Transcriptional regulation during fruiting body formation in the basidiomycete *Coprinopsis cinerea*: promoter analysis of the *cgI2* gene encoding a fruiting body-specific galectin

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH

for the degree of

Doctor of Natural Sciences

presented by

RINALDO CAMILLO BERTOSSA

Dipl. Natw. ETH

born 22.2.1972

citizen of Rossa (GR)

accepted on the recommendation of

Prof. Dr. M. Aebi, examiner

Prof. Dr. U. Kües, co-examiner

Dr. M. Künzler, co-examiner

2004
Acknowledgments

I thank Prof. Dr. Markus Aebi for allowing me to make a PhD work in his group and especially for being a continuous source of new and stimulating ideas. I thank him additionally for his steady efforts to improve working conditions here at the institute and outside.

I thank Prof. Dr. Ursula Kües for supervising the work. In particular, I thank her for teaching me the “short way” when looking for solutions: a great help for my future career in the lab and outside.

I thank Dr. Markus Künzler for taking the supervision after the leave of Ursula Kües. He was a firm reference point and a capable teacher during the last two years.

I thank Dr. Piers Walser with whom I shared lab E25 during these years. I thank him especially for his steady disposition to listen to any sort of difficulty and for showing me the way of the “integrative thinking”.

I thank present and former members of the „Coprinus“ and of the Aebi group for continuous support and for the highly agreeable working atmosphere.

I thank all those who helped in many different ways: all members of the Institute of Microbiology; Alain and Jacques for the great management at the institute; Silvia, Chiara, Antje, Hélène, Paul, Palmira, Carmen and Serife for providing a comfortable working environment.

Finally, I thank many people outside the institute for continuous support. A special thank go to Paul, Hans, Luca and Markus for sharing difficult and happy experiences during the last five years.

I thank my parents for transmitting me the desire to live for great ideals: the hidden anchor in this work.
CONTENTS

Summary .................................................................................................................. 1
Riassunto ................................................................................................................ 3

CHAPTER I

Introduction .......................................................................................................... 5
1. The model organism Coprinopsis cinerea ......................................................... 5
   1.1. Life cycle .................................................................................................... 5
1.2. Fruiting body development in C. cinerea ..................................................... 7
   1.2.1. Mating type genes .............................................................................. 10
   1.2.2. Nutrients ............................................................................................ 11
   1.2.3. Light .................................................................................................... 11
2. Control of gene transcription in fungi during developmental processes .... 11
   2.1. Transcriptional regulation of fungal genes .............................................. 12
   2.2. Trans-acting proteins and their DNA targets in fungal genes ............... 15
      2.2.1. Transcription factors containing basic domains ......................... 15
      2.2.2. Transcription factors containing zinc-coordinating DNA-binding domains 19
      2.2.3. Transcription factors containing the helix-turn-helix domain ....... 23
      2.2.4. β-scaffold transcription factors with minor groove contacts ......... 24
   2.3. Fungal cis- and trans-acting elements in their regulatory networks .... 28
      2.3.1. Mating type genes ........................................................................ 28
      2.3.2. Nutrients ......................................................................................... 34
      2.3.3. Light ................................................................................................ 40
   2.4. Concluding remarks ................................................................................. 42
3. Transcriptional regulation of fruiting body development in C. cinerea .... 43
4. Aim of the thesis ............................................................................................ 45
5. References ...................................................................................................... 47

CHAPTER II

Promoter analysis of cg/2, a galectin encoding gene transcribed during fruiting body formation in Coprinopsis cinerea (Coprinus cinereus) .............................................................................................................. 59
1. Abstract .......................................................................................................... 60
2. Introduction ................................................................................................... 61
3. Materials and methods .................................................................................. 63
4. Results .......................................................................................................... 70
5. Discussion .................................................................................................... 80
6. References .................................................................................................... 88
7. Annex ........................................................................................................... 93
Summary

Basidiomycetous fungi propagate as loose, undifferentiated mycelia in the substrate they live of. The combined action of various stimuli induce on the established mycelium the formation of a variety of highly differentiated structures. One of these, the fruiting body, develops usually on dikaryotic mycelia (mycelia containing two haploid nuclei with complementary mating specificities) and is specialized in the production and dispersal of sexual spores. The formation of the fruiting body is strictly regulated by several factors, of which mating type genes, light and nutrients play a central role. The effects these regulators exert on fruiting body formation are known but the molecular mechanisms through which they act remain yet largely obscure. Regulation of transcription is one possible means to attain control over cellular events. The discussion of some mechanisms of transcriptional regulation during developmental processes in fungi introduces the thesis.

In the basidiomycete Coprinopsis cinerea, two galectins (β-galactoside binding proteins), cgl1 and cgl2, were shown to be exclusively expressed, though at slightly different time points for each protein, in restricted tissues of the young developing fruiting body. To explore to what extent transcriptional control is exploited to regulate fruiting events, the regulation of cgl2 transcription by fruiting-regulating signals was analyzed. For this purpose, C. cinerea was transformed with reporter plasmids containing the gfp coding sequence under control of modified cgl2 promoter fragments. Transcript levels of the endogenous cgl2 and the ectopic integrated gfp gene were subsequently analyzed in C. cinerea transformants by a newly developed analytical method based on real-time PCR technology. The investigation of the cgl2 promoter in a self-compatible homokaryotic strain allowed defining the minimal promoter fragment able to confer wild-type regulated transcription to the gfp reporter gene. In the minimal promoter fragment, a site similar to a Sp1-binding motif functioned as enhancer and a sequence containing a CRE (cAMP responsive element) motif together with AAGG direct repeats was necessary to achieve respectable basal activity as well as transcriptional induction upon prolonged incubation of transformants in continuous darkness. The possible involvement of the CRE motif in cgl2 transcriptional regulation is in agreement with the direct influence cAMP has on fruiting events, as established from previous studies in C. cinerea. The
repression of cg/2 transcription upon incubation of transformants in continuous light could not be linked to any site in the cg/2 promoter and despite the direct control exerted by mating type genes on fruiting, no result indicated presence of docking sites for mating type factors in the cg/2 promoter.

Fruiting body formation on monokaryons is usually repressed as is galectins expression. To explore whether the repression of cg/2 expression is mediated, at transcriptional level, by promoter elements, the gfp reporter constructs were transformed and tested in a monokaryotic C. cinerea strain. Preliminary results indicate that the observed repression was due to specific sites in the cg/2 promoter.

At the end of the thesis, a possible model for the regulation of the cg/2 promoter, based on the findings presented in the thesis and backed up by additional data from the literature, is proposed.
Riassunto

I basidiomiceti si sviluppano sottoforma di micelio indifferenziato nel substrato di cui si nutrono. L’azione combinata di diversi stimoli inducono nel micelio la formazione di varie strutture altamente differenziate. Una di queste, il corpo fruttifero, si sviluppa solitamente su micelio dicariotico (contenente due nuclei aploidi a diversa polarità) ed è specializzato nella produzione e disseminazione delle spore sessuali. Lo sviluppo del corpo fruttifero è strettamente regolato da molteplici fattori dei quali i geni di incrocio, la luce e il nutrimento hanno un ruolo fondamentale. L’influenza che questi hanno nella formazione del corpo fruttifero è nota ma i meccanismi molecolari attraverso i quali essi agiscono sono ancora sconosciuti. La regolazione della trascrizione è uno dei modi attraverso il quale eventi cellulari possono essere controllati. La presentazione di alcuni meccanismi di regolazione della trascrizione depurante i processi di sviluppo dei funghi introduce la tesi.

Nel basidiomicete Coprinopsis cinerea due galectine (proteine che legano il β-galattosio), cg/1 e cg/2, sono espresse esclusivamente, anche se in momenti leggermente distinti per ogni proteina, in tessuti circoscritti del giovane corpo fruttifero in sviluppo. Per capire in che modo la trascrizione è sfruttata come livello di controllo nell’organizzazione degli eventi della fruttificazione, la regolazione della trascrizione di cg/2 da parte dei fattori che inducono la fruttificazione è stato analizzato a livello di promotore in C. cinerea. A tale scopo, C. cinerea è stato trasformato con plasmidi contenenti la sequenza codificante di gfp sotto controllo di frammenti del promotore cg/2 modificati e il livello del trascritto del gene endogeno cg/2 e di gfp è stato analizzato in trasformanti di C. cinerea per mezzo di un nuovo procedimento analitico basato sulla tecnologia “real-time PCR”. L’analisi del promotore di cg/2 in un ceppo omocariotico fruttificante ha permesso di definire il frammento di promotore minimo capace di conferire una trascrizione “wild-type” al gene gfp. Nel frammento del promotore minimo, un sito simile al motivo riconosciuto dal fattore Sp1 funziona come “enhancer” e una sequenza contenente un motivo CRE (elemento che risponde a cAMP) insieme a una ripetizione diretta della sequenza AAGG sono necessari sia per raggiungere livelli appropriati di attività basale sia per provocare un’induzione di trascrizione dopo prolungata incubazione dei trasformanti all’oscurità. La possibile implicazione del motivo CRE nella
regolazione della trascrizione del gene cg/2 è in accordo con l'influenza diretta che cAMP possiede sulla fruttificazione, come stabilito precedentemente da studi con C. cinerea. La repressione della trascrizione di cg/2 dopo incubazione di trasformanti in luce continua non ha potuto essere collegata a nessun sito nel promotore di cg/2 e, nonostante il controllo diretto esercitato dai geni d'incrocio sulla fruttificazione, nessun risultato ha potuto indicare la presenza di possibili motivi di riconoscimento per i prodotti di questi geni nel promotore di cg/2.

La formazione del corpo fruttifero in miceli primari è solitamente repressa come pure l'espressione delle galectine. Per capire se la repressione dell'espressione di cg/2 è mediata, a livello della trascrizione, da elementi nel promotore, i vettori contenenti il gene gfp sono stati trasformati e analizzati in un ceppo aploide di C. cinerea. Risultati preliminari indicano che la repressione potrebbe essere raggiunta con l'ausilio di sequenze specifiche all'interno del promotore di cg/2.

Alla fine della tesi è proposto un possibile modello della regolazione del promotore del cg/2 basato sui risultati ottenuti nella tesi e appoggiato da dati addizionali presi dalla letteratura.
CHAPTER I

Introduction

1. The model organism *Coprinopsis cinerea*

*Coprinopsis cinerea* (Schaeff.) Redhead, Vilgalys & Moncalvo, is a basidiomycetous fungus belonging to the family of the Coprinaceae (order Agaricales). Filamentous fungi are eukaryotic, heterotrophic, absorptive organisms that develop a rather diffuse, branched, tubular (i.e. hyphae, sing. hypha) thallus (mycelium, pl. mycelia) and reproduce by means of spores (Kendrick, 2000). In nature, *C. cinerea* is commonly found on horse dung (Kües, 2000). Because of its tractable genetics and ease of experimental manipulation, *C. cinerea* is used as a model system for studying multicellular and sexual development in basidiomycetes (Kües, 2000).

1.1. Life cycle

The life cycle of *C. cinerea* is outlined in Fig. 1. After germination on a suitable substrate, a meiotic spore (basidiospore), containing a haploid nucleus, gives rise to the monokaryon, an undifferentiated mycelium containing one haploid nucleus per cell segment (Fig. 1). The monokaryon grows and colonizes the substrate. Monokaryons can readily fuse with each other (anastomosis) and, if they differ in their mating type genes (see below), can give rise to a stable dikaryon. The dikaryon is a mycelium in which every cell segment contains two haploid nuclei, one from each respective mating partner. The presence of distinct haploid nuclei in the same cell segment (dikaryosis) is a feature specific to fungi, assured in basidiomycetes, through the synchronous division of the nuclei at every new cell division and partition of the daughter nuclei by specialized cells called "clamp cells" (Fig. 1). The dikaryon develops as a mycelial colony in the substrate. A feature unique to the dikaryon is the ability to form fruiting bodies, differentiated multicellular structures specialized in the production and dispersal of sexual spores (basidiospores, Fig. 1). These are formed
on particular cells termed basidia (sing. basidium) in which karyogamy and meiosis have occurred. Once released in the environment, under suitable conditions each basidiospore will eventually germinate and give rise to a monokaryon. In addition to meiotic spores, other reproductive units are produced on the monokaryotic as well as on the dikaryotic mycelium: oidia are mitotic spores formed on specific sporophores (oidiophores) within aerial hyphae. Haploid uninuclear oidia are constitutively formed on the monokaryon and upon light stimulation on the dikaryon. Large, thick-walled mitotic spores produced within submerged aging hyphae of monokaryons and dikaryons are called chlamydospores (Kües, 2000). In contrast to chlamydospores formed on the monokaryon that have one type of haploid nucleus, chlamydospores formed on the dikaryon contain two different haploid nuclei. Chlamydospores from dikaryons may germinate into two different monokaryons or into a dikaryon, whilst chlamydospores of monokaryons as oidia from both monokaryons and dikaryons germinate to give monokaryons. Depending whether formed on monokaryons or dikaryons, sclerotia (sing. sclerotium) — resting multicellular bodies formed by the aggregation of hyphae — will germinate as monokaryon or dikaryon, respectively (Fig. 1).

