcG10-02 CGGCATGGATTTGAATGATAAACG	160 bp
DDB_G0275689 ( <i>abcG2</i> )	abcG2-01 CCATTCTTTTACCACTCCTTCATCTC abcG2-02 TAGCGATTGAAGCTGGTGG	175 bp
DDB_G0279751	DDB0206011-01 ACATTTAGCATCCTCACCACC DDB0206011-02 GTGAGCAATACCATAGAGATGATG	175 bp
DDB_G0267528	DDB0189346-01 TGATGGTACTCGTGTCCAG DDB0189346-02 GAGAGTTGAAGAGATACGCC	193 bp
<i>cbfA</i> (attB-CbfA <sup>724-998</sup> )	attB-CMBF01 GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATG GGTTTAATGTCTCATTTTTCTGG attB-CMBF02 GGGGACCACTTTGTACAAGAAAGCTGGGTTATTTTTTA TTATTATAAAGTTCTTGTTTTACATATAAAC	893 bp

## Anhang A

Gen	Sequenz 5' → 3'	Produkt
<i>cbfA</i> ( <i>CbfA</i> <sup>724-998</sup> )	CMBF-37.3 GGggtaccGGTTAATGTCTCATTTTTCTGG ( <i>KpnI</i> ) CMBF-40.1 CCctcgagTTATTTTTATTATTATAAAAGTTCTGTTTTAC ( <i>XhoI</i> )	844 bp
<i>cbfA</i> ( <i>CbfA</i> <sup>793-1000</sup> )	CMBF-124 GgcggccgcAATGGATGTGCATTCCAAGAAAAGACC ( <i>NotI</i> ) CMBF-125 GgcggccgcTAAATTATTTTTATTATTATAAAAGTTCTTG ( <i>NotI</i> )	645 bp
<i>cbfA</i> -Mutagenese (Xa-Schnittstelle)	CMBF-139 CCGAATTCGAGCTCCAATCTCGGCATTGAGGGCCGAATGGAT GTGCATTCC	
<i>cbfA</i> -Mutagenese (K799N R800G)	CMBF-153 ATCATTATCTAAAACAATTGGTCCATTCTTGAATGCACATCCA TTGG	
<i>cbfA</i> -Mutagenese (R812A)	CMBF-155 CTTTTAAATTCTTTGGTGGTGCACCTTTGGTTTATCATTATCTA AAAC CMBF-156 GATGTGCATTCCAAGAATG (Screening-Primer)	
<i>cbfA</i> -Mutagenese (P801K)	CMBF-161 GGTTTATCATTATCTAAAACAATTTTCTTTCTTGGAAATGCAC ATCCATTGG	
<i>cbfA</i> -Mutagenese (K808A)	CMBF-169 GTGGTGCACCTTTGGTGCATCATTATCTAAAACAATTTTC CMBF-170 TGTTTTAGATAATGATGC (Screening-Primer)	
<i>cbfA</i> -Mutagenese (K810A)	CMBF-171 CTTTGGTGGTGCACCTGCTGGTTTATCATTATCTAAAAC CMBF-172 AGATAATGATAAACCAGC (Screening-Primer)	
<i>cbfA</i> -Mutagenese (K815A)	CMBF-173 GTCCATTCTTTAAATTCGCTGGTGGTGCACCTTTGG CMBF-174 CCAAAGGTGCACCACCAGC (Screening-Primer)	
<i>cbfA</i> -Mutagenese (K818A)	CMBF175 GAGTTGATGTCCATTCTGCTAAATCTTTGGTGGTGCA CMBF-176 GCACCACAAAGAATTTAGC (Screening-Primer)	
Y2h (Beute- amplifizierung)	Y2h-5' TTCACCCAAGCAGTGGTATCAACGCAGAGTGG Y2h-3' GTATCGATGCCACCCTCTAGAGGCCGAGGCGGCCGACA	95 bp + cDNA
Y2h (Bank- konstruktion)	CDSIII ATTCTAGAGCCGAGGCCGCCGACATG-d(T) <sub>30</sub> VN CDSIII/6 ATTCTAGAGCCGAGGCCGCCGACATG-NNNNNN SmartIII AAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCCGGG	
Y2h (Ausschluss Selbstaktivierer)	pGAD426seq GTATAACGCGTTTGAATCACTACAGG Y2h-1#1 CCCTAAATTTATTTGAATAATTCTTCATCTC	430 bp
Biotinylierte PCR-Produkte ( <i>acaAP</i> und C-Modul)	pGEM-for(biotin) Biotin-GGCCCGACGTCGCATGCTCC pGEM-rev GGTCGACCTGCAGGCGGCC KO-bsr-rev CCAACGCGTTGGGAGCTCTCCGGATC <i>acaAP</i> -04 GGTACAACAAATTAATTTACTAGTAA	222 bp 337 bp 342 bp 386 bp
<i>purF</i>	<i>purF</i> -01 CCTCTCCACCAGTTTTACATCCATGTC <i>purF</i> -02 GTTGGGGATGAAGTTGAAAACATGG	220 bp
<i>purD</i>	<i>purD</i> -01 GGTGAAACAGCAGAAATGCCAGGAATG <i>purD</i> -02 GTTTTACTTGAATCGAATGGTGAATTG	243 bp
<i>purN</i>	<i>purN</i> -01 CATCAATTGAAACACGTGTATTCTCAC <i>purN</i> -02 CATTGAAAGCACGTTTCGATTGCATGTGC	273 bp
<i>purL</i>	<i>purL</i> -01 GTCGTTTGAATCTCGTTGGGTCAATG <i>purL</i> -02 GAGTACCAGATGGATTGAATGGATAGG	234 bp
<i>purC/E</i>	<i>purC/E</i> -01 CAAACCATTGCACCAGCTCAATTGCCAG <i>purC/E</i> -02 GAATCAATTGTTCAAATTGAGAGGTTAC	238 bp
<i>purH</i>	<i>purH</i> -01 GAAGGTCGTCCTGAAAATAAACCAAC <i>purH</i> -02 GAGCTAAATCGATTGAATCGGCCAATGG	254 bp
<i>purA</i>	<i>purA</i> -01 GTTAGATGCTGCTGTTTTGCGTTACAC <i>purA</i> -02 TTGGGAGTTGATCATATTCGGTAACATG	248 bp

## Anhang A

Gen	Sequenz 5' → 3'	Produkt
<i>purB</i>	purB-01 CCGCTAAATTACCAATCTCCCGTCTTC purB-02 GGTTTTGGGAAACCTTCACGACGAAGG	229 bp
<i>guaA</i>	guaA-01 GGTTTTCATTTTCTCTCAAAATGCTAC guaA-02 CAGGAACACCAGTCATGAAATCATTTG	278 bp
<i>pyr1-3</i>	pyr13-01 CATTCTTTGGTTCGTCTTTAGCAAAC pyr13-02 GAAGCTTTAGTGAGAGTATGTGGAGTG	266 bp
<i>pyr4</i>	pyr4-01 CAACAAGCTGAAAGAGCTGGCGTTGCC pyr4-02 CAGAATTTAAAAAGTCATCAAAATGCTC	262 bp
<i>pyr5-6</i>	pyr56-01 GTTTCAAATCAGGTTTGAAAGAGTATG pyr56-02 GTATTATATTGTTGACCCATAGCATCAC	247 bp
<i>thyA</i>	thyA-01 GAAGAAGCATTGGAGCCAAAAACCGCC thyA-02 CAACTGGTACAATTGGACGAATGAGAG	277 bp

## B Plasmidliste

Nachfolgende Liste enthält alle im Rahmen dieser Arbeit erstellte Vektoren. Die Plasmidnummer bezieht sich auf die lehrstuhlinterne Datenbank. Vektor und Insertgrößen sind in Basenpaaren angegeben.