Fig. 1: Life cycle of Coprinopsis cinerea (modified after Kües, 2000).
1.2.  **Fruiting body development in C. cinerea**

The fruiting body is the multicellular structure on the dikaryon specialized in the production and dispersal of sexual spores. From a developmental point of view, the morphogenesis of fruiting bodies is regulated by two key parameters: the origin of the hyphal tip cell (i.e. its appropriate placement and orientation on the parental hypha) and the growth direction of the new hyphal apex. The former seems to be the formal equivalent to the morphogenetic growth of plants, achieved by orienting the plane of division and the new cross-wall. The latter has much in common with the morphogenetic cell migrations involved in body and structure formation in animals (Moore and Novak Frazer, 2002). In this sense, fruiting body formation in mushrooms is seen as employing morphogenetic processes that have affinities with both of the other major eukaryote kingdoms. In Fig. 2, a schematic representation of fruiting body development in *Coprinopsis cinerea* is presented. The first steps of fruiting body formation (reviewed in Kües, 2000) are characterized by intense hyphal branching. This process takes place in the dark and gives rise to the primary hyphal knot. After the primary hyphal knot is formed, a light pulse is required for the further development into a secondary hyphal knot (also known as initial) in which cellular differentiation begins. After initial formation, other light pulses induce correct progression into the next developmental stages: primordia formation and maturation. In the primordia (1-10 mm sized aggregated hyphal structures) the typical mushroom shape is recognizable and the different hyphal tissues that will constitute the mature fruiting body are present. Maturation of the fruiting body (stipe elongation, cap expansion) proceeds with karyogamy and meiosis within the specialized basidia. Four basidiospores per basidium are finally produced that will be dispersed within a brown-yellow liquid when the cap auto-lyses at maturation. Usually, four to five days are necessary to complete maturation from the first sign of fruiting on the dikaryon (Kües, 2000). The process described corresponds to the normal development of the fruiting body as observed in nature: the events, although allowing certain flexibility, are precisely scheduled by the natural alternation of light and dark periods. Deviations from this developmental pattern are possible under artificial conditions. For instance, if primary hyphal knots are maintained in constant darkness, secondary hyphal knot formation is repressed. Instead, sclerotia will develop from the primary hyphal knots (Waters et al., 1975; Bouliananne et al., 2000; Kües, 2000). Under
constant light, in contrast, primary hyphal knot and fruiting body formation are repressed (Fig. 2) (Kües et al., 1998). There is no universal set of conditions that leads to fructification in all fungi (Kües and Liu, 2000), however, signals important for sexual reproduction in fungi fall into two categories. The first embrace signals of environmental nature like temperature, light, humidity, but also nitrogen and carbon concentration in the substrate. The second comprises signals of genetic origin, like mating type genes (Gow and Gadd, 1995; Kües and Liu, 2000). Accordingly, fruiting body formation in C. cinerea is controlled by both genetic and non-genetic signals. Among genetic signals, the mating type genes are the master regulators, whereas light and nutrient signals exert the most important non-genetic influence (Kües, 2000). Hereafter, the main effects of these three regulators of fruiting body formation in C. cinerea are briefly discussed.

Fig. 2 (next page): Fruiting body development in C. cinerea (modified from Boulianne et al., 2000).
Mycelium → Light → Carbon & nitrogen → A mating type genes → Primary hyphal knot → Light → Secondary hyphal knot → Light → Primordium → Light → Karyogamy → Light → Meiosis → Mature fruiting body

Sclerotium: 4.5 cm

Meiosis:

Primordium: 2.4 cm

Karyogamy: 0.8 cm

Meiosis: 1 cm

Light: 0.3 cm

Light: 0.1 cm
1.2.1. Mating type genes

In *C. cinerea*, genetic control of sexual development is exerted by the mating type genes as the master regulators. The function of the mating type genes in the fungi is to impose barriers on self-mating and thereby maintain genetic variability within the population (Casselton and Olesnicky, 1998). Around 90% of the homobasidiomycetes are heterothallic, i.e. they need a mating partner to undergo sexual reproduction. The rest is homothallic (self-compatible). Around 61% of heterothallic homobasidiomycetes have two unlinked mating type loci, designed $A$ and $B$, both controlling the mating events. *C. cinerea* belongs to this group. The rest has a single mating type locus (Kamada, 2002). The *C. cinerea* $A$ mating type genes code for homeodomain transcription factors (Kües and Casselton, 1992; Kües et al., 1992; Kües and Casselton, 1993; Kües et al., 1994; Kües et al., 1994) and the $B$ genes for pheromone and pheromone receptors (O'Shea et al., 1998; Olesnicky et al., 1999; Halsall et al., 2000). The products of the mating type gene exert different functions on dikaryon formation and maintenance. $A$ gene transcription is not pheromone dependent and is constitutive both in monokaryons and dikaryons. Moreover, no evidence is present for increased transcription levels being required to maintain the dikaryon (Richardson et al., 1993). After successful anastomosis, the $B$ gene function is needed to control nuclear migration of haploid nuclei invading monokaryotic hyphae of different mating type and, later in dikaryon development, for clamp cell fusion. After mating, the $A$ mating type gene products are responsible for pairing of the two genetically distinct nuclei in hyphal segments, synchronous nuclear division and clamp cell formation in the newly formed dikaryon (Swiezynski and Day, 1960a, 1960b). Regarding fruiting body formation, a pivotal role is played by the $A$ mating type genes, whose products, under normal environmental conditions, are essential for primary and secondary hyphal knot formation (Tymon et al., 1992; Kües et al., 1998). The $B$ mating type gene products, though having an enhancing effect on the action of the $A$ mating factors in the initial stages of fruiting body development, are alone not able to induce fruiting when activated by transformation in a monokaryon. However, once a primordium is fully established, the $B$ genes are needed for induction of fruiting body maturation at the stage of karyogamy (Kües et al., 2002).
1.2.2. Nutrients

It is assumed that nutrient depletion in the substrate is the triggering signal for fruiting body formation on dikaryons (Rao and Niederpruem, 1969; Stewart and Moore, 1974). Indeed, fruiting is repressed by high concentrations of carbon and nitrogen sources in the substrate (Madelin, 1956; Uno and Ishikawa, 1974; Boulianne et al., 2000; Bottoli, 2001; Fig. 2). More generally, nutrients are known in fungi to play a pivotal role in controlling morphological as well as metabolic transitions (see chapter 2.3.2.).

1.2.3. Light

Light is probably the most important physical signal upon which developmental choices are made when other requirements are met (Kües, 2000). On fruiting body formation, light (i.e. blue light) seems to have both positive and negative effects, depending on the developmental stage and the time and energy with which light is given. Primary hyphal knot formation (the first step in fruiting body formation) is carried out in the dark and is repressed by light (Fig. 2 and Kües et al., 1998). Subsequently, a light pulse is required to induce secondary hyphal knot (initial) formation once the primary hyphal knot is formed, whilst incubation under constant light blocks also secondary hyphal knot formation. Primordia will then arise if secondary hyphal knots are kept under a light-dark regime mimicking the natural day/night rhythm. Otherwise, by incubation of secondary hyphal knots in constant darkness, dark stipes (i.e. fruiting body-like structures with an extended stipe and without a properly developed cap) will be formed (Kües, 2000). Light is further needed to promote primordial maturation and karyogamy, and another dark period is required for completion of meiosis (Fig. 2, Lu, 1974,2000; Kües, 2000b).

2. Control of gene transcription in fungi during developmental processes

The regulators listed above (i.e. products of mating type genes, light and nutrients) exert their control on sexual development by directly or indirectly activating target genes. The resulting dynamic network of gene activation/repression will finally produce the desired phenotype (i.e. fruiting body development). Regulation of gene
expression in fungi can be achieved at different levels, of which transcription is one of the most important (Moore and Novak Frazer, 2002). How do mating type genes, light and nutrients regulate gene transcription as a mean to guide fruiting body development? Despite the high relevance of fungi to understand basic questions in developmental biology, the regulation of dikaryon formation and fruiting body development in basidiomycetes at the gene level is still poorly understood. In chapter 2.1., a short overview on transcriptional regulation in fungi is presented. Afterwards, a brief list of fungal transcription factors with their DNA target sequence (chapter 2.2.) and their functioning in a natural context (chapter 2.3.) is given. Where possible, emphasis is put on transcriptional regulation exerted by mating type genes, nutrients and light signals.

2.1. Transcriptional regulation of fungal genes

Although one of the best studied eukaryotic organisms, Saccharomyces cerevisiae, is an ascomycete, there is still little knowledge on gene regulation in filamentous (ascomycetous and basidiomycetous) fungi. However, the mechanisms involved appear to be conserved in all eukaryotes (Kornberg, 1999, 2001). The presentation in the following is not intended to provide a thorough understanding of the general architecture of eukaryotic promoters but rather to indicate transcriptional regulatory features found usually in fungal genes. There is abundant evidence that transcriptional control is a major level of gene regulation in eukaryotes (Moore and Novak Frazer, 2002) and this applies also in filamentous fungi (Gurr et al., 1987). Regarding DNA regulatory sequences (known as cis-acting elements), eukaryotic promoters contain roughly three types of elements: firstly, sequences found in most genes involved in basal transcription; secondly, the “response elements” which are found only in genes transcribed in a particular tissue or in response to a specific signal and are likely to produce a specific pattern of expression; finally, elements, like enhancers, able to influence the level of transcription in an unspecific way (Latchman, 1998).

The TATA box (consensus TATAA/TAA/T) is a classical eukaryotic element of the first type (Berk, 2000). Its function is to direct the general transcription machinery to the right place on the promoter. In some genes there is a unique transcriptional start site (+1), whilst in others there is a tendency towards multiple initiation. The TATA
Chapter I

box is found in several but by no means in all filamentous fungal promoters (Gurr et al. 1997). In S. cerevisiae, the TATA box is commonly located 60 to 120 base pairs upstream from the transcription start point compared to the 30 base pairs found in mammals (Moore and Novak Frazer, 2002). In promoters, containing a TATA box or not, the minimal activity of the promoter is dramatically increased by other basal elements located upstream in the promoter sequence and found in a wide variety of genes. One of these is the CCAAT box, found in many genes irrespective of their specific regulation patterns (Latchman, 1998). In genes of filamentous fungi, putative CCAAT boxes, where present, are usually observed between 60 and 120 base pairs upstream of the transcription initiation site (Moore and Novak Frazer, 2002). Another typical ubiquitous eukaryotic motif is a GC-rich sequence known as the Sp1 box (consensus KRGGCKRRK, http://www.gene-regulation.de/ and Hapgood et al., 2001), located about 100 base pairs further upstream to the CCAAT box (Moore and Novak Frazer, 2002). In addition to these classical motifs, other minor sites can be found to influence basal transcriptional activity in different eukaryotic genes. Motifs resembling TATAAA, in that they possess AT-rich regions, are often found upstream at around -275 bp to the +1 site (Gurr et al., 1987 and references therein). In yeast, the level of transcription may be markedly affected by the presence of a pyrimidine-rich sequence preceding the transcriptional initiation site and a number of filamentous fungal transcription initiation sites appear in or immediately downstream from such pyrimidine-rich sequences. These CT-rich sequences are particularly noticeable in genes lacking TATAAA and CCAAT motifs (Gurr et al., 1987).

Transcription is a rather complex process in which the action of many different proteins have to be precisely co-ordinated in time and space. Constitutive – like factors belonging to the general transcription machinery – and transient (cell- and developmental-specific) proteins have to interact in order to produce the correct gene transcription at the right moment. The first steps of transcription include the binding of the general transcription factors (GTF) to the core promoter elements (i.e. TATAAA and CCAAT boxes). This provides low levels of accurate transcription, called basal transcription, using a variety of additional transcription factors (TFIID, TFIIA, B, E, F and H) and RNA polymerase II itself. Gene, cell or stage-specific gene transcription is achieved by transcription factors binding at “response elements”. A more comprehensive discussion of these fungal transcription factors, their target sequences and mode of action are presented in chapters 2.2. and 2.3.. Proteins
binding at enhancers constitute another class of transcription factors, acting by increasing the promoter activity by several orders of magnitude. Enhancers can be found many thousands of nucleotides away from the promoter, upstream or downstream or within a transcription unit and function in either orientation relative to the start site of transcription (Latchman, 1998). Proteins binding at enhancers, however, are unable to drive transcription by themselves and need another class of transcription factors called co-activators. These do not have site-specific DNA-binding activity, but act as intermediaries for the action of transcription activators on the basal transcription machinery. Finally, as interface between activators and polymerase II, an additional co-activator activity is required. This is provided by a factor called mediator which in yeast consists of a 20-subunit protein complex (Moore and Novak Frazer, 2002).

Further control in gene transcription is provided at the level of chromatin. Chromosomes in eukaryotes consist of about one-third genomic DNA, one-third histone proteins and one-third non-histone proteins. Function of histone proteins is the packaging of the DNA in chromatin which makes up the chromosomes. In eukaryotes, the normal structure of chromatin is entirely sufficient to maintain transcription at the minimal, basal level (Moore and Novak Frazer, 2002). Heterochromatin is highly condensed chromatin that silences gene transcription completely. Even basal transcription is lacking, as long as the silencing structure is in place. Chromosomal regions silenced by heterochromatin are often associated with DNA methylation. Even though the extent of DNA methylation is consistently low in fungi compared with mammals and plants, it seems to play an important role in gene inactivation in at least some filamentous fungi (Moore and Novak Frazer, 2002), including *C. cinerea* (Freedman and Pukkila, 1993).

The majority of genes involved in transcriptional control identified so far in filamentous fungi appear to produce proteins with a positive mode of action (activator proteins). Most activators seem to require dimerization to be active (Moore and Novak Frazer, 2002). Repression of transcription can act at different levels. Repressors may interact with general transcription factors, thereby affecting assembly of the transcription complex. They may also induce chromatin packaging, and thereby gene silencing, by positively influencing histone assembly. Some repressors bind to the same DNA sites as activator proteins or to the activators themselves preventing in this way their activating function. It has become evident that
many transcription factors can act as either activator or repressor, depending on the
gene being regulated and the cell type in which it is expressed. Covalent
modifications of already synthesized transcription factors (e.g. phosphorylation or
acetylation) is a widely used way by which specific modulation of gene transcription
can be achieved in a comparable short time (e.g. Tootle et al., 2003).

2.2. Trans-acting proteins and their DNA targets in fungal
genes

Trans-acting factors are proteins influencing the transcriptional activity of a gene
by binding at cis-acting elements in a gene’s promoter. Cis-acting elements are DNA
motifs localized in the vicinity of a gene’s coding sequence that are required for its
transcription. It is the dynamic interaction between trans- and cis-acting elements
which provides the molecular basis through which transcriptional regulation of a gene
is achieved. Hereafter, a short overview of eukaryotic trans- and cis-acting elements
is provided. The more general knowledge is gained from studies involving also non-
fungal organisms. Whenever possible, however, information from fungal systems is
included. Here, emphasis is put on fungal proteins somehow involved in transmitting
mating type genes-, light- and nutrient-derived signals and thus binding at cis-acting
sites of the “response element” type. According to the DNA binding domain they
contain, trans-acting proteins (and their binding motifs) are first listed individually and
then, by means of classical examples, considered in their natural context (chapter
2.3.).

2.2.1. Transcription factors containing basic domains

Leucine zipper (bZip) domain

The leucine zipper domain consists of a dimer of two homologous polypeptides.
Every polypeptide contains an α-helix in which leucine residues occur every two turns
on the same side of the helix. The leucine residues extending from one polypeptide
interdigitate with those of the analogous helix of the second polypeptide, forming a
motif known as the leucine zipper, which would result in the dimerization of the factor
(Latchman, 1998). The DNA itself is bound by two regions (one per polypeptide)
adjacent to the α-helix which, in the animal transcription factors C/EBP (see below),
Fos and Jun, are rich in basic amino acids. The leucine zipper is believed therefore
to serve an indirect structural role in DNA binding, facilitating dimerization and thus correct positioning of the two basic DNA-binding domains (Latchman, 1998). The leucine zipper can also be found as a dimerization motif in proteins which use DNA-binding motifs other than the basic region (Latchman, 1998). An important group of bZip transcription proteins is represented by the CRE-binding factors. CRE stands for cAMP responsive element and, as the name says, it mediates transcriptional regulation on stimulation of the adenylate cyclase signaling pathway (Daniel et al., 1998; De Cesare and Sassone-Corsi, 2000). Classical eukaryotic trans-acting factors binding to CRE motifs were discovered in animal systems and include:

- CREB (cAMP response element binding protein, Mayr and Montminy, 2001; Conkright et al., 2003);
- CREM (cAMP response element modulatory protein, De Cesare and Sassone-Corsi, 2000);
- ATF-1 (activating transcription factor 1, Rehfuss et al., 1991).

Homo- and heterodimers composed of these factors bind best to the classical CRE motif defined by the eight-base pair palindromic sequence TGACGTCA (Daniel et al., 1998; Conkright et al., 2003), although variations of this motif (e.g. the half palindrome CGTCA) are also effective (Hyman et al., 1989; Kageyama et al., 1991; Conkright et al., 2003). As stated above, phosphorylation is a largely used way to regulate the activity of transcription factors. Different mechanisms have been described by which various signaling routes converge on CREB and CREM proteins and control their function by modulating their phosphorylation state (Mayr and Montminy, 2001). Phosphorylation does not seem to be required for CREB to bind DNA. It rather works as a molecular switch in that it dictates the capability of these factors to interact with ubiquitous co-activators, such as the CBP (CREB binding protein), that mediate their contacts with the basal transcriptional machinery (Mayr and Montminy, 2001). CBP, known to play a key role as a co-activator for transcription factors activated by other signaling pathways, has been shown to have histone acetyltransferase activity (Latchman, 1998). Hence, its binding to CREB may also result in a change of the chromatin structure allowing transcription to occur (Latchman, 1998; De Cesare and Sassone-Corsi, 2000). Finally, data point to the existence of tissue-specific co-activators able to stimulate the transcriptional activity of CREB and CREM in a phosphorylation-independent manner (Fimia et al., 1999). CRE-binding proteins can act as activators as well as repressors (De Cesare and
Sassone-Corsi, 2000). CREB is able to mediate both basal and kinase-inducible transcription in a variety of cells and tissues (Quinn, 2002).

bZip proteins binding to CRE motifs do also exist in fungi. In *S. cerevisiae*, the family of ATF/CREB transcriptional regulators consists of a repressor, Acr1 (Sko1), and two activators, Aca1 and Aca2, all binding to the CRE consensus site TGACGTCA or variants thereof (Vincent and Struhl, 1992; Garcia-Gimeno and Struhl, 2000). Like other ATF/CREB proteins, they do not bind the AP-1 consensus (TGACTCA), in contrast to another important yeast transcription factor, GCN4 (see below), which prefers this site over the ATF/CREB site (Sellers et al., 1990). Accordingly, GCN4 does not activate transcription through ATF/CREB sites *in vivo*, even though it binds these sites *in vitro* (Garcia-Gimeno and Struhl, 2000).

ATF/CREB activators in *Schizosaccharomyces pombe* and mammalian cells play an important role in mediating the response to activation of PKA and to a wide variety of environmental stresses. The remarkably conserved molecular basis for the stress response in these eukaryotes implies that these ATF/CREB proteins are directly phosphorylated by stress-responsive MAP kinases, whereupon they activate target genes containing ATF/CREB sites. In contrast, the ATF/CREB family in *S. cerevisiae* does not appear to mediate the responses to PKA or to stress (Garcia-Gimeno and Struhl, 2000). Nevertheless, Aca2 is important in *S. cerevisiae* for a variety of biological functions including growth on non-optimal carbon sources. Acr1 represses transcription both by directly competing with Aca1 and Aca2 for target sites and by recruiting the Cyc8-Tup1 complex, a co-repressor complex required for repression by a variety of pathway-specific DNA-binding proteins such as α2 and MIG1 (see chapter 2.3.2.), (Garcia-Gimeno and Struhl, 2000). In *S. pombe*, the *pcr1* gene, encoding a bZip protein with CRE-binding activity, functions in an early stage of sexual development in association with the cAMP cascade. Disruption of *pcr1* reduced expression of *fbp1*, a glucose-repressible gene negatively regulated by PKA (see below, Watanabe and Yamamoto, 1996). In *S. pombe*, the *cdc10* gene product constitutes a transcription factor complex with other proteins (e.g. Res1/Sct1) and activates the transcription of different genes in correspondence of the “START” point, the moment in the G1 phase of the cell cycle at which the cells assess the environment and enter the pathway best suited for survival (Nojima et al., 1994). In this context, Nojima et al. (1994) described a *S. cerevisiae*’s multicopy suppressor of the *cdc10* mutant gene called *hac1*. *hac1* encodes a bZIP protein of 240 AA with
close homology to the mammalian ATF/CREB transcription factor (HAC1 is an acronym of "homologous to ATF/CREB") and gel-retardation assays showed that it binds specifically to a CRE site (Nojima et al., 1994). The caffeine sensitivity of the HAC1 disruptant suggests that it may be involved in a signal transduction pathway using cAMP as a second messenger and the existence of a potential phosphorylation site by cAMP-dependent protein kinase in the bZIP region suggests indeed that Hac1 transcriptional activity has to be located in a cAMP dependent pathway (Nojima et al., 1994).

Additional relevant fungal members of the bZip family are SWI6 and SWI4. These proteins form the SBF (SCB binding factor) complex, which, together with many other factors, is involved in chromatin remodeling in S. cerevisiae (see below, Sidorova and Breeden, 1993; Aalfs and Kingston, 2000). The binding site of SBF is called SCB, for SWI4,6-dependent cell cycle box, and has the consensus sequence CACGAAA (Harrington and Andrews, 1996).

**Basic helix-loop-helix (bHLH) domain**

The basic DNA-binding domain of the just described leucine zipper factors has also been identified in a number of other transcription factors which do not contain a leucine zipper. In these cases, the basic DNA-binding domain is juxtaposed to a region that can form a helix-loop-helix motif, which, like the leucine zipper, allows dimerization with other molecules thereby facilitating DNA binding (Latchman, 1998). It seems as if hetero- and homodimers formed by HLH transcription factors would bind to a conserved core sequence (CANNTG) called the E box (Murre et al., 1989; Peleg and Metzenberg, 1994). The basic DNA-binding domain can be additionally involved in protein-protein interaction (Latchman, 1998). Two fungal examples of helix-loop-helix containing transcription factors are PHO4 (S. cerevisiae) and NUC-1 (Neurospora crassa). The product of the PHO4 gene binds to a sequence containing the core bHLH-binding site CACGTG, which is found in the promoter of several genes involved in phosphate utilization in S. cerevisiae. One of these genes, PHO5, contains two PHO4 binding sites and binding of PHO4 to these sites is required for the around 1000-fold induction of the PHO5 gene expression in response to phosphate starvation (Robinson and Lopes, 2000). The transcription factor NUC-1 is a positive regulatory protein which directly controls the transcription of several unlinked target genes involved in phosphorus uptake in N. crassa. NUC-1, containing
Chapter I

the bHLH domain, binds to DNA sequences harboring the core sequence CACGTG (Peleg and Metzenberg, 1994).

2.2.2. Transcription factors containing zinc-coordinating DNA-binding domains

Transcription factors with one or usually more domains containing a zinc ion coordinated by different amino acids, build up another important group of trans-acting proteins. DNA-binding is mediated by a loop (from here the name “finger”) formed in the polypeptide when the zinc ion is coordinated. Usually, many adjacent zinc fingers are required for DNA-binding, explaining why zinc finger-containing transcription factors always have multiple copies of the zinc finger motif. According to the amino acids coordinating the zinc ion, different DNA-binding domains exist (Latchman, 1998).

Cys2His2 zinc finger domain

In this first group, two cysteines and two histidines are the amino acids coordinating the zinc ion, hence the name “Cys2His2 zinc finger”. Two anti-parallel β-sheets with an adjacent α-helix are the secondary structures found in correspondence of the finger. The first identified transcription factor containing the Cys2His2 zinc finger was TFIIIA, which contains 9 fingers and regulates the transcription of the 5S ribosomal RNA genes by RNA polymerase III (Latchman, 1998). Subsequently, similar Cys2His2-containing zinc finger motifs were identified in a number of RNA polymerase II transcription factors, such as Sp1, which contains three contiguous zinc fingers (Latchman, 1998; Hapgood et al., 2001). DNA-binding specificity is determined by amino acids at the N-terminus of the α-helix, most notably the amino acids immediately preceding the first histidine residue. It seems that, in each successive finger, these amino acids interact with three bases of DNA within the recognition sequence (Latchman, 1998). In principle, according to whether there are two, three or more zinc fingers, the DNA target sequence should be composed of two, three or more nucleotide triplets.

A typical group of fungal transcription factors of the Cys2His2 zinc finger class are those involved in carbon catabolite repression (see below, chapter 2.3.2.B). Homologous genes in this group are the Aspergillus nidulans CREA, the S.
*cerevisiae* MIG1 and the CRE1 of *Trichoderma reesei*. The products of these genes bind all to the consensus sequence SYGGGRG (Lundin et al., 1994; Takashima et al., 1996) which, however, seems to allow some variability. Thus, Espeso and Penalva (1994) found that some of the consensus sequences in the *A. nidulans* ipnA gene were not recognized by CREA, while some non-consensus sequences were. Similarly, *Trichoderma reesei* CRE1 was shown to bind both consensus and consensus-like sites in the cellobiohydrolase 1 (*cbh1*) promoter (Takashima et al., 1996). In addition to the core SYGGGRG sequence, in a saturation mutagenesis of a natural MIG1 site, Lundin and coworkers (1994) found that an AT-rich region 5' to the GC box was also important for MIG1 binding. Although the AT-rich region is present in all natural MIG1 sites, no single base within the AT box is essential for binding. It is possible that the role of flanking DNA is a unique development among fungal zinc finger proteins that have only two fingers and therefore may require additional DNA contacts for high-affinity binding. The mammalian GC box-binding proteins frequently bind to very GC-rich sequences (GC islands) and are therefore unlikely to require flanking AT-rich DNA (Lundin et al., 1994). As already stated in chapter 2.1., in higher eukaryotes the Sp1 box constitutes a well known GC-rich motif. The discovery of functional Sp1 motifs in fungal gene promoters (Zhang et al., 1999) extends now their importance as transcriptional control DNA elements to the fungal kingdom.

Other examples of fungal Cys2His2 zinc finger transcription factors are the *S. cerevisiae* ADR-1 (activator of peroxidase genes), the *A. nidulans* PACC (involved in pH regulated transcription) and SW5, which, together with the above-discussed SW6 and SW4 (see above, chapter 2.2.1.) are in *S. cerevisiae* involved in chromatin remodeling (Aalfs and Kingston, 2000). A well studied gene on which SW5 operates is the *HO* gene encoding for an endonuclease involved in mating type switching. Transcription of the *HO* gene is controlled by different SWI (mating type switching) genes. SW5 and SBF (bZip type transcription factor, see above, chapter 2.2.1.) bind to distinct sites upstream of the *HO* open reading frame, and both are required for the normal expression of the *HO* gene. SW5 binding to the *HO* promoter is believed to be the triggering event leading to *HO* promoter remodeling and transcription (Aalfs and Kingston, 2000). The consensus sequence for SW5-binding corresponds to the sequence USNNYNKGCTGR (Tebb et al., 1993; Kovacech et al., 1996).
In *S. cerevisiae*, another important Cys2His2 zinc finger protein involved in development is RME1 (regulator of meiosis 1, see chapter 2.3.1.). RME1, which contains three zinc fingers, binds as a dimer to a direct repeat of the sequence TCAARARRWCC (Covitz and Mitchell, 1993).

**Cys6 zinc cluster domain**

Another zinc finger motif is the Cys6 zinc cluster, in which the zinc ion is coordinated by four cysteine residues rather than the two-cysteine-two-histidine structure discussed above (Latchman, 1998). Differently from the modular structure of the Cys2His2 finger, where each finger constitutes an independent structural element whose configuration is unaffected by adjacent fingers, the fingers of multicysteine zinc finger factors form one single structural motif (Latchman, 1998). Nuclear receptors (e.g. steroid receptor in animals) constitute an important family of multicysteine zinc finger factors. In these proteins, DNA-binding specificity is determined by adjacent amino acids in consecutive fingers as well as in the linker sequence between fingers (Latchman, 1998). Most binding sites of Cys6 zinc finger containing proteins consist of terminal trinucleotides separated by an internal spacer sequence. The terminal trinucleotides can be in an inverted orientation as well as direct repeats (Todd et al., 1998). Members of the nuclear receptor family can bind both as homo- and heterodimers, each dimer interacting with one repeat. Depending on the amino acid composition in the fingers, the optimal spacing between DNA repeats can differ. The same amino acids are additionally responsible for interaction between the polypeptides composing the dimer complex (Latchman, 1998). Examples of fungal Cys6 zinc finger transcription factors are the *S. cerevisiae* GAL4 (involved in galactose metabolism, see below, chapter 2.3.2.B) and HAP1, the *N. crassa* NIT4 (involved in nitrogen metabolism) or the *A. nidulans* FACB (involved in acetate utilization) and AMDR (involved in omega-amino acid metabolism). GAL4 binds to DNA containing the sequence CGGN_{11}CCG. In contrast, FACB binding sites seem to be of two classes: class A has the consensus sequence TCSN_{8-10}SGA and class B the consensus sequence GCAGNTN_{2}CCN_{1-2}GGC (Todd et al., 1998).
Cys4 zinc finger domain and the GATA factors

Transcription factors containing the Cys2Cys2 zinc finger domain and binding to the DNA consensus sequence HGATAR are generally called "GATA" factors (Scanzocchio, 2000). GATA factors are present in fungi, metazoans and plants. In invertebrates, proteins containing two as well as one Cys2Cys2 domain can be found. In contrast, plant and fungal GATA factors have typically a single Cys2Cys2 binding domain (Scanzocchio, 2000). Fungal GATA factors can act both as activators and repressors (Marzluf, 1997) and are involved in different cellular processes, ranging from nitrogen source utilization (see below, chapter 2.3.2.C), mating-type switching and light signaling (Marzluf, 1997). In Ustilago maydis, URBS1, which contains exceptionally two zinc fingers, acts in iron regulation (Voisard et al., 1993). WC-1 and WC-2 are involved in light-dependent transcription in N. crassa (Ballario and Macino, 1997, see chapter 2.3.3.) and the S. cerevisiae DAL80 is involved in nitrogen catabolite repression (see below, chapter 2.3.2.C). GATA factors in higher metazoan species have typically two zinc fingers. Here, the carboxy-terminal finger, responsible for sequence-specific DNA binding, shows a high degree of homology to the single zinc finger of the A. nidulans AREA, N. crassa NIT2 and S. cerevisiae GLN3 (Marzluf, 1997), all transcription factors involved in nitrogen regulation in these fungi. As found already for the bHLH domain (see above, chapter 2.2.1.), the DNA-binding domain in both the mammalian and the fungal GATA factors is also able to interact with other proteins, including other zinc fingers. The N-terminal zinc finger of the mammalian GATA-1 protein, for instance, involved in erythropoietic cell lineage development and function, appears to promote dimerization as well as functional interactions with Sp1 (Marzluf, 1997). In fungi, the domain of AREA/NIT2 can interact with an identical domain of the NMR (negative acting protein) and the specific regulator NRA/NIT4 (Scanzocchio, 2000). Similar as found for other cis-acting sequences, the presence of a GATA sequence in a yeast gene’s promoter does not necessarily imply GATA factor regulation of its expression. For example, expression of DAL5 and DAL7 is highly regulated by GATA factors, and their promoters possess multiple GATA sequences. However, only two of the GATA elements found on the DAL5 promoter function at full capacity; six other cannot be shown to function under any condition assayed so far (Bysani et al., 1991). Similarly, in DAL7 only three of the six GATAs present appear to function (van Vuuren et al., 1991). Likewise, on the
promoter of the *A. nidulans* lysF gene, encoding homoaconitase, two GATA sites were shown to positively affect the expression of a lysF reporter gene, but were not bound by AREA, the GLN3 homologous nitrogen master regulator in *A. nidulans* (Weidner et al., 2001). These observations suggest that GATA elements might consist of more than the core GATA motif (Cooper, 2002).