Plasmid Name	Nummer	Inserts	Vektor	Insert
pPT132 Cmbf1-2	1.1639	PCR attrBcmbf1 und 2 (893 bp) an genomischer DNA, Zwischenklonierung in pEntry1-2 (1.1631), Rekombination in pPT132 (1.1622) → GFP-CbfA <sup>724-998</sup> (PB1)	6000	893
pGEM-Linker	2.0001	Annealing CMBF-133 mit 134 zur Vektormodifikation	3000	58
pGEM-CTD	2.0002	PCR CMBF-124 und 125 (645 bp) an pGEM-SS01/02 (2.0008) → CbfA <sup>793-1000</sup>	3000	642
pGEM-CTD (N816S)	2.0003	PCR CMBF-124 und 125 an pGEM-SS01/02 (2.0008), Mutation in Sequenz (N816S)	3000	642
pET33b(+)/CTD	2.0020	NotI-Fragment (635bp) aus pGEM-CTD (2.0002) → (His) <sub>6</sub> -CbfA <sup>793-1000</sup> -Expressionskonstrukt	5378	642
pGADT7-Bgl-Linker +3	2.0025	Annealing aus Bgl-Linker+3_1 und Bgl-Linker+3_2	7988	33
pGADT7-Bgl-Linker +3 (aus +3/1)	2.0026	Annealing aus Bgl-Linker+3_1 und Bgl-Linker+3_2	7988	33
pGEM-GST (-Bgl2)	2.0028	3' BglII-Site von pGEM-GST (1.1860) zerstört	3000	731
pGBKT7/CTD	2.0030	NotI-Fragment (635bp) aus pGEM-CTD (2.0002) → Köderplasmid für Hefe-2-Hybrid	7304	642
pGBKT7/CTD	2.0031	NotI-Fragment (635bp) aus pGEM-CTD (2.0002)	7304	642
pGADT7-Bgl-Linker +2	2.0051	Annealing aus Bgl-Linker+2_1 und Bgl-Linker+2_2	7988	32
pMAL-c2x (NotI) / CTD	2.0052	NotI-Fragment (635bp) aus pGEM-CTD (2.0002) → MBP- CbfA <sup>793-1000</sup> -Expressionskonstrukt	6639	642
pGEM-SS01_02 ura bgl2 / GST	2.0054	GST mit BglII aus pGEM-GST (1.1860) in pGEM-SS01/02 ura bgl (2.0008) kloniert	9392	723
pGADT7-REC-AX2-Bank	2.0067	cDNA-Beute-Bank aus vegetativen <i>D.discoideum</i> AX2 mRNA → oligo (dT)-Bank für Hefe-2-Hybrid	8059	div
pGBKT7/ORF1	2.0120	NotI-Fragment aus pGEM-ORF1 NotI (1.1757)	7304	1351
pGEM7zf(-)/CTD-x	2.0121	SphI/XhoI-Fragment (1081bp) aus pET33b(+)/CTD (2.0020) in pGEM7zf(-) (1.0943)	2987	1081
pGADT7-REC-AX2-Bank2	2.0125	cDNA-Beute-Bank aus vegetativen <i>D.discoideum</i> AX2 mRNA → (N) <sub>6</sub> -Bank für Hefe-2-Hybrid	8059	div
pGEM7zf(-)/CTD-Xa	2.0145	Mutagenese zur Einführung Faktor Xa-Schnittstelle	2987	1081
pGEM-acaAP (EMSA)	2.0161	PCR mit acaAP_EMSA01/02 (172 bp) an pGEM-acaAP (1.1797)	3000	172
pET33b(+)/CTD-Xa	2.0162	CTD aus CbfA mit N-terminaler Xa-Schnittstelle → finales (His) <sub>6</sub> - CbfA <sup>793-1000</sup> -Xa-Expressionskonstrukt	4937	1081
pGEM7zf(-)-pET-PB1	2.0244	XbaI/NsiI-Fragment aus pPT132-Cmbf1-2 (1.1639) in 2.0121	3170	2300
pGEM7zf(-)-PB1	2.0245	XbaI/NsiI-Fragment aus pPT132-Cmbf1-2 (1.1639) in 1.0259	2920	2300
pGEM7zf(-)-PB1-K799N R800G	2.0249	Mutagenese für PB1 K799N R800G (vordere NLS zerstört)	2920	2300
pGEM7zf(-)-ctd-PB1-R812A (20)	2.0250	Mutagenese für PB1 R812A (AT-Haken; hintere NLS zerstört)	3170	2300

## Anhang B

Plasmid Name	Nummer	Inserts	Vektor	Insert
pGEM7zf(-)-PB1-R812A (29)	2.0251	Mutagenese für PB1 R812A (AT-Haken; hintere NLS zerstört)	2920	2300
pGEM7zf(-)-ctd-PB1-P801K	2.0252	Mutagenese für PB1 P801K (vordere NLS verbessert)	3170	2300
pGEM7zf(-)-ctd-PB1-P801K R812A	2.0253	Mutagenese für PB1 P801K (vordere NLS verbessert) R812A (AT-Haken, hintere NLS zerstört)	3170	2300
pPT132-CMBF1-2 K799N R800G	2.0257	GFP-PB1 K799N R800G (vordere NLS zerstört)	5000	2000
pPT132-CMBF1-2 R812A	2.0258	GFP-PB1 R812A (hintere NLS / AT-Haken zerstört)	5000	2000
pPT132-CMBF1-2 P801K	2.0259	GFP-PB1 P801K (vordere NLS optimiert)	5000	2000
pPT132-CMBF1-2 P801K R812A	2.0260	GFP-PB1 P801K (vordere NLS optimiert) R812A (hintere NLS / AT-Haken zerstört)	5000	2000
pGEM7zf(-)-ctd-PB1-K799N R800G R812A	2.0291	Mutagenese für PB1 K799N R800G (vordere NLS zerstört) R812A (AT-Haken; hintere NLS zerstört)	3170	2300
pGEM7zf(-)-ctd-PB1-P801K R812A K808A	2.0292	Mutagenese für PB1 P801K (vordere NLS verbessert) R812A (AT-Haken, hintere NLS zerstört) K808A (Lysin um AT-Haken ausgetauscht)	3170	2300
pGEM7zf(-)-ctd-PB1-P801K R812A K810A	2.0293	Mutagenese für PB1 P801K (vordere NLS verbessert) R812A (AT-Haken, hintere NLS zerstört) K810A (Lysin um AT-Haken ausgetauscht)	3170	2300
pGEM7zf(-)-ctd-PB1-P801K R812A K815A	2.0294	Mutagenese für PB1 P801K (vordere NLS verbessert) R812A (AT-Haken, hintere NLS zerstört) K815A (Lysin um AT-Haken ausgetauscht)	3170	2300
pGEM7zf(-)-ctd-PB1-P801K R812A K818A	2.0295	Mutagenese für PB1 P801K (vordere NLS verbessert) R812A (AT-Haken, hintere NLS zerstört) K818A (Lysin um AT-Haken ausgetauscht)	3170	2300
pPT132-CMBF1-2 K799N R800G R812A	2.0296	GFP-PB1 K799N R800G (vordere NLS zerstört) R812A (hintere NLS / AT-Haken zerstört)	5000	2000
pPT132-CMBF1-2 P801K R812A K808A	2.0297	GFP-PB1 P801K (vordere NLS optimiert) R812A (hintere NLS / AT-Haken zerstört) K808A (Lysin um AT-Haken ausgetauscht)	5000	2000
pPT132-CMBF1-2 P801K R812A K810A	2.0298	GFP-PB1 P801K (vordere NLS optimiert) R812A (hintere NLS / AT-Haken zerstört) K810A (Lysin um AT-Haken ausgetauscht)	5000	2000
pPT132-CMBF1-2 P801K R812A K815A	2.0299	GFP-PB1 P801K (vordere NLS optimiert) R812A (hintere NLS / AT-Haken zerstört) K815A (Lysin um AT-Haken ausgetauscht)	5000	2000
pPT132-CMBF1-2 P801K R812A K818A	2.0300	GFP-PB1 P801K (vordere NLS optimiert) R812A (hintere NLS / AT-Haken zerstört) K818A (Lysin um AT-Haken ausgetauscht)	5000	2000
pDXA-Tcep-PB1	2.0304	<i>KpnI/XhoI</i> -Fragment aus pDXA-PB1 (1.1604, CbfA <sup>724-998</sup> ) in modifizierten pDXA (2.0241)	6171	831
pDXA-Tcep-Dfas CTD	2.0305	<i>KpnI/XhoI</i> -Fragment aus pET-D.fas. CTD (2.0222) → <i>D. fasciculatum</i> CbfA-CTD	6171	1449
pGEM-GFP22_23	2.0306	PCR mit GFP22 und 23 an A15-GFP (1.1566)	3000	770
pGEM-gDNA	2.0316	Genomische DNA aus AX2 und JH.D <i>Sau3AI</i> geschnitten und in <i>Bam</i> HI-site von 2.0001 ligiert	3058	div

### C Differentiell regulierte Gene zwischen AX2 und JH.D sowie ihre Antwort auf GFP-CbfA<sup>724-998</sup> in JH.D

Gen ID ermöglicht eine Zuordnung der für ein Gen genutzten Sonde auf dem Microarray. DDB ID entspricht der zum Zeitpunkt des Experimentes zugeordneten Identifikationsnummer, DDB\_G ID der aktuellsten. NA = nicht ausgewiesen; NS = nicht signifikant

Cluster	Gen ID	DDB ID	DDB_G ID	Genname	AX2 / JH.D		JH.D CbfA <sup>(724-998)</sup> / JH.D	
					Score(d)	Regulation(2 <sup>f</sup> )	Score(d)	Regulation(2 <sup>f</sup> )
	VSE513	DDB0188003	DDB_G0288579		5,518	2,783	NS	NS
	SLG565	DDB0185090	DDB_G0273175	cf50	4,911	2,460	NS	NS
	SLK481	DDB0169192	DDB_G0277383		4,014	2,346	NS	NS
	SSC695	DDB0169400	DDB_G0277761		3,383	2,278	NS	NS
	SSF363	DDB0191161	DDB_G0269248		4,410	2,258	NS	NS
	SSH822	DDB0169268	DDB_G0277473		3,546	2,120	NS	NS
	VSG620	NA	NA		3,483	2,112	NS	NS
	VSH675	NA	NA		3,707	2,092	NS	NS
	VSK133	NA	NA		3,891	2,081	NS	NS
	VSC157	DDB0205158	DDB_G0282153		3,298	2,062	NS	NS
	SLB276	DDB0215005	DDB_G0279187	cprG	2,918	1,967	NS	NS
	SLF133	DDB0219654	DDB_G0291191		2,591	1,907	NS	NS
	SSF663	DDB0237483	DDB_G0295657	ndufs4	2,890	1,797	NS	NS
	SLC357	DDB0232276	DDB_G0277615		2,910	1,783	NS	NS
	SLB420	DDB0185017	DDB_G0277503	amtA	3,222	1,744	NS	NS
	SSJ380	DDB0186397	DDB_G0285215		2,942	1,743	NS	NS
	VSK159	DDB0231380	DDB_G0277455	plbE	2,608	1,702	NS	NS
	SLG322	DDB0229903	DDB_G0280105		2,920	1,692	NS	NS
	SLD112	DDB0185207	DDB_G0277401	vatB	2,424	1,679	NS	NS
	SSK264	DDB0191305	DDB_G0277403	clc	2,841	1,671	NS	NS
	VSC724	DDB0169107	DDB_G0277195		2,695	1,666	NS	NS
	SSE713	DDB0231542	DDB_G0277173	vamp7B	2,223	1,623	NS	NS
	U23957	DDB0191165	DDB_G0267380	argE	2,781	1,588	NS	NS
	SLE371	DDB0233715	DDB_G0286651		2,004	1,585	NS	NS
	VSA111	DDB0192011	DDB_G0293566		2,516	1,582	NS	NS
	SSA309	DDB0191178	DDB_G0287125	proA	1,941	1,582	NS	NS
	SLE433	DDB0191163	DDB_G0267400	hspD	1,795	1,578	NS	NS
	SLC522	DDB0215398	DDB_G0278581	dicA1	2,298	1,568	NS	NS
	SSL215	DDB0233838	DDB_G0280187		2,111	1,568	NS	NS
	VSJ849	DDB0237753	DDB_G0277213		2,033	1,562	NS	NS
	VSE332	DDB0203967	DDB_G0281037		2,257	1,533	NS	NS
	SSA582	DDB0233836	DDB_G0283401	ctsZ	2,222	1,531	NS	NS
	SLD246	DDB0216526	DDB_G0267368		1,851	1,525	NS	NS
	VSC658	DDB0185526	DDB_G0283449		2,373	1,523	NS	NS
	VSC589	DDB0232014	DDB_G0277227		2,060	1,514	NS	NS
	SLK595	DDB0237754	DDB_G0276067	ddcB	2,030	1,507	NS	NS
	SSA265	DDB0191267	DDB_G0277139	apm2	2,299	1,507	NS	NS

Anhang C

Cluster	Gen ID	DDB ID	DDB_G ID	Genname	AX2 / JH.D		JH.D CbfA <sup>(724-998)</sup> / JH.D	
					Score(d)	Regulation(2')	Score(d)	Regulation(2')
II	SSD673	DDB0237660	DDB_G0281961		2,993	1,808	-4,213	0,670
	SSM263	DDB0204558	DDB_G0281609		3,147	1,727	-6,689	0,645
	SSB389	DDB0238293	DDB_G0283523		2,395	1,631	-5,137	0,629
	SSK687	DDB0215459	DDB_G0291646	V4-7	2,852	1,589	-4,189	0,668
	SLJ343	DDB0218067	DDB_G0278623		2,233	1,548	-4,321	0,703
	SSE294	DDB0219879	DDB_G0293714		2,673	1,546	-5,062	0,736
	SLD796	DDB0187993	DDB_G0288563		2,076	1,535	-4,022	0,735
	SLE503	DDB0233995	DDB_G0277493	cct6	2,225	1,529	-3,946	0,864
III	VSB592	DDB0214941	DDB_G0282815	orfSGP	7,096	6,091	20,791	7,975
	SLB184	DDB0214912	DDB_G0281823	V4-7	7,352	4,359	5,744	1,570
	SSD815	DDB0167552	DDB_G0274181		6,246	3,529	14,053	2,577
	VSB633	DDB0217719	DDB_G0275911		4,367	3,055	5,521	1,969
	SSF615	DDB0231278	DDB_G0275119	alyB	5,014	2,950	15,567	2,144
	SSM789	DDB0215394	DDB_G0291255		5,072	2,863	8,620	2,942
	VSC651	DDB0187209	DDB_G0286953		5,680	2,818	5,332	2,026
	VSB701	DDB0235310	DDB_G0284549		4,950	2,747	6,569	1,420
	SSL266	DDB0185971	DDB_G0284365		3,776	2,608	9,569	2,114
	SSH545	DDB0185092	DDB_G0272815	cprE	4,010	2,362	7,483	1,593
	SLJ671	DDB0215583	DDB_G0293690		4,047	2,273	13,779	4,705
	VSD610	DDB0191902	DDB_G0293366		3,784	2,155	7,709	2,344
	VSB639	DDB0232085	DDB_G0272238		4,052	2,137	4,148	1,562
	SSL389	DDB0233609	DDB_G0274459		3,535	2,129	4,202	1,549
	SSF731	DDB0237972	DDB_G0269616		3,861	2,062	8,520	1,826
	SSF705	DDB0185020	DDB_G0275693		3,390	2,004	3,304	1,331
	SLE471	NA	NA		2,675	1,775	3,677	1,250
	VSG827	DDB0218414	DDB_G0282905		2,917	1,758	7,964	1,996
	SSK228	DDB0220657	DDB_G0291834	cpiA	2,498	1,750	2,875	1,368
	SSH534	DDB0233476	DDB_G0270978		3,223	1,725	4,689	1,359
	VSB565	DDB0237633	DDB_G0284955		3,039	1,708	6,860	1,556
	SLJ376	DDB0191099	DDB_G0292378	fhbA	3,101	1,707	7,897	1,934
	SSE485	DDB0219864	DDB_G0293492		2,534	1,679	4,365	1,267
	SSM682	DDB0214999	DDB_G0278721	cprD	2,549	1,665	4,505	1,378
	SLB609	DDB0229833	DDB_G0284363	cf60	2,562	1,648	4,787	1,458
	X61581	NA	NA		2,817	1,640	4,955	1,233
	SLB855	DDB0215400	DDB_G0273063	dscA	2,193	1,630	11,746	2,756
	VSJ810	DDB0235252	DDB_G0291301		2,297	1,600	6,206	2,250
	SSL817	NA	NA		2,535	1,600	4,043	1,475
	AB029922	DDB0216234	DDB_G0289391	pkiA	2,604	1,599	10,676	2,126
	SLB804	DDB0217967	DDB_G0277377		2,569	1,575	4,665	1,211
	VSC144	DDB0191097	DDB_G0275007		2,421	1,562	3,705	1,232
VSA633	DDB0231511	DDB_G0288935		2,409	1,550	4,212	1,351	
SSJ467	DDB0233455	DDB_G0277209		1,988	1,550	3,132	1,378	
SSC321	DDB0233867	DDB_G0289439		2,192	1,542	8,259	1,588	