### 2.2.3. Transcription factors containing the helix-turn-helix domain

#### The homeodomain

In homeodomain-containing transcription factors, the DNA interaction is mediated by the homeobox, a 60 amino acids conserved sequence found in all homeodomain-containing proteins (Kornberg, 1993; Latchman, 1998). Main constituents of the homeodomain are three α-helices, of which the carboxy-terminal-most helix (recognition helix) inserts along the major groove making thus sequence-specific contacts with the DNA bases (Latchman, 1998; Khorasanizadeh and Rastinejad, 1999). As revealed from studies with the Bicoid and Antennapedia proteins, which control developmental processes in *Drosophila*, it seems that the DNA target sequence is primarily determined by the ninth amino acid in the recognition helix (Kornberg, 1993; Latchman, 1998). Additionally, a short N-terminal arm contributes significantly to the DNA-binding ability of the homeodomain, probably by contacting the ATTA motif common to the DNA-binding sites of several homeodomain proteins (Kornberg, 1993; Latchman, 1998).

Important fungal transcription factors of the homedomain type are the mating type proteins of which the best studied representatives are MATα1 and MATα2 of *S. cerevisiae* (Johnson and Herskowitz, 1985). Importantly, crystal structures of MATα2 revealed how a homeodomain protein can collaborate with either the related family member MATα1, or the MADS box protein MCM1 (see below, chapter 2.2.4.), to select the appropriate DNA target site for the proteins' function in the regulation of the yeast cell type (see chapter 2.3.1., Khorasanizadeh and Rastinejad, 1999). The MATα1 and MATα2 corresponding transcription factors in the basidiomycetes are the b mating type proteins in *U. maydis* and the A mating type proteins in *C. cinerea* and *S. commune* (Casselton and Olesnicky, 1998). By characteristic amino acid differences, their homeodomains fall into two classes which have been termed HD1 and HD2 (Kües and Casselton, 1993). To activate sexual development in *C. cinerea*
and the other basidiomycetes, dimerization of the homeodomain mating type transcription factors with subsequent DNA target site recognition is required (Banham et al., 1995; Casselton and Olesnicky, 1998). HD1 and HD2 mating type factors might however exert other functions by binding, alone or in complex with other proteins, to additional DNA sequence motifs, as known from MATα2 in yeast (Kües and Casselton, 1992; Casselton and Olesnicky, 1998). While the analogous HD1-HD2 heterodimer in S. cerevisiae (the MATα1-MATα2 heterodimer) act as repressor (Herskowitz, 1998; see chapter 2.3.1.), in the basidiomycetes it is not clear yet whether the proteins have activator, repressor or both functions (Casselton and Olesnicky, 1998; Kamada, 2002).

Another homeodomain transcription factor of yeast is HCM1. HCM1, another target of the already described SBF complex (see more on top), is believed to have a role in cell cycle progression (Horak et al., 2002). The consensus binding sequence of HCM1 is WAAYAAACAAW (Zhu and Davis, 1998).

2.2.4. β-scaffold transcription factors with minor groove contacts

The MADS box domain

MADS box containing transcription factors are found from yeast to humans. The name refers to four of the originally identified members of this transcription family: MCM1, AG, DEFA and SRF. In recent years, MADS box transcription factors earned increased attention because of the fundamental role played during flower development in plants (Ng and Yanofsky, 2001; Irish, 2003). The MADS box is a contiguous conserved sequence of 56 amino acids, of which 9 are identical in all family members described so far. Several members have been shown to form dimers. The N-terminal half is the major determinant of DNA-binding specificity whilst the C-terminal half is necessary for dimerization. Even though each MADS box containing protein apparently possesses a distinct binding specificity, the majority bind similar sites based on the consensus sequence CCW6GG. Moreover, the specific recruitment, among several MADS box proteins, of other transcription factors into multi-component regulatory complexes, appears to be a common theme within this family and play a pivotal role in the regulation of target genes (Shore and Sharrocks, 1995). To the MADS box class of transcription factors belongs the already mentioned MCM1 of S. cerevisiae (Ammerer, 1990), which, by interacting with either
MATα1 or MATα2, is central to the transcriptional control of cell-type specific genes including genes of the mating pheromone response (see below). Interestingly, the requirements within the MCM1 MADS box domain for interaction with MATα1 and MATα2 are different and do not affect each other (Mead et al., 2002).

Another important group of β-scaffold factors are the TATA box-binding proteins, common to all eukaryotes (Berk, 2000).

The HMG box domain

HMG box factors are a group of transcription factors containing the HMG box (for “High Mobility Group”) DNA-binding domain. It consists of three α-helical domains and an irregularly structured C-terminus (van Beest et al., 2000). Two types of HMG box proteins can be distinguished. The first type usually contains multiple HMG boxes that bind DNA in a structure-dependent but sequence-independent fashion. Examples for this type are the human eukaryotic upstream binding factor (UBF, Jantzen et al., 1990) and the rat HMG1 (Weir et al., 1993). The second type of protein contains a single HMG box that can bind to DNA in a sequence-specific way. In these proteins, the HMG domain can recognize DNA sequences of up to 12 base pairs with high specificity but the length and composition of the bound DNA motif can diverge for different HMG domains (van Beest et al., 2000). The core recognition DNA motif of HMG boxes corresponds to the sequence AACAAAGG or variants thereof. Additionally, sequence-specific contacts outside the core DNA motif are necessary for discrimination (van Beest et al., 2000). Fig. 3 reports the consensus sequence for different HMG box proteins.
<table>
<thead>
<tr>
<th></th>
<th>Sry/Sox-5</th>
<th>Ste11</th>
<th>Rox1</th>
<th>MatMc</th>
<th>Sox-4</th>
<th>Tcf-1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7/9 29/36 34/40 40/40 40/40 39/40 40/40 39/40 40/40 34/40</td>
<td>26/33 26/33 29/33 32/33 33/33 33/33 33/33 33/33 33/33 33/33 31/33</td>
<td>23/45 22/45 40/45 41/45 45/45 45/45 45/45 38/45 45/45 40/45</td>
<td>29/33 25/33 24/33 25/33 30/33 24/33 30/33 30/33 26/33 33/33 27/33 26/33</td>
<td>42/42 41/42 42/42 - - - - 40/42 40/42 24/42 24/42</td>
<td></td>
</tr>
</tbody>
</table>
Important fungal HMG box proteins are encoded by several genes involved in fungal mating type determination, such as \textit{prf1} in \textit{U. maydis}, \textit{mat} genes from mating type loci in filamentous ascomycetes or the \textit{S. pombe} genes \textit{mat1-Mc} and \textit{ste11} (van Beest et al., 2000). \textit{S. pombe} has one mating type with two different alleles, \textit{mat1-Mc} and \textit{mat1-Pc} (see chapter 2.2. for details), of which the activation of M or P cell-specific genes, respectively, depends. During mating in \textit{S. pombe}, Ste11 controls the expression of M-specific genes as well as genes that are expressed in both cell types by binding to a specific site on the genes' promoter. By comparing the Ste11-binding sites in these two type of genes, Kjaerulff et al. (1997) propose that M-specific gene activation results from Mat1-Mc-dependent Ste11 binding to a weak TR-box (AACAAAGAGA) whereas ubiquitously expressed genes harbor a strong TR-box (AACAAAAGAAA), to which Ste11 can bind on its own. This situation is reminiscent to MCM1 action in yeast (see below, chapter 2.3.1.).

Heteromeric CCAAT transcription factors

The CCAAT motif is one of the most common elements in eukaryotic promoters (Latchman, 1998; Mantovani, 1998). In general, CCAAT boxes lie about 50 to 200 bp upstream to the transcriptional start point and can be present in either orientation (Brakhage et al., 1999). Two families of CCAAT box binding proteins exist. The first is represented by C/EBP (CCAAT/Enhancer binding protein) as its prototype, whose activity is probably more restricted to specific tissues or developmental stages (Latchman, 1998). The second, more important, whose members act as heteromeric polypeptides (like NF-Y, CBF and HAP), is likely to play a role in constitutive expression. In agreement with this distinction, the vast majority of the CCAAT boxes have invariably been shown to significantly contribute to overall promoter strength and, sometimes, to be strictly required for activity. In the heteromeric polypeptide group, the CCAAT motif is bound by a regulatory complex named NF-Y in human and mouse, CBF in rat and HAP in \textit{S. cerevisiae}. The consensus sequence for these factors, determined by the analysis of 178 bona fide NF-Y binding sites, is CRRCCAA TVR (Mantovani, 1998). However, only for the NF-Y complex the motif has been shown to absolutely require all 5 nucleotides, whereas efficient binding needs specific additional flanking nucleotides (Mantovani, 1998). HAP-like complexes are also known from filamentous fungi (e.g. the AnCF complex in \textit{A. nidulans}, Kato et al., 1998), where, in addition to their activation function, they are
required for gene regulation by pathway specific regulators (Brakhage et al., 1999). Regarding the way of action, the current view is that binding to the CCAAT motif facilitates the binding of further transcriptional factors to DNA, allowing activation of gene expression. Consistent with this view, for several genes which are regulated by CCAAT-binding complexes, additional regulatory proteins recognizing DNA sequences adjacent to or overlapping the CCAAT sequence were discovered. The A. nidulans homoaconitase gene lysF, for instance, was shown to be negatively regulated by a CCAAT box and positively regulated by an adjacent GATA site (see above, chapter 2.2.2.) (Weidner et al., 2001). Caruso et al. (2002) report the isolation of AnBH1, a novel bHLH transcription factor of A. nidulans able to bind as a homodimer to an asymmetric E-box (see above) within the aatA promoter, that overlaps with the AnCF-binding site. AnBH1 represses, whereas AnCF promotes aatA gene expression. Indeed, in most cases analyzed, multimeric CCAAT-binding factors positively activate gene expression whereas repression mediated by CCAAT-binding complexes has rarely been reported (Weidner et al., 2001).

2.3. **Fungal cis- and trans-acting elements in their regulatory networks**

The fashion by which distinct signals influence gene transcription is not straightforward. Rather, multiple signals are acting in parallel in space and time to provide a complex transcriptional network which will finally influence the fate of an organism (Lengeler et al., 2000; D'Souza and Heitman, 2001; Sanchez-Martinez and Perez-Martin, 2001; Lee et al., 2003). As stated in chapter 1.2., mating type genes, nutrients and light are the major regulators controlling sexual development in C. cinerea (Casselton and Olesnicky, 1998; Kües, 2000b). Most information on how these factors regulate gene transcription comes from few well established fungal model organisms. In the following, the principal cis- and trans-acting fungal elements analyzed in chapter 2.2. are considered in regulatory networks involving mating types, nutrients and light.

2.3.1. **Mating type genes**

Hereafter, the molecular way of action of transcription factors encoded by mating type genes is discussed on the basis of some model fungi. Even though there are
some differences between *C. cinerea* and other fungi (ascomycetes, yeast-like basidiomycetes), like the missing need in mushroom fungi of pheromone signaling to attract mates or to promote development of mating structures in the already filamentous mating partners (Casselton and Olesnicky, 1998), the principal molecular mechanisms are to be considered similar, as already demonstrated in some cases.

*Saccharomyces cerevisiae*

*S. cerevisiae* has a single mating type locus with two allelic forms (idiomorphs), *MATa* and *MATa*. The *MATq* locus contains two genes, *a1* and *a2*. The *MATa* locus encodes two genes, the functional *a1* and the unfunctional *a2*. The products of *MATa1* and *MATa2* genes belong to the homeodomain transcription factor family, whilst the product of *MATa1* is a *a1* domain-containing transcription factor (see above, chapter 2.2.3.). Yeast haploid cells transcribe haploid specific genes (*hsp*), among others those encoding for mating type-specific pheromones and pheromone receptors, required during mating events. Thus, *a*-cells transcribe *a*-specific genes and *α*-cells transcribe *α*-specific genes, respectively (Fig. 4). This differential transcription is achieved by an elegant combinatorial strategy involving three proteins: MCM1, a MADS box transcription factor expressed in all cell types (see above, chapter 2.2.4.), *a1* and *a2*. Promoters of *a*-specific genes contain nearly perfect versions of the palindromic MCM1-binding site, the P box (TTTCCTAATTAGAAA), flanked by *a2*-binding sites (Johnson and Herskowitz, 1985; Kronstad et al., 1987; Sauer et al., 1988). *α*-specific gene promoters contain the so called P'Q binding motif (Jarvis et al., 1988). The P' box (TTTCCTAATTAGTNCA) is a modified version of the P box and thus a weak binding site for MCM1. The Q box (TCAATGACAG) is a weak binding site for *a1* (Hagen et al., 1993). In haploid *α*-cells, the *a1* gene product binds cooperatively with MCM1 to the P'Q box and promotes expression of *α*-specific genes (Hagen et al., 1993). In addition, the product of the *a2* gene and MCM1 bind together to the *a2*-P box-α2 site on the promoter of *a*-specific genes and repress their transcription (Fig. 4). In *a*-cells, the same DNA site, bound only by MCM1, is not enough to repress expression of α-specific genes. In these cells, because of the lack of *a1* gene product, α-specific genes are not transcribed (Fig. 4). Diploid (*a/a*) cells contain both alleles of both mating type genes, but have novel properties because of the interaction between the
a1 and α2 polypeptides. The α1/α2 heterodimer represses a set of haploid-specific genes (hsg genes), including α1 and RME1, a repressor of meiosis (Goutte and Johnson, 1988). Because these cells express also α2, α-specific genes are repressed, and because they do not express α1, α-specific genes are not induced (Fig. 4). Therefore, diploids do not produce or respond to mating pheromones and are unable to mate. However, because in diploids, under certain nutritional conditions, RME1 is repressed, diploids are able to undergo meiosis and sporulation (Gow and Gadd, 1995). A sequence alignment of 17 potential α1/α2 binding sites found in the promoters of several hsg genes, including RME1 and MATα1, more precisely defined the consensus sequence for the α1/α2 heterodimer as being SCATGTWAAWWSYRCTCA (Jin et al., 1999).

Fig. 4: Mating type control of cell type in S. cerevisiae (adapted from Mead et al. 2002).

Ustilago maydis

Unlike the ascomycetes, which have single mating type loci (some of which encode two or more gene products, Coppin et al., 1997), many basidiomycetes have two distinct loci. In U. maydis, the a locus (the equivalent of the C. cinerea B locus) has two alleles (α1 and α2), but the b locus (the equivalent of the C. cinerea A locus) is multiallelic, having at least 25, possibly up to 33 alleles (Gow and Gadd, 1995). The α1 and α2 alleles encode each a pheromone precursor (Mfa1 and Mfa2,
respectively) and a pheromone receptor (Pra1 and Pra2, respectively) (Fig. 5, Böcker et al., 1992). In U. maydis, haploid cells that differ at the a locus can sense the presence of a potential mating partner with their pheromone and pheromone receptors, eventually mate and form a dikaryon. At this stage, genetic recognition (self versus non-self) occurs (Fig. 5). If the mating partners contain also different alleles of the multiallelic b locus a stable dikaryon is formed, which grows as a long straight filament and is able to infect corn plants (Fig. 5, Böcker, 2001).

Fig. 5: Early steps of the mating process in U. maydis (from Böcker 2001).
Most knowledge about the pheromone response pathway in basidiomycetes comes from studies with *U. maydis*. In *U. maydis*, all genes located in the *a* and *b* loci are induced upon pheromone stimulation (Urban et al., 1996). This guarantees amplification of the pheromone signal during mating. A short DNA sequence (ACAAAGGGGA) termed "pheromone response element" (PRE), found in the promoter of pheromone-inducible genes (comprising the *a* and *b* genes), is both necessary and sufficient for pheromone induction (Urban et al., 1996). This induction is mediated by the product of the *prf1* gene (for pheromone response factor), encoding for a HMG protein (Prf1) that activates pheromone-inducible genes by binding specifically to PRE motifs in their promoters (Hartmann et al., 1996). The mating process is also known to be under control of the PKA and MAPK pathways: *a* genes require PKA stimulation, whereas *b* genes require both MAPK and PKA stimulation (Kaffarnik et al., 2003). Kaffarnik et al. (2003) demonstrated recently that these two pathways mediate their effect by regulating Prf1 phosphorylation: induction of *a* genes requires PKA sites in Prf1 while induction of *b* genes is dependent on the integrity of both PKA and MAPK sites (Fig. 6).

The *b* locus contains a pair of divergently transcribed genes, *bW* and *bE*, which both code for transcription factors of the homeodomain family (Fig. 5, Casselton and Olesnicky, 1998; Bölker, 2001). A pair of bE and bW proteins from different mating type genes can form a heterodimer thought to act as a transcriptional activator or repressor of several target genes, such as those necessary for meiosis and pathogenesis (Fig. 5, Bölker, 2001). The bE1-bW2 heterodimer was shown in vivo to activate *iga2*, a gene of unknown function located in the *a2* locus. In vitro, an artificially fused bE1-bW2 protein bound to the *bbs1* (for *b* binding sequence 1) promoter region. *bbs1* contains a direct repeat of the sequence ASTGTG, (present also in the *mfa2* gene promoter) and a hsg-like motif (see above, chapter 2.3.1.).
Fig. 6: The Prf1 phosphorylation state determines a or b gene activation in *U. maydis* (from Kaffarnik et al. 2003).
Differently from yeast, mutations in the core sequence of the hsg operator as well as alteration in spacing did not cause a significant reduction in binding affinity (Romeis et al., 2000). Brachmann et al. (2001), in their search for genes belonging to the bE/bW cascade, identified frb52, a gene whose deduced product displays sequence similarities to DNA polymerases. The frb52 gene is up-regulated after b gene stimulation and its promoter contains a region similar to the lga2 bbs1 sequence. EMSA and DNaasel footprint assays indicated this sequence as possible target of the bE1-bW2 fusion protein (Brachmann et al., 2001).

2.3.2. Nutrients

Nutrient concentration in the medium seems to be one of the major environmental signals able to dramatically influence metabolism and morphology in fungi.

A. Nutrient-driven dimorphic transitions

The study of dimorphism, i.e. the condition of having two distinct morphological growth forms (single cells versus hyphae or pseudohyphae) within a species, has focused on the signaling pathways that connect the external environment with a change in cell differentiation. It became evident that these pathways are broadly conserved within fungi. Here, effects exerted by nutrients on transcriptional regulation in two ascomycetes and in the plant pathogen basidiomycete U. maydis are presented.

Saccharomyces cerevisiae

Under conditions of nitrogen limitation, diploid strains of the yeast S. cerevisiae switch from growth as individual oval cells to filamentous growth. A similar phenotype (invasive growth) is observed in haploid cells in response to glucose starvation. In these processes both MAPK and cAMP pathways are involved (Rupp et al., 1999; Pan et al., 2000; D'Souza and Heitman, 2001; Gancedo, 2001). The cell surface protein flocculin FLO11 plays here an important role being required for both haploid adhesive growth and for diploid filamentous growth (Fig. 7). The FLO11 promoter appears to be one of the largest in the yeast genome, integrating MAPK, cAMP, mating type and nutritional signals. Regarding cAMP and MAPK regulation, the
FLO11 promoter contains at least four upstream activating sequences (UAS) and nine repression elements (URS) which together span at least 2.8 kb (Rupp et al., 1999). The precise nature of these motifs has not yet been identified. cAMP regulation of FLO11 requires the PKA catalytic subunit TPK2 which control both FLO8, an activator, and SFL1, a repressor of FLO11, respectively. FLO8 and SFL1 compete for the same binding site on the FLO11 promoter. TPK2 enhances filamentous growth by phosphorylating FLO8 and SFL1, thus enabling FLO8 to bind to the FLO11 promoter, and inhibiting SFL1 dimer formation and binding to the FLO11 promoter (Fig. 7, Pan and Heitman, 2002). Beside cAMP and MAPK, other pathways are able to control yeast filamentous growth. Recently, Braus and co-workers (2003) discovered amino acid starvation as a further nutritional signal activating adhesive growth in a FLO11-dependent manner and leading to induced expression of the FLO11 gene, even in the presence of high glucose and ammonium in the medium. In the FLO11 promoter two distinct types of regulatory elements involved in control by amino acid starvation could be identified. Interestingly, although GCN4, an element of the general control system for amino acid biosynthesis, is required for amino acid starvation-induced adhesive growth and activation of FLO11, sequence analysis of the FLO11 promoter did not predict any GCN4-responsive element sites nor was GCN4 found to bind to any region of the FLO11 promoter in vitro (Braus et al., 2003).
Chapter I

S. cerevisiae

\[
\text{glucose} \rightarrow \text{Gpr1} \rightarrow \text{Gpa} \rightarrow \text{Ras} \rightarrow \text{AC} \rightarrow \text{cAMP} \rightarrow \text{Pde} \rightarrow \text{Bcy} \rightarrow \text{Tpk1,3} \rightarrow \text{Tpk2} \rightarrow \text{Sfl} \rightarrow \text{Flo8} \rightarrow \text{FL011} \rightarrow \text{pseudohyphal growth}
\]

S. pombe

\[
\text{glucose} \rightarrow \text{git3} \rightarrow \text{nitrogen} \rightarrow \text{gpa} \rightarrow \text{gpb} \rightarrow \text{AC} \rightarrow \text{cAMP} \rightarrow \text{cgs} \rightarrow \text{pka} \rightarrow \text{ste11} \rightarrow \text{fbp1} \rightarrow \text{mating sporulation} \rightarrow \text{gluconeogenesis}
\]

Fig. 7: Details of the nutrient-controlled cAMP cascade regulating dimorphic transitions in S. cerevisiae and S. pombe (adapted from D'Souza and Heitman, 2001)

**Schizosaccharomyces pombe**

S. pombe employs both cAMP signalling and MAP kinase cascades to co-ordinately regulate the processes of sexual-driven dimorphism, gluconeogenesis, and stress tolerance (Fig. 7). When glucose is plentiful, cAMP levels and PKA activity increase, inhibiting these processes. On the other hand, when starved for glucose (or nitrogen), reduced cAMP signalling and PKA activity induce these processes (D'Souza and Heitman, 2001). As for the FLO11 promoter in S. cerevisiae, in S. pombe the cAMP and MAPK signaling cascades converge on the promoter of the fbp1 gene, responsible for gluconeogenesis. PKA, differently as for the FLO11 gene, represses, and MAPK activates fbp1 transcription. Two elements on the fbp1
promoter were found to be necessary for its activation. UAS1, the first upstream activation site (TGACGTAG), resembles a CRE motif and is the binding site for the atf1-pcr1 heterodimeric transcriptional activator. Binding of atf1-pcr1 to UAS1 is supported by the MAPK pathway and negatively regulated by PKA. The second upstream activation site UAS2, contains the sequence AAAAAACGAGGGG which resembles to the S. cerevisiae stress response element (STRE), AGGGG, bound by the stress-induced activators Msn2 and Msn4. The sequence also resembles the binding site for three S. cerevisiae glucose repressors MIG1, MIG2 and NRG1. UAS2 is bound by transcriptional activators and repressors regulated by both the PKA and MAPK pathways (Neely and Hoffman, 2000).

**Ustilago maydis**

In *U. maydis*, the agent of corn smut disease, a prerequisite for infection is the formation of a filamentous dikaryon, a process controlled by the mating type genes (see above, chapter 2.3.1.). Environmental signals, however, do also control the fate of cell morphology. High glucose, for instance, induces spore formation, whereas nutrient limitation, air exposure, or acidic pH promotes filamentous growth in haploid cells (D'Souza and Heitman, 2001). The pivotal role of prf1 for mating-associated morphological transitions was already explained. Besides promoter elements responding to cAMP and MAPK stimulation, the analysis of the prf1 promoter defined a cis-acting element which mediates transcriptional activation by carbon sources, such as glucose and fructose, as well as repression by high cAMP levels (Hartmann et al., 1999). The minimal UAS sequence is 40 bp long and comprises the palindromic sequence CACACTATACGTTGGTG (Hartmann et al., 1999). Since mating in this fungus usually occurs on the plant surface, it would be advantageous to sense the presence of the plant, a process which could involve cAMP signaling. Elevated cAMP levels would result in PKA-dependent Prf1 expression, thereby increasing pheromone and pheromone receptor genes, rising thus the chances of being recognized by a mating partner (Kaffarnik et al., 2003).
B. Carbon catabolite repression

*S. cerevisiae* and many other yeasts may thrive on a variety of carbon sources (e.g. galactose, sucrose and maltose), but glucose and fructose are the preferred ones. When one of these sugars is present, the production of enzymes required for the utilization of alternative carbon sources diminishes or ceases at all. This phenomenon is called carbon catabolite repression (Gancedo, 1998). The several pathways with partially opposite metabolic fluxes involved in the utilization of the different carbon sources require a complex pattern of gene expression to allow optimal adaptation to the most convenient substrate available in a certain situation. Consequently, distinct upstream activation sites (UAS) are involved in the transcriptional regulation of the genes concerned (Schüller, 2003).