Anhang C

Cluster	Gen ID	DDB ID	DDB_G ID	Genname	AX2 / JH.D		JH.D CbfA <sup>(724-998)</sup> / JH.D	
					Score(d)	Regulation(2 <sup>r</sup> )	Score(d)	Regulation(2 <sup>r</sup> )
IV	VSA360	DDB0237696	DDB_G0274207		-3,653	0,501	8,131	2,604
	X15387	DDB0220110	DDB_G0291121	cinB	-3,666	0,524	7,083	2,489
	SSC656	DDB0237922	DDB_G0287609		-2,768	0,546	9,601	2,779
	SLD691	DDB0214895	DDB_G0292986	abcG10	-2,637	0,572	6,596	1,860
V	SSM424	DDB0218347	DDB_G0282075		-7,913	0,111	-9,842	0,194
	SSL845	NA	NA		-7,500	0,128	-9,449	0,208
	SSE346	DDB0167826	DDB_G0274289		-7,731	0,133	-10,328	0,199
	SSA288	NA	NA		-5,916	0,157	-8,260	0,211
	SSJ758	DDB0167628	DDB_G0274423	cprE	-7,386	0,166	-12,017	0,228
	SSH187	NA	NA		-5,945	0,169	-8,897	0,245
	SSL878	NA	NA		-7,368	0,172	-11,819	0,272
	VSI401	DDB0191230	DDB_G0287461	abcG3	-7,730	0,189	-10,059	0,267
	SLI159	DDB0206011	DDB_G0279751		-5,431	0,267	-15,331	0,395
	SLE486	DDB0167523	DDB_G0274153		-6,510	0,322	-16,290	0,379
	SLA566	DDB0215341	DDB_G0275323	tipD	-3,171	0,349	-10,273	0,368
	SSG739	NA	NA		-4,571	0,382	-7,640	0,514
	VSG644	DDB0203824	DDB_G0276383	cf50	-4,091	0,413	-3,777	0,671
	VSC645	NA	NA		-3,222	0,449	-5,178	0,622
	SSC146	DDB0219898	DDB_G0293788		-3,719	0,454	-3,555	0,633
	SLD676	DDB0238609	DDB_G0287491		-3,970	0,457	-5,950	0,664
	SLJ337	DDB0238299	DDB_G0285617		-2,834	0,502	-4,960	0,616
	SLD725	NA	NA		-3,170	0,507	-7,586	0,514
	SSD285	DDB0191228	DDB_G0284047	abcF2	-3,012	0,509	-12,580	0,422
	SSC606	DDB0237856	DDB_G0277429		-3,201	0,512	-5,622	0,649
	VSG586	DDB0219241	DDB_G0287575		-3,754	0,518	-7,447	0,584
	SSK662	DDB0190552	DDB_G0269786		-3,737	0,519	-6,074	0,645
	SSJ785	DDB0230996	DDB_G0287515	tat	-3,486	0,526	-4,569	0,750
	SLA466	DDB0231577	DDB_G0268028		-2,968	0,537	-7,513	0,573
	SSM755	DDB0191229	DDB_G0275689	abcG2	-3,073	0,541	-4,691	0,582
	SSA222	DDB0238634	DDB_G0269092		-2,778	0,549	-4,603	0,671
	SSK318	DDB0237769	DDB_G0287229	chdh	-2,742	0,552	-11,799	0,410
	VSF328	NA	NA		-2,595	0,555	-5,065	0,488
	SSK605	DDB0230088	DDB_G0283987	purC/E	-2,405	0,556	-5,592	0,740
	SSD563	DDB0231652	DDB_G0286925		-3,136	0,560	-4,300	0,765
	SSF224	DDB0238290	DDB_G0272718		-2,834	0,568	-8,170	0,673
	VSH428	DDB0229909	DDB_G0275501		-2,415	0,570	-6,690	0,487
	SSL494	DDB0185839	DDB_G0284091		-3,388	0,572	-11,021	0,694
	SSE412	DDB0185193	DDB_G0291253	dia2	-2,498	0,575	-4,977	0,703
	VSB586	NA	NA		-2,706	0,578	-3,946	0,751
	SSA273	NA	NA		-2,492	0,585	-4,702	0,755
	SSC633	DDB0231608	DDB_G0278155	mai	-2,761	0,589	-11,958	0,602
	SSK833	DDB0233832	DDB_G0288091		-2,886	0,590	-9,455	0,649
	SSJ747	DDB0233340	DDB_G0284167		-1,969	0,598	-4,724	0,585
	SSL464	DDB0189610	DDB_G0267874		-2,433	0,601	-6,397	0,574
	SSG526	DDB0191227	DDB_G0290483	abcE1	-2,523	0,603	-7,655	0,653
	SSB349	DDB0230128	DDB_G0283419	gpal	-2,919	0,608	-9,240	0,631
SSI408	DDB0202191	DDB_G0268398		-2,193	0,609	-6,493	0,649	
SSI496	DDB0214930	DDB_G0277901	apm3	-2,232	0,620	-4,888	0,671	
SLB620	DDB0201597	DDB_G0294016	p17	-2,219	0,623	-6,504	0,563	
SSA106	DDB0230072	DDB_G0277947		-2,053	0,628	-8,182	0,609	
VSG761	DDB0191167	DDB_G0268630	cprE	-2,718	0,631	-13,412	0,599	

Anhang C

Cluster	Gen ID	DDB ID	DDB_G ID	Genname	AX2 / JH.D		JH.D CbfA <sup>(724-998)</sup> / JH.D	
					Score(d)	Regulation(2')	Score(d)	Regulation(2')
V	SLA540	DDB0230094	DDB_G0288333		-2,072	0,633	-4,656	0,752
	SSA754	DDB0215334	DDB_G0281551		-2,224	0,633	-6,326	0,759
	VSE458	DDB0167477	DDB_G0275111		-2,839	0,634	-11,768	0,385
	SLA332	DDB0216411	DDB_G0285425		-2,213	0,635	-7,763	0,573
	SLE885	DDB0237662	DDB_G0293202		-2,287	0,638	-7,046	0,454
	VSD427	DDB0187160	DDB_G0286853		-2,092	0,641	-3,284	0,850
	SSG701	DDB0230084	DDB_G0290121	purD	-2,492	0,642	-5,200	0,714
	VSC555	NA	NA		-2,142	0,645	-5,188	0,459
	SLG272	DDB0231108	DDB_G0271678	pckA	-1,934	0,645	-3,696	0,728
	VSJ853	DDB0201652	DDB_G0291856	gerD	-2,148	0,647	-7,574	0,650
	VSD139	DDB0252602	DDB_G0295779		-1,993	0,652	-4,804	0,621
	VSJ569	DDB0214937	DDB_G0279921	gp130	-2,099	0,655	-7,796	0,443
	VSH763	DDB0205823	DDB_G0279527		-1,987	0,659	-3,665	0,710
VI	X15381	DDB0216253	DDB_G0280543	lmcA	-8,498	0,130	NS	NS
	X15380	DDB0215343	DDB_G0280533	lmcB	-9,445	0,161	NS	NS
	VSA166	NA	NA		-5,034	0,210	NS	NS
	VSH732	DDB0189346	DDB_G0267528		-3,699	0,373	NS	NS
	SLA602	DDB0191292	DDB_G0267386		-2,984	0,590	NS	NS
	SSM236	DDB0216231	DDB_G0270836	fbp	-2,960	0,597	NS	NS
	VSE478	DDB0192224	DDB_G0293956		-2,784	0,599	NS	NS
	SSI137	DDB0189956	DDB_G0268696		-2,583	0,604	NS	NS
	M26017	DDB0191363	DDB_G0288373		-2,067	0,620	NS	NS
	VSH415	DDB0185217	DDB_G0276331		-2,269	0,631	NS	NS
	VSG570	DDB0204128	DDB_G0281281	cf50	-2,315	0,631	NS	NS
	SSI248	DDB0231632	DDB_G0276287		-2,461	0,632	NS	NS
	VSK149	DDB0233521	DDB_G0287363	sibA	-2,268	0,637	NS	NS
	VSA167	NA	NA		-2,249	0,641	NS	NS
	SLK653	DDB0234094	DDB_G0269380		-2,219	0,644	NS	NS
	SSK888	DDB0230208	DDB_G0274559	udkB	-2,320	0,646	NS	NS
	SLD548	DDB0191239	DDB_G0269206	abcG21	-1,887	0,654	NS	NS
VSA566	DDB0230098	DDB_G0283701	guaB	-2,015	0,656	NS	NS	

## D Liste der Anreicherungen biologischer Prozesse, molekularer Funktionen und zelluläre Kompartimente in den differentiell exprimierten Genen zwischen AX2 und JH.D

Ergebnis der Genontologie-Analyse mithilfe des Programmpaketes **Gene Ontology Analysis Tool** (GOAT, Xu und Shaulsky 2005) für die durch CbfA-Suppression differentiell exprimierten Gene. GO Level gibt hierarchische Ordnung der Annotation an, wobei 1 der allgemeinsten entspricht, GO ID und Annotation definieren den Prozess, die Aktivität und die Kompartimentierung. Anreicherung (Anr.) ergibt sich aus dem Verhältnis der Gene in der Liste und aller Gene (Total) mit dieser Annotation in Bezug auf die Gesamtzahl der Gene in der Liste und der annotierten Gene, p-Wert entspricht der statistischen Signifikanz entsprechend der Hypergeometrischen Verteilung modifiziert nach Benjamini und Hochberg (1995).

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>AX2 &gt; JH.D biologischer Prozess</b>	6	GO:0042594	response to starvation	6	158	4	0,005
	5	GO:0031667	response to nutrient levels	6	159	4	0,005
	4	GO:0009991	response to extracellular stimulus	6	161	4	0,006
	3	GO:0009605	response to external stimulus	6	230	3	0,018
	3	GO:0019954	asexual reproduction	6	151	4	0,005
	5	GO:0000270	peptidoglycan metabolic process	4	12	36	0
	5	GO:0016052	carbohydrate catabolic process	4	48	9	0,002
	6	GO:0009253	peptidoglycan catabolic process	4	10	44	0
	3	GO:0051674	localization of cell	4	85	5	0,011
	4	GO:0055084	fruiting body development in response to starvation	6	150	4	0,004
	3	GO:0030582	fruiting body development	6	150	4	0,004
	4	GO:0030587	sorocarp development	6	150	4	0,004
	6	GO:0031158	negative regulation of aggregate size involved in sorocarp development	2	7	31	0,004
	6	GO:0006508	proteolysis	11	212	6	0
	7	GO:0006907	pinocytosis	2	20	11	0,016
	7	GO:0016998	cell wall catabolic process	4	14	31	0
	6	GO:0010382	cell wall metabolic process	4	14	31	0
	5	GO:0007047	cell wall organization and biogenesis	4	23	19	0
	4	GO:0045229	external encapsulating structure organization and biogenesis	4	23	19	0
	3	GO:0006928	cell motility	4	85	5	0,011
	7	GO:0042593	glucose homeostasis	2	5	44	0,002
	6	GO:0033500	carbohydrate homeostasis	2	5	44	0,002
	5	GO:0048878	chemical homeostasis	3	31	11	0,005
	4	GO:0042592	homeostatic process	3	60	5	0,018
	7	GO:0031157	regulation of aggregate size involved in sorocarp development	2	13	17	0,009
	6	GO:0031156	regulation of sorocarp development	2	18	12	0,013
	5	GO:0031155	regulation of fruiting body development	2	18	12	0,013
	6	GO:0051248	negative regulation of protein metabolic process	2	9	24	0,005
	4	GO:0016337	cell-cell adhesion	2	18	12	0,013