*Saccharomyces cerevisiae*

The functioning of carbon catabolite repression in yeast is considered as the prototype of eukaryote gene regulation and is represented by the regulation of the GAL genes. The GAL genes, coding for the enzymes required for the catabolism of galactose, are subject to dual control of expression: they are induced by galactose and repressed by glucose. Metabolization of galactose as a carbon source is exerted by the three structural genes GAL1, GAL7, and GAL10. GAL1 encodes the enzyme galactokinase, which phosphorylates galactose to galactose 1-phosphate. GAL7 and GAL10 cooperate to convert galactose 1-phosphate to glucose 1-phosphate, which can be further metabolized (Moore and Novak Frazer, 2002). All these enzymes-encoding genes are essential for metabolism of galactose and are activated by the GAL4-encoded protein GAL4 (see above, chapter 2.2.2.), which binds to upstream activating sequences in the promoters of the other GAL genes (Fig. 8). Since glucose has an overriding effect on galactose, even when galactose is added to a glucose-containing medium, glucose causes repression of the GAL genes through the MIG1 repressor (Gancedo, 1998). MIG1 is a general negative regulator of GAL but also of other glucose-repressed genes like those involved in sucrose and maltose metabolism. MIG1 is a transcription factor containing two N-terminal zinc fingers of the Cys2His2 type and binds to a GC-rich core DNA sequence with consensus WWWWNNSYGGRG (Lundin et al., 1994) found upstream of several glucose-
regulated structural and regulatory genes (Schüller, 2003). In addition to the GC box, an AT-rich region 5' to the GC box is also important for MIG1 binding (Lundin et al., 1994). In the absence of MIG1, however, MIG1 binding sites can also function as activating elements (Wu and Trumbly, 1998). To execute gene repression in the presence of glucose, MIG1 has to recruit the general co-repressor complex Cyc8-Tup1 (Wu and Trumbly, 1998), a process that seems to be dependent on the phosphorylation state of MIG1 (DeVit and Johnston, 1999). Repression of\textit{GAL4} by MIG1 is mediated by a single MIG1 binding site in the \textit{GAL4} promoter (Nehlin et al., 1991). Thus, carbon catabolite repression in yeast is achieved by two mechanisms: on the one hand MIG1 represses \textit{GAL4} expression, on the other hand, MIG1 inhibits \textit{GAL4} function by binding at an upstream repression sequence (URS, Fig. 8) in the \textit{GAL} genes promoters (Moore and Novak Frazer, 2002). Once glucose has been used up, the gene \textit{SNF1} is activated which turns off the MIG1 complex, thus allowing expression of the \textit{GAL} genes (Gancedo, 1998).

\begin{center}
\includegraphics[width=\textwidth]{carbon_catabolite_repression.png}
\end{center}

\textbf{Fig. 8: Carbon catabolite repression in \textit{S. cerevisiae} (from Gancedo 1998).}

\section*{C. Nitrogen catabolite repression (NCR) in \textit{Saccharomyces cerevisiae}}

\textit{S. cerevisiae}, like most microorganisms, distinguishes between preferred (e.g. glutamine, asparagine, ammonium) and non-preferred nitrogen sources (other N-containing amino acids, GABA, allantoin). Nitrogen catabolite repression (NCR) – in following the “carbon catabolite repression” diction – is the mechanism for achieving this selectivity and is the preeminent control exerted over nitrogen catabolic gene expression. In the presence of a preferred nitrogen source in adequate supply,
transcription of genes encoding the proteins needed to metabolize non-preferred nitrogen sources is repressed. This control is exerted at transcriptional level by the GATA factors, so called for their characteristic to bind to a GATA-like motif on the promoter of the genes they control. GATA factors (see above, chapter 2.2.2.) that have clearly been shown in yeast to mediate nitrogen-responsive gene expression include a pair of transcriptional activators, GLN3 and GAT1/NIL1, and a pair of transcriptional repressors, DAL80 and DEH1/GZF3 (Cooper, 2002). Expression of all GATA factor encoding genes, except GLN3, is regulated by all of the GATA factors, whereas GLN3, which can be considered to be the initial regulator of the NCR response, has to date not been demonstrated to be transcriptionally regulated to any significant extent (Cooper, 2002). Tor proteins, responsible for the sensing of nitrogen sources, influence the cellular localization of GLN3: in the cytoplasm GLN3 is inactive, in the nucleus it initiates NCR by activating GAT1. However, most GATA factor-activated genes require both GLN3 and GAT1 for expression. GLN3 and DAL80 can bind to the same GATA elements in the DAL3 promoter, which prompted the hypothesis that DAL80 and DEH1 repress gene expression by competing with the GLN3 and GAT1 activators for binding to the same target GATA sites (Coffman et al., 1997). However, GLN3 (and likely GAT1) binds to single GATA elements while DAL80 and DEH1, which both form dimers through leucine zipper motifs at their C-termini, require two GATA elements for binding (Svetlov and Cooper, 1998). For DAL80, additionally, these GATA elements must be 15-35 bp apart, in either orientation in the promoter, and oriented head-to-tail or tail-to-tail, but not head-to-head (Cunningham and Cooper, 1993).

2.3.3. Light

The influence of light on transcriptional activity in fungi is particularly well studied in the ascomycete *N. crassa*. Several developmental processes are in this organism under blue light control: induction of carotenoid biosynthesis in the mycelium, conidia production and circadian clock regulation are some examples (Ballario and Macino, 1997). *albino-3* (*al-3*), encoding the carotenoid biosynthetic enzyme geranylgeranylpyrophosphate synthetase, was one of the first genes in *N. crassa* shown to be induced by light (Carattoli et al., 1994). APE (for *al-3* proximal element), the identified promoter motif on the *al-3* gene responsible for light-induction, has the consensus
sequence GAANNTTGCC and can be found as well in other light-regulated *N. crassa* genes (Ballario and Macino, 1997). On the clock control gene 2 (ccg-2) promoter, for instance, deletion of one of two APE motifs abolishes the light induction of the gene suggesting that the presence of a single APE site is not sufficient to confer light induction to the gene (Ballario and Macino, 1997). The analysis of two *N. crassa* mutants impaired in blue light response revealed the two white collar genes, wc-1 and wc-2, as being the elements responsible for transmitting the light signal in this fungus (Ballario and Macino, 1997). WC-1 and WC-2 are nuclear transcription factors containing trans-activation and zinc finger DNA-binding domains characteristic of the vertebrate GATA factors (see on top, Ballario et al., 1996; Linden and Macino, 1997). WC-1 and WC-2 can form the White Collar Complex (WCC) by heterodimerizing via PAS domains and act as positive elements in light signalling (Linden and Macino, 1997; Talora et al., 1999; Froehlich et al., 2002). Light causes phosphorylation of WC-1 but not of WC-2 and the phosphorylation level of WC-1 depends on light exposure (Talora et al., 1999). Interestingly, WC-1 and WC-2 were shown to bind two canonical GATA motifs on the *al-3* promoter, found in close proximity of the APE sites (Ballario et al., 1996; Linden and Macino, 1997). That WC-1 can directly act as photoreceptor was only recently demonstrated by exploiting WC-1’s role in the regulation of the circadian clock in *N. crassa*. WC-1 is essential for light resetting of the circadian cycle by the induction of *frq* (frequency gene) transcription upon a pulse of blue light, whereas WC-2 is not required for light-induced transcription of *frq* (Ballario and Macino, 1997). On the *frq* promoter two sequences termed LRE (light responding element) contain each two GATA motifs. Similar sequences are found on the *al-3* promoter (Fig. 9). Both LREs in the *frq* promoter are bound by WCC and deletion of the LRE sequence proximal to the start codon specifically affects light input to the *N. crassa* clock. The presence of the GATA sequence repeats in the LREs is consistent with the WCs binding as a heterodimer with the zinc finger of each WC interacting with one GATA motif (Froehlich et al., 2002). Alignment of the imperfect repeats found in the LREs with the LRE from *albino-3* highlights CGATN-CCGCT as the consensus binding sequence, with the GATN motif having been shown for all three LREs to be necessary for binding of the WC-1/WC-2 complex (Fig. 9). Whether the APEs and the GATA-containing LREs are distinct elements or are part of a common motif remains to be shown.
proximal
 CGAT    CCGCT
 CGAT    CCCCT

distal
 TGATG   CCGCT
 CGATG   ACGCT

al-3
 CGATA   CCGCA
 CGATAATACGCT

Fig. 9: Alignment of the two halves of the imperfect repeats of every LRE found in frq and al-3 (from Froehlich et al. 2002).

2.4. Concluding remarks

It is evident that the contribution to transcriptional regulation as a consequence of the interaction between cis- and trans-acting elements is not the simple result of a "key and lock" mechanism. Rather, the process is influenced at different levels by many parameters. It is the sum of these that will ultimately decide on the effect on gene transcription exerted by the interaction of cis- and trans-acting elements. Thus, even though the binding of a transcription factor to its DNA target sequence is mediated by defined protein domains, the same DNA-binding domain allows many transcription factors to bind to several dissimilar sequences. Similarly, it is possible for a defined DNA sequence to be recognized by more than one factor. For instance, whilst the transcription factors CTF/NFI and C/EBP both bind to the CCAAT box, they do so via completely different DNA-binding domains, with C/EBP having a basic DNA-binding domain, whereas CTF/NFI has a DNA-binding domain distinct from that of any other factor (Latchman, 1998). It is therefore unlikely that the existence of several distinct DNA-binding domains reflects the need of the factors to bind to distinct types of DNA sequences. Rather, it seems perfectly possible that one DNA-binding motif could be present in different factors binding to distinct DNA sequences, because of factor-specific amino acid variations (Latchman, 1998). Regarding the contribution to transcriptional regulation by cis-acting sequences, the presence of a particular motif on a gene's promoter does not necessarily mean its involvement in the regulation of transcription. Thus, despite binding of an ARE-GST fusion protein to different GATA sites on the A. nidulans otaA gene promoter, point mutations in these sites did not change OTAse activity in vivo under the same conditions (Dzikowska et al., 2003). Similarly, GCN4 does not activate transcription through ATF/CREB sites in
vivo, even though it binds these sites in vitro (Garcia-Gimeno and Struhl, 2000). Finally, despite the existence of several defined core DNA recognition sequences, many transcription factors still bind to unique DNA motifs that show no similarity with already known cis-acting sites and that have to be defined individually.

Therefore, to understand the role exerted by cis-acting sequences on the transcriptional activity of a gene, the specific analysis of each putative cis-acting element and, moreover, the systematic exploration of every short promoter stretch for their contribution to transcriptional regulation of a gene in a particular context, is a must.

3. Transcriptional regulation of fruiting body development in C. cinerea

In C. cinerea there are few tentative experiments done to reveal the genetic interactions regulating sexual processes and, ultimately, fruiting body formation. An interesting result in this direction comes from the analysis of the two C. cinerea mutants, pcc1 and clp1. pcc1 was identified by a mutation that induces pseudo-clamp formation and fruiting body development in a homokaryon without previous mating (Murata et al., 1998). The phenotype is suppressed by complementing the mutant with the pcc1 gene, predicted to encode a protein containing an HMG box and a putative nuclear localization signal (Murata et al., 1998). The clp1 mutant, isolated in an AmutBmut background, does not show clamp formation (Inada et al., 2001). Transcripts of the corresponding clp1 gene are normally detected only in the presence of A protein heterodimers formed when homokaryons with compatible A genes are mated (Inada et al., 2001). Moreover, forced expression of the wildtype clp1 gene by promoter exchange induces clamp development without the need for a compatible A gene combination (Kamada, 2002). These data, together with the presence, in the clp1 promoter, of a putative recognition motif for the A mating type products heterodimer (Inada et al., 2001), suggest that clp1 is likely a target of the HD1/HD2 heterodimer and unpublished results are pointing in this direction (T. Kamada, personal communication). Fig. 10 shows a working model on the first unraveled genetic steps regulating fruiting body development in C. cinerea based on findings with the clp1 and pcc1 genes (Fig. 10, Kamada, 2002).
Chapter 1

A-protein heterodimer

\[ \text{clp1 gene} \rightarrow \text{clp}1 \]

\[ \text{pcc1 gene} \rightarrow \text{pcc1} \]

The model predicts the \textit{pcc1} gene product to be a putative transcription factor inhibiting genes required for clamp cell and fruiting body formation. In an Aon background (i.e. a strain, in which the A mating type pathway is activated), the A-protein heterodimer would indirectly release the PCC1 inhibition by activating the expression of CLP1, an inhibitor of PCC1 (Fig. 10, Kamada, 2002).

In higher basidiomycetes, the current lack of information on transcriptional gene regulation is due, on the one hand, to the missing suitable fast assays and, on the other hand, to the absence of suitable marker genes accompanying fruiting events. In the last years, the improvement of the available technical tools and the progress made to understand the genetics of basidiomycetes encourage trying new ways to understand the mechanisms of transcriptional regulation during fruiting body formation.

Fig. 10: Model illustrating the first unraveled genetic events regulating fruiting body development in \textit{C. cinerea} (adapted from Kamada, 2002).
4. Aim of the thesis

Two β-galactoside binding lectins (galectins) CGL1 and CGL2 were purified from fruiting mycelia in *C. cinerea* (Cooper et al., 1997). In higher eukaryotes, the ability of galectins to bind β-galactoside containing sugars appears to be exploited for various biological functions which go from pre-mRNA splicing and regulation of cell growth to extracellular cell adhesion, embryogenesis and cancer (Walser, 2004). Recently, the function of the two *C. cinerea* lectins has been investigated (Walser et al., 2004a; Walser et al., 2004b). By virtue of their cellular localization within fruiting body tissues, one possible role is the cross-linking of extracellular matrix components during fruiting events. The epitopes recognized by the *C. cinerea* galectins are shown to be lipid-borne and the substrate site is strictly conserved between both galectins. However, if present in two distinct cellular environments, different roles for CGL1 and CLG2 in *C. cinerea* are plausible (Walser et al., 2004b).

Expression patterns of the two galectins in *C. cinerea* were shown to highly correlate with various events in the fruiting process (Boulianne et al., 2000). cg/2 gene expression starts at the stage of primary hyphal knot formation and continues during the following developmental stages of primordia formation. CGL1 is first detected in secondary hyphal knots (initials) which arise from primary hyphal knots after light stimulation. Therefore, cg/1 might be directly or indirectly controlled by a light signal as are secondary hyphal knots (Boulianne et al., 2000).

Like primary hyphal knot formation, CGL2 expression is under control of the major regulators of the fruiting process: mating type genes, light and nutrients. CGL2 is detected mainly in dikaryons and is induced in the dark, apparently upon nutrient depletion. No expression is observed in cultures kept in continuous light. High glucose or nitrogen concentrations in the substrate inhibit sclerotia and primary hyphal knots as well as CGL2 expression. Moreover, the regulation of the cg/2 expression was already shown to be controlled at transcriptional level (Boulianne et al., 2000). These facts prompted us to use the cg/2 gene as a marker of the fruiting process and, by dissecting the cg/2 promoter, understand how fruiting signals are interpreted at the promoter level to regulate transcription.
The aim of this work was the set up of a method to analyze gene transcription in *C. cinerea* and the analysis, by the developed method, of the *cgl2* promoter *in vivo* in *C. cinerea*. 
5. References


Cooper, T. G., 2002. Transmitting the signal of excess nitrogen in *Saccharomyces cerevisiae* from the Tor proteins to the GATA factors: connecting the dots. FEMS Microbiol. Rev. 26, 223-238.


Kjaerulff, S., Dooijes, D., Clevers, H., Nielsen, O., 1997. Cell differentiation by interaction of two HMG-box proteins: Mat1-Mc activates M cell-specific genes in
S. pombe by recruiting the ubiquitous transcription factor Ste11 to weak binding sites. EMBO J. 16, 4021-4033.


Tebb, G., Moll, T., Dowzer, C., Nasmyth, K., 1993. SWI5 instability may be necessary but is not sufficient for asymmetric HO expression in yeast. Genes Dev. 7, 517-528.


Uno, I., Ishikawa, T., 1974. Effect of glucose on the fruiting body formation and adenosine 3',5'-cyclic monophosphate levels in *Coprinus macrorhizus*. J. Bacteriol. 120, 96-100.


Wu, J., Trumbly, R. J., 1998. Multiple regulatory proteins mediate repression and activation by interaction with the yeast Mig1 binding site. Yeast 14, 985-1000.


CHAPTER II

Promoter analysis of cgl2, a galectin encoding gene transcribed during fruiting body formation in Coprinopsis cinerea (Coprinus cinereus)

Rinaldo C. Bertossa¹, Ursula Kües², Markus Aebi¹, Markus Künzler¹

Submitted to Fungal Genetics and Biology

¹ Institute of Microbiology, ETH Zurich, Schmelzbergstrasse 7, 8092 Zurich, Switzerland.