## Anhang D

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>AX2 &gt; JH.D molekulare Funktion</b>	7	GO:0015662	ATPase activity\, coupled to transmembrane movement of ions\, phosphorylative mechanism	2	29	8	0,033
	6	GO:0016298	lipase activity	2	34	7	0,039
	6	GO:0003796	lysozyme activity	4	10	49	0
	5	GO:0004553	hydrolase activity\, hydrolyzing O-glycosyl compounds	4	78	6	0,007
	4	GO:0016798	hydrolase activity\, acting on glycosyl bonds	4	87	6	0,01
	6	GO:0004197	cysteine-type endopeptidase activity	7	29	29	0
	5	GO:0008234	cysteine-type peptidase activity	7	58	15	0
	5	GO:0004175	endopeptidase activity	7	106	8	0
	4	GO:0008233	peptidase activity	9	197	6	0
3	GO:0016787	hydrolase activity	18	1022	2	0,002	

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>AX2 &gt; JH.D zelluläres Kompartiment</b>	4	GO:0031988	membrane-bound vesicle	9	258	4	0,001
	3	GO:0031982	vesicle	9	260	4	0,001
	7	GO:0045335	phagocytic vesicle	7	193	4	0,002
	7	GO:0030136	clathrin-coated vesicle	2	19	12	0,02
	4	GO:0005615	extracellular space	2	16	14	0,009
	3	GO:0044421	extracellular region part	2	20	11	0,02
	2	GO:0005576	extracellular region	6	88	8	0
	7	GO:0015630	microtubule cytoskeleton	6	139	5	0,002
	7	GO:0005764	lysosome	2	30	7	0,032
	7	GO:0030139	endocytic vesicle	8	200	4	0
	7	GO:0030135	coated vesicle	2	36	6	0,044
	7	GO:0030662	coated vesicle membrane	2	32	7	0,036
	7	GO:0030125	clathrin vesicle coat	2	15	15	0,008
	7	GO:0030659	cytoplasmic vesicle membrane	2	38	6	0,046
	5	GO:0048475	coated membrane	2	29	8	0,032
	7	GO:0005856	cytoskeleton	7	264	3	0,009
	7	GO:0000323	lytic vacuole	2	33	7	0,036
	7	GO:0005813	centrosome	6	114	6	0,001
	7	GO:0030118	clathrin coat	2	16	14	0,009
	7	GO:0030120	vesicle coat	2	28	8	0,026
	7	GO:0016023	cytoplasmic membrane-bound vesicle	9	255	4	0,001
	7	GO:0044430	cytoskeletal part	6	204	3	0,02
	7	GO:0005773	vacuole	3	59	6	0,02
	7	GO:0005815	microtubule organizing center	6	117	6	0,001
	7	GO:0030117	membrane coat	2	29	8	0,032
	7	GO:0031410	cytoplasmic vesicle	9	257	4	0,001
	7	GO:0044444	cytoplasmic part	20	1210	2	0,008
	6	GO:0005737	cytoplasm	20	1431	2	0,021
	5	GO:0005625	soluble fraction	2	31	7	0,033

## Anhang D

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>AX2 &lt; JH.D biologischer Prozess</b>	4	GO:0042493	response to drug	2	31	8	0,038
	3	GO:0006807	nitrogen compound metabolic process	5	196	3	0,038
	5	GO:0016051	carbohydrate biosynthetic process	2	33	7	0,038
	3	GO:0009058	biosynthetic process	10	656	2	0,039
	5	GO:0016051	carbohydrate biosynthetic process	2	33	7	0,038
	5	GO:0006855	multidrug transport	2	6	41	0,002
	4	GO:0015893	drug transport	2	15	16	0,01
	3	GO:0006810	transport	12	741	2	0,015
	5	GO:0019752	carboxylic acid metabolic process	7	283	3	0,01
	4	GO:0006082	organic acid metabolic process	7	284	3	0,01
	7	GO:0009126	purine nucleoside monophosphate metabolic process	4	16	31	0
	7	GO:0009161	ribonucleoside monophosphate metabolic process	4	16	31	0
	6	GO:0009123	nucleoside monophosphate metabolic process	4	17	29	0
	7	GO:0009144	purine nucleoside triphosphate metabolic process	2	26	10	0,038
	7	GO:0009199	ribonucleoside triphosphate metabolic process	2	26	10	0,038
	6	GO:0009141	nucleoside triphosphate metabolic process	2	32	8	0,038
	7	GO:0009150	purine ribonucleotide metabolic process	5	40	15	0
	6	GO:0006163	purine nucleotide metabolic process	6	43	17	0
	6	GO:0009259	ribonucleotide metabolic process	5	42	15	0
	5	GO:0009117	nucleotide metabolic process	6	97	8	0
	4	GO:0046483	heterocycle metabolic process	3	60	6	0,015
	7	GO:0009127	purine nucleoside monophosphate biosynthetic process	4	14	35	0
	7	GO:0009156	ribonucleoside monophosphate biosynthetic process	4	14	35	0
	6	GO:0009124	nucleoside monophosphate biosynthetic process	4	15	33	0
	7	GO:0009144	purine nucleoside triphosphate metabolic process	2	26	10	0,038
	8	GO:0009201	ribonucleoside triphosphate biosynthetic process	2	25	10	0,017
	7	GO:0009142	nucleoside triphosphate biosynthetic process	2	29	9	0,038
	7	GO:0009152	purine ribonucleotide biosynthetic process	5	38	16	0
	6	GO:0006164	purine nucleotide biosynthetic process	6	41	18	0
	6	GO:0009260	ribonucleotide biosynthetic process	5	40	15	0
	5	GO:0009165	nucleotide biosynthetic process	6	68	11	0
	4	GO:0044249	cellular biosynthetic process	10	503	2	0,01
	7	GO:0006559	L-phenylalanine catabolic process	2	4	62	0
	7	GO:0006572	tyrosine catabolic process	2	5	49	0,002
	5	GO:0019439	aromatic compound catabolic process	2	8	31	0,002
	6	GO:0046112	nucleobase biosynthetic process	2	11	22	0,01
	6	GO:0006144	purine base metabolic process	2	11	22	0,01
	5	GO:0009112	nucleobase metabolic process	3	24	15	0,002
	4	GO:0006725	aromatic compound metabolic process	5	64	10	0
	4	GO:0006519	amino acid and derivative metabolic process	5	171	4	0,014
	7	GO:0006558	L-phenylalanine metabolic process	2	4	62	0
7	GO:0006570	tyrosine metabolic process	2	5	49	0,002	
6	GO:0009072	aromatic amino acid family metabolic process	2	12	21	0,01	
7	GO:0006563	L-serine metabolic process	2	9	27	0,01	
6	GO:0009069	serine family amino acid metabolic process	2	24	10	0,017	

## Anhang D

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>AX2 &lt; JH.D biologischer Prozess</b>	5	GO:0006520	amino acid metabolic process	5	152	4	0,01
	4	GO:0009308	amine metabolic process	5	181	3	0,016
	7	GO:0009074	aromatic amino acid family catabolic process	2	7	35	0,002
	6	GO:0009063	amino acid catabolic process	2	34	7	0,038
	7	GO:0006090	pyruvate metabolic process	2	14	18	0,01
	7	GO:0006006	glucose metabolic process	2	35	7	0,038
	7	GO:0019319	hexose biosynthetic process	2	17	15	0,01
	6	GO:0046364	monosaccharide biosynthetic process	2	17	15	0,01
	5	GO:0046165	alcohol biosynthetic process	2	17	15	0,01
	5	GO:0007032	endosome organization and biogenesis	2	7	35	0,002
	6	GO:0006897	endocytosis	3	88	4	0,039
5	GO:0010324	membrane invagination	3	88	4	0,039	

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>AX2 &lt; JH.D molekulare Funktion</b>	6	GO:0015405	P-P-bond-hydrolysis-driven transmembrane transporter activity	7	117	8	0
	5	GO:0015399	primary active transmembrane transporter activity	7	117	8	0
	4	GO:0022804	active transmembrane transporter activity	7	151	6	0
	3	GO:0022857	transmembrane transporter activity	7	248	4	0,004
	2	GO:0005215	transporter activity	9	357	3	0,002
	6	GO:0042626	ATPase activity\, coupled to transmembrane movement of substances	7	113	8	0
	5	GO:0016820	hydrolase activity\, acting on acid anhydrides\, catalyzing transmembrane movement of substances	7	116	8	0
	7	GO:0017111	nucleoside-triphosphatase activity	10	352	4	0
	6	GO:0016462	pyrophosphatase activity	10	358	4	0
	5	GO:0016818	hydrolase activity\, acting on acid anhydrides\, in phosphorus-containing anhydrides	10	358	4	0
	4	GO:0016817	hydrolase activity\, acting on acid anhydrides	10	361	4	0,001
	3	GO:0016787	hydrolase activity	14	1022	2	0,02
	5	GO:0016831	carboxy-lyase activity	2	24	11	0,016
	4	GO:0016830	carbon-carbon lyase activity	2	32	8	0,025
	3	GO:0016829	lyase activity	5	96	7	0,001
	2	GO:0003824	catalytic activity	33	2955	1,4	0,004
	7	GO:0005524	ATP binding	12	783	2	0,016
	6	GO:0032559	adenyl ribonucleotide binding	12	797	2	0,017
	5	GO:0030554	adenyl nucleotide binding	13	837	2	0,011
	5	GO:0032555	purine ribonucleotide binding	14	1007	2	0,018
4	GO:0017076	purine nucleotide binding	16	1059	2	0,006	
4	GO:0032553	ribonucleotide binding	14	1007	2	0,018	
3	GO:0000166	nucleotide binding	17	1189	2	0,008	

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>AX2 &lt; JH.D zelluläres Kompartiment</b>	4	GO:0031988	membrane-bound vesicle	9	258	4	0,001
	3	GO:0031982	vesicle	9	260	4	0,001
	7	GO:0045335	phagocytic vesicle	8	193	5	0
	7	GO:0030139	endocytic vesicle	8	200	4	0,001
	7	GO:0016023	cytoplasmic membrane-bound vesicle	9	255	4	0,001
	7	GO:0031410	cytoplasmic vesicle	9	257	4	0,001

## E Liste der Anreicherungen biologischer Prozesse, molekularer Funktionen und zelluläre Kompartimente in den differentiell exprimierten Genen der Clusteranalyse

Ergebnis der Genontologie-Analyse mithilfe des Programmpaketes **Gene Ontology Analysis Tool** (GOAT, Xu und Shaulsky 2005) für die unter Kapitel 3.3 definierten Cluster der transkriptionellen Antwort auf CbfA und dessen Carboxy-terminalen Domäne. Erläuterung siehe Anhang D.

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
Cluster 1 biologischer Prozess	6	GO:0042594	response to starvation	5	158	7	0,003
	5	GO:0031667	response to nutrient levels	5	159	7	0,003
	4	GO:0009991	response to extracellular stimulus	5	161	7	0,003
	3	GO:0009605	response to external stimulus	5	230	5	0,009
	3	GO:0006950	response to stress	6	392	3	0,011
	2	GO:0050896	response to stimulus	6	513	3	0,033
	2	GO:0000003	reproduction	5	236	5	0,009
	2	GO:0032501	multicellular organismal process	5	288	4	0,011
	4	GO:0006629	lipid metabolic process	4	276	3	0,036
	3	GO:0044238	primary metabolic process	18	2752	1,4	0,033
	3	GO:0051674	localization of cell	3	85	8	0,011
	2	GO:0051179	localization	8	806	2	0,033
	4	GO:0030587	sorocarp development	5	150	7	0,002
	3	GO:0048856	anatomical structure development	5	299	4	0,013
	4	GO:0055084	fruiting body development in response to starvation	5	150	7	0,002
	3	GO:0030582	fruiting body development	5	150	7	0,002
	3	GO:0007275	multicellular organismal development	5	273	4	0,011
	2	GO:0032502	developmental process	5	411	3	0,036
	6	GO:0006508	proteolysis	7	212	7	0
	6	GO:0006665	sphingolipid metabolic process	2	27	16	0,011
	5	GO:0006643	membrane lipid metabolic process	3	84	8	0,011
	7	GO:0006907	pinocytosis	2	20	22	0,009
	3	GO:0006928	cell motility	3	85	8	0,011
	7	GO:0042593	glucose homeostasis	2	5	87	0,001
	6	GO:0033500	carbohydrate homeostasis	2	5	87	0,001
	5	GO:0048878	chemical homeostasis	3	31	21	0,002
	4	GO:0042592	homeostatic process	3	60	11	0,008
	3	GO:0065008	regulation of biological quality	4	115	8	0,006
	5	GO:0051246	regulation of protein metabolic process	2	60	7	0,033
	6	GO:0051248	negative regulation of protein metabolic process	2	9	49	0,003
	5	GO:0009892	negative regulation of metabolic process	2	35	12	0,013
	4	GO:0016337	cell-cell adhesion	2	18	24	0,009

## Anhang E

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>Cluster 1 molekulare Funktion</b>	7	GO:0022890	inorganic cation transmembrane transporter activity	2	57	9	0,035
	6	GO:0016298	lipase activity	2	34	14	0,018
	5	GO:0004091	carboxylesterase activity	2	50	10	0,029
	6	GO:0004197	cysteine-type endopeptidase activity	4	29	34	0
	5	GO:0008234	cysteine-type peptidase activity	4	58	17	0
	5	GO:0004175	endopeptidase activity	4	106	9	0,002
	4	GO:0008233	peptidase activity	6	197	7	0
	3	GO:0016787	hydrolase activity	10	1022	2	0,012

<b>Cluster 1 zelluläres Kompartiment</b>	7	GO:0030136	clathrin-coated vesicle	2	19	23	0,004
	7	GO:0045335	phagocytic vesicle	5	193	6	0,002
	4	GO:0031988	membrane-bound vesicle	7	258	6	0
	3	GO:0031982	vesicle	7	260	6	0
	2	GO:0043226	organelle	15	1919	2	0,009
	2	GO:0005576	extracellular region	5	88	12	0
	7	GO:0030665	clathrin coated vesicle membrane	2	15	29	0,002
	7	GO:0015630	microtubule cytoskeleton	3	139	5	0,025
	7	GO:0005764	lysosome	2	30	14	0,009
	7	GO:0030139	endocytic vesicle	6	200	6	0,001
	7	GO:0030135	coated vesicle	2	36	12	0,013
	7	GO:0030662	coated vesicle membrane	2	32	14	0,009
	7	GO:0030125	clathrin vesicle coat	2	15	29	0,002
	5	GO:0031090	organelle membrane	4	245	4	0,025
	6	GO:0012506	vesicle membrane	2	41	11	0,018
	5	GO:0048475	coated membrane	2	29	15	0,009
	7	GO:0005856	cytoskeleton	4	264	3	0,031
	7	GO:0000323	lytic vacuole	2	33	13	0,013
	7	GO:0005813	centrosome	3	114	6	0,015
	7	GO:0030118	clathrin coat	2	16	27	0,004
	7	GO:0030120	vesicle coat	2	28	15	0,007
	7	GO:0030659	cytoplasmic vesicle membrane	2	38	11	0,013
	7	GO:0016023	cytoplasmic membrane-bound vesicle	7	255	6	0
	6	GO:0043229	intracellular organelle	15	1918	2	0,009
	7	GO:0005773	vacuole	3	59	11	0,004
	7	GO:0005815	microtubule organizing center	3	117	6	0,018
	7	GO:0030117	membrane coat	2	29	15	0,009
	7	GO:0044433	cytoplasmic vesicle part	2	39	11	0,014
	7	GO:0031410	cytoplasmic vesicle	7	257	6	0
	7	GO:0044444	cytoplasmic part	14	1210	3	0
6	GO:0005737	cytoplasm	14	1431	2	0,002	

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>Cluster 3 biologischer Prozess</b>	5	GO:0000270	peptidoglycan metabolic process	3	12	65	0
	6	GO:0009253	peptidoglycan catabolic process	3	10	77	0
	5	GO:0016052	carbohydrate catabolic process	3	48	16	0,002
	4	GO:0009057	macromolecule catabolic process	3	155	5	0,027
	3	GO:0009056	catabolic process	4	266	4	0,025
	7	GO:0016998	cell wall catabolic process	3	14	55	0
	6	GO:0010382	cell wall metabolic process	3	14	55	0
	5	GO:0007047	cell wall organization and biogenesis	3	23	34	0
	4	GO:0045229	external encapsulating structure organization and biogenesis	3	23	34	0



## Anhang E

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>Cluster 3 molekulare Funktion</b>	6	GO:0003796	lysozyme activity	3	10	86	0
	5	GO:0004553	hydrolase activity\, hydrolyzing O-glycosyl compounds	3	78	11	0,003
	4	GO:0016798	hydrolase activity\, acting on glycosyl bonds	3	87	10	0,004
	6	GO:0004197	cysteine-type endopeptidase activity	2	29	20	0,005
	5	GO:0008234	cysteine-type peptidase activity	2	58	10	0,017
	5	GO:0000287	magnesium ion binding	2	45	13	0,011

<b>Cluster 3 zelluläres Kompartiment</b>	7	GO:0005815	microtubule organizing center	3	117	7	0,009
	7	GO:0005813	centrosome	3	114	7	0,009