² Molecular Wood Biotechnology, Institute of Forest Botany, Georg-August-University Göttingen, Büsgenweg 2, 37077 Göttingen, Germany.
1. Abstract

In the homobasidiomycete *Coprinopsis cinerea* the two galectins CGL1 and CGL2 are specifically expressed in the developing fruiting body, from the first stages of hyphal aggregation until completion of tissue differentiation in the primordia. In particular, CGL2 expression starts in the dark when primary hyphal knots form within the mycelium. Onset of CGL1 expression correlates with the light induced formation of the more compact secondary hyphal knots. In mycelia cultivated in constant light, primary hyphal knot production and fruiting body formation is inhibited and so is galectin expression. In cultures kept in constant darkness, fruiting is also inhibited and CGL2 expression arrests with aging of the mycelium. In this study, we analyzed the promoter of the *cg/l2* gene by measuring transcript levels by quantitative real-time PCR and show that regulation of CGL2 expression occurs at the transcriptional level. A minimal promoter sufficient to confer regulated expression of a heterologous reporter gene and comprising 627 base pairs from the start codon was defined. On the minimal promoter we identified a 120 bp sequence mediating induction of the *cg/l2* gene in constant darkness. Along with direct repeats (TGGAAG/TGGAAG/GGAA), the sequence contains a CRE consensus site (cAMP-responsive element, TGCGTCA) suggesting the involvement of cAMP signaling in *cg/l2* activation. No specific elements responsible for light repression and mating type regulation were found in the promoter.

**Key words:** promoter analysis, galectin, *Coprinopsis cinerea*, fruiting body, real-time PCR, filamentous fungi, CRE element.
2. Introduction

The homobasidiomycete Coprinopsis cinerea (formerly Coprinus cinereus) is used as a model organism to study fruiting body formation and other developmental aspects in basidiomycetes. Fruiting body formation is a complex developmental process regulated by both environmental and genetic signals (Moore, 1998; Kües, 2000; Kamada, 2002). As a first, yet unspecific step towards fruiting body development, primary hyphal knots develop in the vegetative mycelium within darkness by localized multiple branching combined with restricted hyphal tip growth. Without a stimulating light signal, primary hyphal knots mature into globular sclerotia, multicellular resting bodies. In contrast, when a light signal is given, primary hyphal knots appear to transform into secondary hyphal knots, small compact round hyphal aggregates. These plectenchymatic structures are specific for the fruiting pathway and, within them, tissue differentiation initiates primordia development. Light is needed for correct formation of the cap and stipe tissues within the fruiting body primordia and for induction of karyogamy within the basidia present in the fully established primordia. Upon karyogamy, meiosis and basidiospore formation occurs parallel to fruiting body maturation i.e. stipe elongation and cap expansion (Walser et al., 2003; Kües et al., 2004).

Nutrient depletion and light-dark changes are major triggering signals for initiation and continuation of fruiting body formation on the dikaryon (Kües, 2000; Kües et al., 2004). Among the genetic signals, the mating type genes are known to play a pivotal role in the activation of sexual development (Kües et al., 1998; Kües et al., 2002). Like the majority of homobasidiomycetes, C. cinerea has two unlinked mating type loci, A and B, encoding homeodomain transcription factors and pheromone and pheromone receptors, respectively. The A and B mating type pathways are activated in the dikaryon, where mating type gene products of different specificity are present (Casselton and Olesnicky, 1998). In monokaryons, primary hyphal knots may form in the absence of activated mating pathways. However, activation of the A mating type pathway strongly increases primary hyphal knot formation. Moreover, it induces development of secondary hyphal knots that develop further until tissue differentiation in the primordia is completed (Kües et al., 1998). Additional activation of the B mating type pathway enhances the effects of the A mating type genes on
initiation of fruiting body development and it promotes maturation of primordia into complete fruiting bodies. Activation of the B pathway alone, however, is not sufficient to induce fruiting (Kües et al., 2002). Strains carrying specific mutations in the mating type genes (Amut and Bmut) are self-compatible by activation of both mating type pathways and therefore can undergo the complete fruiting pathway (Swamy et al., 1984). Because they have only one haploid genome and produce abundant asexual oidia (Kertesz-Chaloupková et al., 1998), such strains are easily accessible by transformation and thus widely used in genetic and developmental studies of fruiting body formation (Granado et al., 1997; Boulianne et al., 2000; Inada et al., 2001).

At the molecular level, the process of fruiting body formation is poorly understood. Besides the mating type genes, only few other genes have been described to play a role in this developmental process (Kües, 2000; Kamada, 2002). For a better understanding of the pathway, both regulatory and structural genes essential for the process have to be identified. Regarding the latter, two closely related galectins (β-galactoside binding lectins), CGL1 and CGL2, were previously shown to be specifically expressed in the developing fruiting body (Boulianne et al., 2000). The galectin ORFs share 87% DNA sequence identity and are arranged in tandem, spaced from each other by 1367 base pairs (from the CGL1 stop to the CGL2 start codon). Both genes have two introns located at conserved positions in the 5’ and the 3’ untranslated regions. Over a length of about 700 base pairs (bp), the regulatory regions 5’upstream to the start codon show 61% DNA identity to each other. CGL2 expression initiates in the dark at the time of primary hyphal knot formation. CGL1 appears later, at the light-induced stage of secondary hyphal knot formation. Expression of both galectins increases and continues during primordia development to end at the stage of meiosis (Charlton et al., 1992; Boulianne et al., 2000). Galectin expression, like fruiting, is induced by the A mating type pathway and repressed by constant light and high nutrient levels in the medium (Boulianne et al., 2000). In constant darkness, CGL2 expression declines with aging of the mycelium and maturation of sclerotia (Bottoli, 2001). In order to dissect the signaling pathways governing this regulation, we analyzed the cg/2 promoter by quantitative real-time PCR in an Amut Bmut background.
3. Materials and methods

C. cinerea cultivation and transformation

The self-compatible *C. cinerea* homokaryotic strain AmutBmut (*A43mut, B43mut, pab-1*, Swamy et al., 1984; May et al., 1991) was transformed following the protocol of Granado et al. (1997). Plasmid pPAB1-2 (Granado et al., 1997), carrying the para-aminobenzoic acid (PABA) synthase gene (*pab-1*) was used for selection in cotransformation experiments (Kües et al., 2001; Kües et al., 2002) with reporter constructs. For analysis of promoter activities, 10 different transformants were inoculated with constant spacing to each other on a 4.5 cm Ø circle centered in the middle of a 9 cm Ø Petri dish containing solid YMG/T complete medium (Granado et al., 1997). Plates were then incubated for different number of days (depending on the experiment) at 37°C in constant light (Kertesz-Chaloupková et al., 1998) or in constant darkness (in ventilated closed black boxes).

Reporter constructs

Clonings were performed according to standard procedures (Sambrook and Russell, 2001). All reporter constructs are based on plasmid pPW-6 (kindly provided by P. J. Walser), a pBluescript® II KS (Stratagene Europe, Amsterdam, The Netherlands) derivative containing a codon-optimized green fluorescent protein gene (*gfp*) (Spellig et al., 1996) under the control of 1290 bp upstream of the *cg/2* start codon. For cloning reasons, 93 bp (transcribed) immediately upstream (including the 5' intron) and around 800 bp downstream (including the 3' intron) of the GFP coding region (Boulianne et al., 2000) originate from the *cg/1* gene. The number in the name of reporter constructs (e.g. in “p1290”) indicates the length in base pairs of the 5' sequence to the ATG start codon present in the respective construct. 5'-deletion constructs were generated by PCR-amplifying promoter fragments using appropriate primers. The tail of the distal primer (relative to the start codon) contained a Sapl restriction site which was also present in the pPW-6 plasmid backbone, right 5' upstream of the *cg/2* promoter sequence. The obtained PCR fragments were used to replace the original Sapl-x promoter fragment in the pPW-6 plasmid, causing the 5'-shortening of the promoter (x indicates an arbitrary chosen restriction site in the
Chapter II

promoter fragment, downstream of SapI). Linker-scanning constructs (arbitrary linker numbers indicated by the second number separated by a dash in the construct names) were generated by standard PCR procedures. The linkers replaced 10 or 20 bp of the original promoter sequence in the p742 construct with an unrelated sequence including a HindIII restriction site. Replaced sequences were selected according to any of the following criteria: regular spacing of the linkers along the promoter sequence; striking homology or divergence between cg/1 and cg/2 promoters; similarity to promoter sequences of fungal genes involved in fruiting; homology to known eukaryotic promoter motifs defined by blasting the cg/2 promoter against the public version of the Transfac database 6.0 (http://www.gene-regulation.com/cgi-bin/pub/databases/transfac/search.cgi?) using Sigscan Version 4.05 (http://bimas.dcrtp.nih.gov/molbio/signal/). Table 1 shows the linker sequences introduced, the original sequences that were replaced, their position and the rationale for each replacement. Internal deletion constructs were generated by combining promoter fragments of respective linker-scanning constructs using the introduced HindIII site and additional unique restriction sites in the promoter region. Plasmids p742-(11/AatII) and p742-(AatII/7) were generated by removing the respective HindIII-AatII (AatII at position -559/-554) fragments from construct p742. All constructs were verified by DNA sequencing.
Table 1: Linkers used in the linker scanning analysis.

<table>
<thead>
<tr>
<th>Linker</th>
<th>Linker sequence</th>
<th>Original sequence</th>
<th>Position</th>
<th>Rationale for replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TAAGCTTCCA</td>
<td>CGTATAAAAG</td>
<td>-135/-126</td>
<td>Putative TATA box (found in cgfl at a similar position)</td>
</tr>
<tr>
<td>2</td>
<td>CCGAAGCTTG</td>
<td>GATGCAATGA</td>
<td>-183/-174</td>
<td>Putative mating type proteins binding sequence, present also in cgfl and clp3</td>
</tr>
<tr>
<td>3</td>
<td>ATCCAAGCTTT</td>
<td>TGTGTCCGGG</td>
<td>-226/-217</td>
<td>Missing DNA stretch in the cgfl promoter sequence; linker spacing</td>
</tr>
<tr>
<td>4</td>
<td>AAGCTTTGAGT</td>
<td>TTCATCTCCG</td>
<td>-315/-306</td>
<td>Inverted repeats; linker spacing</td>
</tr>
<tr>
<td>5</td>
<td>ACTAAGCTTT</td>
<td>TTCTATTCGC</td>
<td>-406/-397</td>
<td>Linker spacing</td>
</tr>
<tr>
<td>6</td>
<td>CAAGCTTAGT</td>
<td>TGGCATAGTG</td>
<td>-429/-420</td>
<td>Adjacent CAAT and Cre1-target (GGAG) sites; CAAT stretch homologous to clp1 and cgfl promoters; Cre1-target site homologous to pcc1 promoter4</td>
</tr>
<tr>
<td>7</td>
<td>TCAAGCTTCCG</td>
<td>GTCGAACCTC</td>
<td>-496/-487</td>
<td>Linker spacing</td>
</tr>
<tr>
<td>8</td>
<td>CATGTGCAAGCTTACCAGCA</td>
<td>AGGAAAATGAACCGCGCTTT</td>
<td>-524/-505</td>
<td>Homology with DNA stretch in the pcc1 promoter; identical sequence as in the cgfl promoter</td>
</tr>
<tr>
<td>9</td>
<td>TCGAAGCTTGGTACCTGCTA</td>
<td>CTGGGATGATGCAACGAC</td>
<td>-576/-557</td>
<td>Inverted repeats; within sequence responsible for dark induction5</td>
</tr>
<tr>
<td>10</td>
<td>TCACCGTAGGAAGCTTGA</td>
<td>GACTATAACACCAGACTCCC</td>
<td>-602/-583</td>
<td>Within sequence responsible for dark induction</td>
</tr>
<tr>
<td>11</td>
<td>GCTAAGCTTGA</td>
<td>CAGGACTGCG</td>
<td>-616/-607</td>
<td>Putative CRE motif (cAMP responsive element); inverted repeats</td>
</tr>
</tbody>
</table>

1 The HindIII restriction site is underlined.
2 Position relative to the ATG start codon.
3 Inada et al., 2001; Kues et al. 2004.
4 pcc1 accession no. is AB007769 (Murata et al. 1998).
5 Promoter sequence between linkers 11 and 7.
RNA extraction and cDNA synthesis

The mycelium from different plates (with 10 individual transformants each) incubated under identical conditions, was harvested by scraping the agar surface with a scalpel, pooled and immediately frozen in liquid nitrogen. For the assays reported in Fig. 1, the mycelia from two identical plates were pooled (10 individual transformants in total). For the 5'-deletion assay, the mycelia from 10 plates were combined to give 5 pools of 20 individual transformants each (100 individual transformants in total). In all other experiments, a total of 120 different transformants were collected in 4 different pools of 30 individual transformants each (in case of the linker constructs 1, 3, 4, 5, 7 and 11) and 3 pools of 40 individual transformants each (in all other cases), respectively. Total RNA was extracted right after lyophilization of frozen mycelium using a hot phenol extraction procedure (Dudler et al., 1991) optimized for C. cinerea (Mehra, 1996; Wilkins and Smart, 1996; Rapley and Manning, 1998; Bottoli, 2001). Lyophilized mycelium (ca. 50-200 mg) was grinded in a FastPrep machine (BIO101; 3 pulses of 45 sec each at maximal speed) by adding acid-washed glass beads (400-600 μm Ø) in a 1:5 (vol:vol) ratio (glass beads:mycelium). After grinding, ca. 400 μl of a glass beads-mycelium mix (containing ca. 20-40 mg mycelium) was transferred to 2ml-centrifuge tubes for RNA extraction. From this step on, a maximum of 6 samples were handled at a time. For the extraction, 0.9 ml of hot (80°C) phenol [AquaPhenol (QBiogene, Illkirch Cedex, France) with 3.6 μl/ml freshly added β-mercaptoethanol] was added and the samples vigorously shaken for 1 min (by hand and vortexing) before addition of 0.9 ml of hot (80°C) 2xHB buffer (100 mM NaCl, 100 mM LiCl, 5 mM EDTA). After vigorous shaking for 1 min, the samples were successively placed on a shaking heating block set at 75°C until all 6 samples were handled. All samples were shaken for an additional minute and then centrifuged (3700 g) at 4°C for 15 min. After centrifugation, the upper aqueous phase was transferred into a tube containing an equal volume of phenol/chloroform/isoamylalcohol (25:24:1). Samples were shaken during 3-5 min at 75°C and afterwards centrifuged (3700 g) at 4°C for 10 min. The upper aqueous phase was transferred to new tubes containing an equal volume of chloroform/isoamylalcohol (24:1) and stored on ice. After extraction (3 to 5 min by shaking the samples by hand and vortexing), the samples were centrifuged (2100 g) at 4°C for 10 min and the upper aqueous phase was transferred to RNase-free tubes.
Chapter II

RNA was precipitated by adding 1/10 (vol:vol) RNase-free 3M sodium acetate pH 5.2 and 1 volume of cold 100% 2-propanol. After short vortexing, samples were placed immediately at -80°C for at least 30 min. After precipitation, samples were centrifuged (10300 g) at 4°C for 15 min and the supernatant discarded by aspiration. The pellet was washed thoroughly with 300 µl cold 70% ethanol, centrifuged (10300 g) for 10 min at 4°C and dried either at room temperature or 50°C, after discarding of the supernatant. The RNA was dissolved in 20-50 µl sterile RNase-free water by repeated heating (65°C) and vortexing and stored at -80°C until utilization. Purity and integrity were checked by calculating the 260/280 nm absorption ratio (Rapley and Heptinstall, 1998) and by visual check after separation of 10 µg RNA on a 6.7% formaldehyde gel (1% agarose). RNA was diluted to 2 µg/µl for use.

cDNA was synthesized using the SuperScript™ II RNase H- Reverse Transcriptase kit of Gibco (Invitrogen - Life Technologies, Basel, Switzerland) according to the manufacturer’s instructions. 2 µg RNA per reaction volume (20µl) were used and the obtained cDNA was stored at -20°C until use.