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>Cluster 5 biologischer Prozess</b>	3	GO:0009605	response to external stimulus	4	230	3	0,045
	4	GO:0042493	response to drug	2	31	11	0,016
	3	GO:0042221	response to chemical stimulus	4	178	4	0,028
	3	GO:0006807	nitrogen compound metabolic process	4	196	4	0,031
	5	GO:0006855	multidrug transport	2	6	57	0,001
	4	GO:0015893	drug transport	2	15	23	0,013
	5	GO:0019752	carboxylic acid metabolic process	5	283	3	0,028
	4	GO:0006082	organic acid metabolic process	5	284	3	0,028
	7	GO:0009126	purine nucleoside monophosphate metabolic process	3	16	32	0
	7	GO:0009161	ribonucleoside monophosphate metabolic process	3	16	32	0
	6	GO:0009123	nucleoside monophosphate metabolic process	3	17	30	0
	7	GO:0009150	purine ribonucleotide metabolic process	4	40	17	0
	6	GO:0006163	purine nucleotide metabolic process	5	43	20	0
	6	GO:0009259	ribonucleotide metabolic process	4	42	16	0
	5	GO:0009117	nucleotide metabolic process	5	97	9	0
	5	GO:0044270	nitrogen compound catabolic process	2	41	8	0,028
	7	GO:0009127	purine nucleoside monophosphate biosynthetic process	3	14	37	0
	7	GO:0009156	ribonucleoside monophosphate biosynthetic process	3	14	37	0
	6	GO:0009124	nucleoside monophosphate biosynthetic process	3	15	34	0
	7	GO:0009152	purine ribonucleotide biosynthetic process	4	38	18	0
	6	GO:0006164	purine nucleotide biosynthetic process	5	41	21	0
	6	GO:0009260	ribonucleotide biosynthetic process	4	40	17	0
	5	GO:0009165	nucleotide biosynthetic process	5	68	13	0
	7	GO:0006559	L-phenylalanine catabolic process	2	4	86	0
	7	GO:0006572	tyrosine catabolic process	2	5	69	0,001
	5	GO:0019439	aromatic compound catabolic process	2	8	43	0,001
	6	GO:0006144	purine base metabolic process	2	11	31	0,009
	5	GO:0009112	nucleobase metabolic process	2	24	14	0,014
	4	GO:0006725	aromatic compound metabolic process	4	64	11	0,001
	4	GO:0006519	amino acid and derivative metabolic process	4	171	4	0,017
	7	GO:0009074	aromatic amino acid family catabolic process	2	7	49	0,001
	6	GO:0009063	amino acid catabolic process	2	34	10	0,017
	5	GO:0009310	amine catabolic process	2	41	8	0,028
	7	GO:0006558	L-phenylalanine metabolic process	2	4	86	0
	7	GO:0006570	tyrosine metabolic process	2	5	69	0,001
	6	GO:0009072	aromatic amino acid family metabolic process	2	12	29	0,009
	5	GO:0006520	amino acid metabolic process	4	152	5	0,014
	4	GO:0009308	amine metabolic process	4	181	4	0,028
	5	GO:0007032	endosome organization and biogenesis	2	7	49	0,001
	6	GO:0007186	G-protein coupled receptor protein signaling pathway	2	53	6	0,041

## Anhang E

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>Cluster 5 molekulare Funktion</b>	6	GO:0015405	P-P-bond-hydrolysis-driven transmembrane transporter activity	4	117	6	0,006
	5	GO:0015399	primary active transmembrane transporter activity	4	117	6	0,006
	4	GO:0022804	active transmembrane transporter activity	4	151	5	0,012
	3	GO:0022857	transmembrane transporter activity	4	248	3	0,046
	2	GO:0005215	transporter activity	5	357	3	0,045
	5	GO:0030554	adenyl nucleotide binding	9	837	2	0,036
	4	GO:0017076	purine nucleotide binding	11	1059	2	0,026
	3	GO:0000166	nucleotide binding	12	1189	2	0,025
	6	GO:0042626	ATPase activity\, coupled to transmembrane movement of substances	4	113	6	0,005
	5	GO:0016820	hydrolase activity\, acting on acid anhydrides\, catalyzing transmembrane movement of substances	4	116	6	0,006
	7	GO:0017111	nucleoside-triphosphatase activity	6	352	3	0,014
	6	GO:0016462	pyrophosphatase activity	6	358	3	0,015
	5	GO:0016818	hydrolase activity\, acting on acid anhydrides\, in phosphorus-containing anhydrides	6	358	3	0,015
	4	GO:0016817	hydrolase activity\, acting on acid anhydrides	6	361	3	0,016
	4	GO:0016879	ligase activity\, forming carbon-nitrogen bonds	3	97	6	0,019
	5	GO:0016831	carboxy-lyase activity	2	24	15	0,01
	4	GO:0016830	carbon-carbon lyase activity	2	32	11	0,016
3	GO:0016829	lyase activity	4	96	8	0,003	
2	GO:0003824	catalytic activity	22	2955	1,3	0,041	

<b>Cluster 5 zelluläres Kompartiment</b>	4	GO:0031988	membrane-bound vesicle	7	258	4	0,001
	3	GO:0031982	vesicle	7	260	4	0,001
	7	GO:0045335	phagocytic vesicle	6	193	5	0,001
	7	GO:0030139	endocytic vesicle	6	200	5	0,001
	5	GO:0005886	plasma membrane	4	152	4	0,02
	7	GO:0016023	cytoplasmic membrane-bound vesicle	7	255	4	0,001
	7	GO:0005768	endosome	2	35	9	0,02
7	GO:0031410	cytoplasmic vesicle	7	257	4	0,001	

Beschreibung	GO Level	GO ID	GO Annotation	Liste	Total	Anr.	p-Wert
<b>Cluster 6 biologischer Prozess</b>	3	GO:0009058	biosynthetic process	4	656	3	0,024
	2	GO:0051179	localization	4	806	3	0,044
	3	GO:0006810	transport	4	741	3	0,041
	2	GO:0051234	establishment of localization	4	750	3	0,041
	4	GO:0044249	cellular biosynthetic process	4	503	5	0,009

<b>Cluster 6 molekulare Funktion</b>	6	GO:0015405	P-P-bond-hydrolysis-driven transmembrane transporter activity	2	117	10	0,019
	5	GO:0015399	primary active transmembrane transporter activity	2	117	10	0,019
	4	GO:0022804	active transmembrane transporter activity	2	151	8	0,027
	2	GO:0005215	transporter activity	3	357	5	0,022
	6	GO:0042626	ATPase activity\, coupled to transmembrane movement of substances	2	113	10	0,019
	5	GO:0016820	hydrolase activity\, acting on acid anhydrides\, catalyzing transmembrane movement of substances	2	116	10	0,019
	7	GO:0017111	nucleoside-triphosphatase activity	3	352	5	0,022
	6	GO:0016462	pyrophosphatase activity	3	358	5	0,022
	5	GO:0016818	hydrolase activity\, acting on acid anhydrides\, in phosphorus-containing anhydrides	3	358	5	0,022
	4	GO:0016817	hydrolase activity\, acting on acid anhydrides	3	361	5	0,023

## F Sequenzierung von Einzelklonen der Hefe-2-Hybrid-Banken

oben: oligo (dT)-Bank, unten: (dN)<sub>6</sub>-Bank; Größe entspricht der geschätzten PCR-Banden-Größe, Match ist Übereinstimmung mit cDNA, Gesamt die Gesamtlänge der jeweiligen cDNA;

\*Klon 8 der (dN)<sub>6</sub>-Bank mit Fusion aus zwei Proteinen

Klon	Größe	Insert	DDB_G ID	Endet am	Match	Gesamt	%
1	180	leer					
2	580	pyrroline-5-carboxylate reductase	DDB_G0269410	3' Ende	448	900	50
3	380	mtDNA (EST: ddv46g10; ddv46i22)		M:25708-26049	341		
4 untere	400	actin 8	DDB_G0269234	51bp vom 3'	284	1131	25
4 obere	750	LTR-RETROTRANSPOSON SKIPPER, GAG	4 identische	3' Ende	720	1185	61
5	580	mitochondrial large subunit rRNA (rnlA)	DDB_G0294034	3' Ende Exon 1	330	2871	11
6	370	unbekannter ORF	DDB_G0294605	3' Ende	141	141	100
7	400	unbekannter ORF (Citratlyase)	DDB_G0284067	504bp vom 3'	264	1083	24
8 untere	420	pyruvate dehydrogenase E1 alpha subunit (pdh1)	DDB_G0292994	311bp vom 3'	375	1134	33
8 obere	950	tailless complex polypeptide 1 (tcp1)	DDB_G0269190	20bp vom 3'	820	1647	50
9	180	leer					
10	400	ubiquitin	DDB_G0276765	3' Ende	314	465	68
11	1200	transcription factor IIH component (repD)	DDB_G0267414	3' Ende	904	2331	39
12	180	leer					
13	520	S60 ribosomal protein L18a (rpl18a)	DDB_G0285881	3' Ende	503	516	97
14	970	unbekannter ORF	DDB_G0293014	3' Ende	808	1461	55
15	180	leer					

Klon	Größe	Insert	DDB_G ID	Endet am	Match	Gesamt	%
1	180	leer					
2	350	(+/+) extrachromosomal palindromic ribosomal RNA gene (17S rRNA-2)	DDB_G0295647	R:67781-67982	173		
3	350	rap1 gene for ras-related gene product	DDB_G0291237	3' Ende	175	561	31
4	500	RasGTPase-activating protein (rgaA)	DDB_G0287585	644bp vom 3'	439	2469	18
5	400	SCAR1 (scrA)	DDB_G0285253	614bp vom 3'	100	1332	8
6	700	hypothetical protein	DDB_G0288707	1537bp vom 3'	398	2958	13
7	400	hypothetical protein --> putative delta-24-sterol methyltransferase	DDB_G0288907	178bp vom 3'	252	1065	24
8*	900	hypothetical protein	DDB_G0287817	1683bp vom 3'	484	3072	16
		hypothetical protein	DDB_G0278453	1755bp vom 3'	201	2472	8
9	180	leer					
10	450	hypothetical protein	DDB_G0282631	3' Ende	273	672	41
11	450	hypothetical protein Similar to Homo sapiens (Human). Myotubularin-related protein 2	DDB_G0271610	311bp vom 3'	408	5217	8
12	800	hypothetical protein	DDB_G0281279	377bp vom 3'	310	2037	15
13	1200	(+/+) extrachromosomal palindromic ribosomal RNA gene (26S rRNA-2)	DDB_G0295653	R:70702-71829	1127		
14	1200	(+/-) extrachromosomal palindromic ribosomal RNA gene (17S rRNA-2)	DDB_G0295649	R:72282-72952	670		
15	650	(+/+) extrachromosomal palindromic ribosomal RNA gene (17S rRNA-2)	DDB_G0295649	R:71397-71834	437		
16	600	(+/+) extrachromosomal palindromic ribosomal RNA gene (17S rRNA-2)	DDB_G0295649	R:71392-71833	442		
17	1000	(+/+) extrachromosomal palindromic ribosomal RNA gene	DDB_G0295649	R:71392-71833	346		
18	180	leer					

## Liste der wissenschaftlichen Publikationen

**Lucas J., Bilzer A., Moll L., Zündorf I., Dingermann T., Eichinger L., Siol O., Winckler T.** (2009)

The carboxy-terminal domain of *Dictyostelium* C-module-binding factor is an independent gene regulatory entity. *PLOS ONE* **4(4)**: e5012. Epub 2009 Apr 3.

**Lucas J., Dölz H., Bilzer A., Siol O., Winckler T.** (2007)

Two-hybrid screening for binding partners of C-module-binding factor CbfA. Posterpräsentation Annual International *Dictyostelium* Conference 2007, Rotenburg an der Fulda (Deutschland)

**Dölz H. und Lucas J., Winckler T.** (2007)

The multifunctional signal protein CbfA. Posterpräsentation 2<sup>nd</sup> meeting of the Joint Research Center SFB604 2007, Jena (Deutschland)

**Lucas J., Bilzer A., Moll L., Bohnert M., Urushihara H., Eichinger L., Glöckner G., Siol O., Winckler T.** (2009) The carboxy-terminal domain of *Dictyostelium* C-module-binding factor (CbfA) is an independent gene regulatory entity. Vortrag Annual International *Dictyostelium* Conference 2009, Estes Park (USA)

## Danksagung

Die vorliegende Arbeit wurde zwischen März 2006 und Oktober 2009 am Lehrstuhl für Pharmazeutische Biologie, Institut für Pharmazie der Friedrich Schiller Universität Jena durchgeführt. Ich möchte nachfolgend allen Personen danken, welche zum Gelingen der Arbeit beigetragen haben.

Meinem Betreuer Herrn Prof. Dr. Thomas Winckler möchte ich für die Überlassung dieses facettenreichen Themas danken. Er gab mir die Möglichkeit, mich an einem sehr breiten Spektrum von Methoden auszuprobieren. Dabei unterstützte er mich stets durch fachliches Wissen, stetige Motivation und konstruktive Diskussionen.

Für die Finanzierung großer Teile meiner Arbeit möchte ich den Sonderforschungsbereich 604 danken. Nicht weniger wertvoll war die Möglichkeit zur Kontaktaufnahme mit einigen dort vertretenen Arbeitsgruppen in Jena und Halle, ohne deren Hilfe viele der durchgeführten Experimente nicht möglich gewesen wären. In diesem Zusammenhang möchte ich auch meinen Zweitbetreuern Prof. Dr. Thorsten Heinzl des Lehrstuhls für Biochemie der Universität Jena und Dr. Frank Hänel der Abteilung Zell- und Molekularbiologie des Leibniz-Institut für Naturstoff-Forschung und Infektionsbiologie – Hans-Knöll-Institut (Jena) danken.

Des Weiteren möchte ich Dr. Constanze Than-Breithaupt aus der Abteilung Physikalische Biotechnologie der Martin-Luther-Universität Halle-Wittenberg danken. Neben den technischen Möglichkeiten der Proteinbiochemie teilte sie vor allem ihr schier unendliches Wissen über dieses Thema mit mir. Die Aufenthalte in ihrem Labor in Halle und die dortige Zusammenarbeit mit Uwe Fandrich waren nicht nur methodisch sehr aufschlussreich. Für die Aufnahme des Protonenspektrums danke ich Dr. Ulrich Weininger, Institut für Physik, Biophysik, Martin-Luther-Universität Halle-Wittenberg. Die Amino-terminale Sequenzierung der Proteinfragmente wurde durch die Arbeitsgruppe von Prof. Dr. Grosse am Leibniz-Institut für Altersforschung – Fritz-Lipmann-Institut e.V. (Jena) durchgeführt. Für die Möglichkeit der Benutzung der French Press danke ich dem Lehrstuhl für Angewandte und ökologische Mikrobiologie Jena.

Privatdozent Dr. Ludwig Eichinger des Institutes Biochemie I der Universität Köln danke ich für die Möglichkeit der Durchführung der *Microarray*-Experimente. Neben der Organisation meines einmonatigen Aufenthaltes in Köln danke ich ihn für seine stetige, unkomplizierte Hilfe bei der Bearbeitung dieser komplexen Thematik. Für die Einführung in diese Methodik danke ich Lorna Moll.

Der Arbeitsgruppe von Dr. Gernot Glöckner am Leibniz-Institut für Altersforschung – Fritz-Lipmann-Institut e.V. (Jena) danke ich für die Erstellung der Transkriptomdaten sowie den Einblick in einige unveröffentlichte Genomdaten. Den Diplomanten Markus Bohnert und Michael Sauer danke ich für die Zusammenarbeit zur Phylogenie von CbfA. Bei Dr. Oliver Siol möchte ich mich für die zahllosen Diskussionen und damit verbundener Hilfe über die Auswertung von Transkriptomdaten bedanken.

Dr. Alexander Gehrke, Andreas Thywißen und Martin Vödisch aus der Abteilung Molekulare und Angewandte Mikrobiologie des Leibniz-Institut für Naturstoff-Forschung und Infektionsbiologie – Hans-Knöll-Institut (Jena) danke ich für die technische Hilfe bei der Erstellung der mikroskopischen Bilder. Claudia Franke und Dr. Frank Hänel aus der Zell- und Molekularbiologie des HKIs danke ich für die Überlassung des Hefestammes KFY sowie für ihre Unterstützung zum Thema Hefe-2-Hybrid.

Ich danke allen Mitarbeitern des Lehrstuhles für pharmazeutische Biologie für ein wunderbares Arbeitsklima, welches nicht nur theoretisch und praktisch zum Gelingen dieser Arbeit beigetragen hat. Es war mir stets eine Freude mit ihnen arbeiten zu dürfen!

Abschließend möchte ich mich bei meiner Familie bedanken, welche immer hinter mir stand. Inka danke ich ein wenig für das Korrekturlesen und unendlich für ihre Liebe.

## Lebenslauf

Name: Jörg Lucas  
Geburtsdatum: 09.10.1980  
Geburtsort: Eisenach  
Anschrift: Dornburger Str. 83, 07743 Jena

### Schulische Bildung:

09/1987-07/1991 Grundschule Thal  
08/1991-07/1999 „Albert-Schweitzer-Gymnasium“ Ruhla  
30.06.1999 Allgemeine Hochschulreife, Abitur

### Studium:

10/2000-02/2006 Biologie (Diplom) an der Friedrich-Schiller-Universität Jena  
10/2002-02/2005 Hauptfach Mikrobiologie, Nebenfächer Genetik und Ökologie  
03/2005-02/2006 Diplomarbeit zum Thema „Neurodegenerative Krankheiten und Apoptose: potentielle Ziele basierend auf neuen Protein-Protein-Interaktionen“ in der Abteilung Zell- und Molekularbiologie des Leibniz-Institut für Naturstoff-Forschung und Infektionsbiologie – Hans-Knöll-Institut (Jena)  
03/2006-12/2009 Wissenschaftlicher Mitarbeiter am Institut für Pharmazie, Lehrstuhl für Pharmazeutische Biologie, Friedrich-Schiller-Universität Jena; Promotion zum Thema „Struktur und Funktion der Carboxy-terminalen Domäne des Transkriptionsfaktors CbfA aus *Dictyostelium discoideum*“ unter der Betreuung von Prof. Dr. Thomas Winckler

## **Ehrenwörtliche Erklärung**

Hiermit erkläre ich,

- dass mir die geltende Promotionsordnung der Biologisch-Pharmazeutischen Fakultät bekannt ist,
- dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Hilfsmittel, persönlichen Mitteilungen und Quellen angefertigt habe,
- dass ich die Hilfe eines Promotionsberaters nicht in Anspruch genommen habe und
- dass ich die vorliegende Arbeit weder komplett noch in Form einer in wesentlichen Teilen ähnlichen oder anderen Abhandlung bei der Friedrich-Schiller-Universität oder einer anderen Hochschule als Dissertation oder Prüfungsarbeit für eine staatliche oder wissenschaftliche Prüfung vorgelegt habe.

Jena, den 24. November 2009

Jörg Lucas