Real-time PCR assay, quantification and statistics

For quantitative transcription analysis, quantitative real-time PCR with TaqMan technology was used (Gibson et al., 1996). Reaction cocktails containing the TaqMan Universal PCR Master Mix (Applied Biosystems, Rotkreuz, Switzerland), forward and reverse primers, TaqMan probe and distilled water were prepared as a master mix to obtain a final volume of 25 µl per reaction. Aliquots of the reaction cocktail were pipetted into the wells and cDNA (10 to 20 ng per reaction) subsequently added. Reactions were carried out to detect one single transcript species per reaction (“singleplex” reactions), either from gfp reporter constructs, or the C. cinerea endogenous genes cg/2 (GenBank accession number AF130360) or benA (for β1-tubulin, GenBank accession number AB000116). In order to exclude genomic DNA amplification, TaqMan probes were designed to anneal to exon-exon junctions (Table 2). Forward and reverse primers (60°C annealing temperature) hybridized to the two exons 5' and 3' of the TaqMan probe, respectively (Table 2). The real-time PCR reactions were carried out in 0.2-ml 96 wells-plates using the ABI Prism 7700 sequence detector (Applied Biosystems) and applying the universal thermal cycling parameters (1x 2 min at 50°C and 10 min at 95°C; 40x 1 min at 60°C and 15 sec at 95°C) suggested by the supplier. Primers and TaqMan probes concentrations giving
the lowest threshold cycle (C_T) value (Table 2) were selected for further analysis. Extensive controls (real-time reactions on plasmid DNA containing the genomic target sequences and on RNA) were done to ensure that no genomic DNA was detected (data not shown).

Data obtained by real-time PCR were analyzed using the sequence detector 7700 software (SDS software, version 1.9, Software Package, Applied Biosystems). Amplification efficiency of all genes of interest were shown to be equivalent allowing the application of the comparative C_T method for transcripts quantification (described in ABI Prism 7700 Sequence Detection System, User Bulletin #2, 2001). *benA* was used as house-keeping gene to normalize *cgI2* and *gfp* values.

To filter out the pool-intrinsic *cgI2* variation, after normalization against *benA*, the *gfp* result was further normalized against *cgI2* as detected in the respective sample and multiplied by the *cgI2* pool-average (average result of all pools analyzed under the same condition). For the samples grown in constant light this was not possible due to low and variable relative *cgI2* levels (more than the arbitrarily set value of one order of magnitude). Statistical relevance of the differences in transcripts levels was verified by ANOVA (SPSS 11.0 software) and Tukey posthoc tests.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward primer</th>
<th>TaqMan probe¹</th>
<th>Reverse primer</th>
<th>Amplicon size</th>
</tr>
</thead>
<tbody>
<tr>
<td>cgi2</td>
<td>tgcacggtgtactctgcatg, 50 nM</td>
<td>6FAM-tccccctgtgataggttacgtt, 100 nM</td>
<td>gcaatatgtggtggtgtg, 0.9 μM</td>
<td>85 bp</td>
</tr>
<tr>
<td>gfp</td>
<td>tcaagtctgtgaaggtgctgct, 50 nM</td>
<td>6FAM-cccgctgataggtctggctct, 100 nM</td>
<td>tggactggtgtgtgaccc, 0.3 μM</td>
<td>107 bp</td>
</tr>
<tr>
<td>benA</td>
<td>caactcaactctcactgagag, 50 nM</td>
<td>VIC-cgctgtaacagtgtcttcctt, 100 nM</td>
<td>gaagccggtatcgaagaagt, 0.9 μM</td>
<td>78 bp</td>
</tr>
</tbody>
</table>

¹ TaqMan probes span an exon-exon boundary. Nucleotides binding at the 5' or at the 3'-exon (relative to the respective intron) are in italics or in normal characters, respectively. Dyes used (6FAM™ or VIC™) are indicated.
4. Results

Quantitative analysis of transcription in Coprinopsis cinerea by real-time PCR

The regulation of galectin expression during fruiting body formation (i.e. induction in constant darkness and repression in constant light) was shown semi-quantitatively in previous studies both on protein and on transcript level (Charlton et al., 1992; Boulianne et al., 2000; Bottoli, 2001). Aims of this work were the development of a quantitative method to determine the transcriptional activity of specific genes in C. cinerea and a subsequent analysis of the cg/2 promoter.

We used quantitative real-time PCR with TaqMan technology (Giulietti et al., 2001). In this method, transcript levels of genes of interest are determined relative to internal control genes (house-keeping genes) by monitoring the amplification of RNA-derived cDNA in real time, using gene-specific primers and fluorescently labeled TaqMan probes. In this study, the house-keeping gene benA, coding for C. cinerea β1-tubulin (Matsuo et al., 1999), served as internal control. For cg/2 promoter analysis, reporter plasmids expressing the green fluorescent protein (GFP) under the control of modified cg/2 promoter sequences were co-transformed into the homokaryotic C. cinerea strain AmutBmut together with a plasmid containing the para-aminobenzoic acid (PABA) synthase gene for selection (Kües et al., 2001; Kües et al., 2002). cDNA derived from several (10 to 120) individual PABA-prototrophic transformants, pooled in groups of 10 to 40, were analyzed. This is in contrast to other methods favoring the determination of the exact copy number and topology of the co-integrated reporter plasmid in a particular fungal transformant and the separate analysis of a comparably low number of individual transformants (e.g. Carattoli et al., 1994; Dzikowska et al., 2003). In this study, every PABA-prototrophic transformant, regardless of whether and how the co-transformed GFP reporter plasmid was integrated in the genome, was included in the analysis. The transcript level determined is thus the average from a pool of a relative large number of individual transformants (see materials and methods for details).
Chapter II

In pooled transformants, transcription regulation of the gfp reporter gene is similar to that observed for the endogenous cg/l2 gene.

In order to test the chosen methodology, the relative transcript levels of the endogenous cg/l2 locus and the gfp transgene were compared over a 8-days period. Expression of the transgene was driven by the largest cg/l2 promoter fragment available (1290 bp upstream to the start codon). Because of inoculating 10 clones on one 9 cm Ø plate instead of one inoculum per plate as in a previous work (Boulianne et al., 2000), the time point for mycelium to reach the same developmental stage on the plate was slightly different: days 3 and 6 in the present study correspond approximately to days 4 and 7, respectively, in the previous study. At day 3 in this study, the mycelium did not reach the edges of the Petri dish and was still actively growing. At day 6, the Petri dish was fully covered by mycelium and primary hyphal knots were formed in the dark. Two days later, in constant darkness, part of the primary hyphal knots had developed into sclerotia (not shown).

During growth in constant darkness, transcript levels of the endogenous cg/l2 gene increased by about 150-fold from day 3 to day 6 and decreased again between day 6 and day 8 (Fig. 1A). Starting from a similar transcript level as after 3 days of growth in constant darkness, in cultures grown in constant light, cg/l2 transcripts decreased at day 5 by a factor of 10 and stayed at this level until day 8 (Fig. 1A). These results confirm previous observations for which cg/l2 is expressed in C. cinerea cultures grown in constant darkness at the time point of primary hyphal knot formation but not in constant light (Boulianne et al., 2000). Furthermore, they suggest that regulation occurs at the transcriptional level. The results also show that the method developed for transcript level quantification in C. cinerea is indeed useful and can in principle be applied for any other gene of interest.

The time course of transcriptional activity of the reporter gene containing the longest cg/l2 promoter fragment analyzed (1290 bp upstream to the start codon) was similar to that of the endogenous cg/l2 gene, both in constant light and constant darkness (compare Fig. 1A to Fig. 1B). Differences were only observed in the amplitudes of transcription. In constant darkness, the transcriptional induction of the endogenous gene was higher than observed for the reporter gene. In the dark, the gfp transcript levels increased approximately 40-fold from day 3 to day 6 compared to
150-fold in case of *cgl2*. In constant light, as already observed for the endogenous *cgl2*, the levels of the reporter transcript decreased from day 3 to day 6 by a factor of about 10 and remained low during the following incubation. In constant light, the *gfp* transcript levels were always approximately 10 to 50 times higher than the *cgl2* transcript levels. Higher reporter gene transcript levels compared to the endogenous gene were also observed in the dark where, beginning from a maximum (10-fold) at day 3, the difference decreased during further incubation.

We conclude that transgene expression is regulated in the same way as observed for the endogenous *cgl2* allowing us to use the developed experimental set-up to define *cis*-acting elements on the *cgl2* promoter.
Fig. 1 Time course of *cgl2* and *gfp* transcription in constant light and constant darkness. Relative *cgl2* (A) and *gfp* (B) transcript levels in AmutBmut transformed with the p1290 *gfp* construct were determined over a 8-days period in constant light (O) and constant darkness (●) by real-time PCR. Dots represent the average of 2 data points with standard deviation. See Materials and Methods for experimental details.
Dissecting cgl2 promoter regulation: basal activity and induction in constant darkness

To analyze whether the induction of cgl2 transcription in constant darkness between day 3 and day 6 ("dark induction") was dependent on cgl2 promoter elements, we modified the cgl2 promoter in the reporter construct and compared transcript levels between pools of transformants grown in the dark for 3 and 6 days, respectively.

A first set of constructs containing various 5' deletions of the largest promoter fragment analyzed (1290 bp upstream to the start codon; Fig. 1B and 2A: p1290) was tested for induction of transcription. 627 bp upstream to the start codon were sufficient to confer wild-type regulation to the reporter gene (Fig. 2A: p627). Further deletion of the promoter resulted in two phenotypes: the basal activity (transcript levels after 3 days of incubation) dropped by two orders of magnitude and dark induction was no more observed (Fig. 2A: p566, p489 and p376). Interestingly, the construct harboring the 566 bp promoter fragment showed the largest variation in transcriptional activity between different pools, whereas constructs with shorter promoter fragments displayed less variation (Fig. 2A: p489, p376).

To identify specific regulatory regions or elements in the cgl2 promoter sequence, within the 742 bp upstream to the start codon a linker scanning analysis was performed. At eleven different positions, original sequences of either 10 or 20 bp were replaced by an unrelated sequence (see Table 1). Compared to the reference construct p742, the lowest gfp transcript levels were observed for construct p742-5 for which the basal activity was almost 100-fold reduced after 3 days of growth in the dark (Fig. 2B). Nonetheless, this construct revealed the highest dark induction (113-fold). In three other constructs (Fig. 2B: p742-1, -7 and -8), the basal activity was also reduced but not as pronounced as in p742-5. These constructs showed also reduced gfp transcript levels at day 6 (in the dark) as compared to the reference construct p742 (Fig. 2B). Values for constructs p742-2, p742-6, p742-9 and p742-11 were comparable to the reference construct p742, suggesting that in these cases the replacements (Table 1: linkers 2, 6, 9 and 11) did not affect essential promoter sequence motifs. Finally, constructs p742-1, p742-3 and p742-10 (Table 1: linkers 1, 3 and 10) displayed reduced dark induction levels (approximately 7- to 8-fold dark induction; Fig. 2B).
Each replacement sequence in the linker scanning analysis contained a *Hind*III restriction site (Table 1). To further analyze the *cgI2* promoter, an internal deletion analysis was performed over the whole length of the 627 bp promoter fragment, making use of this newly introduced site (see materials and methods for details).

Despite disruption of a putative TATA box by linker 1 (Table 1), deletion of the sequence between linkers 1 and 3 did not seem to really affect the promoter activity [Fig. 2C: p742-(3/1)]. In fact, the transcript levels of construct p742-(3/1) were almost identical to the reference construct p742. Instead, a reduction in basal activity compared to p742 was observed for constructs p742-(4/3) (10-fold reduction), p742-(5/4) and p742-(7/5) (both 60-fold reduction), (Fig. 2C). Nonetheless, in all these constructs the dark induction was still functional. In conclusion, the sequence replaced by linker 5 together with surrounding regions is expected to function in basal promoter activity but appears not to be responsible for dark induction. In contrast, deletion of the promoter sequence between linkers 7 and 11 completely abolished the dark induction [Fig. 2C: p742-(11/7)]. To further characterize this region, additional constructs containing sub-deletions were generated using the unique AatII restriction site in the promoter region and tested for activity [Fig. 2C: p742-(11/AatII), p742-(AatII/7)]. In construct p742-(11/AatII), the dark induction was still functional – although not optimal (9.6-fold) – whereas in p742-(AatII/7) it was not present (3.1-fold; not significant). We conclude that the sequence encompassing the restriction site AatII and linker 7 is necessary for the dark induction.
Fig. 2 (next page): *cg/2* promoter analysis in constant darkness. Schematic view of the promoter constructs used in the 5' deletion analysis (A), linker scanning analysis (B) and internal deletion analysis (C) and corresponding relative *gfp* transcript levels in constant darkness. The bars in the graphs indicate transcript levels (geometric mean ±SD) at day 3 (grey), respectively day 6 (black) post-inoculation. Numbers next to the graphs indicate the fold induction of the transcript levels between days 3 and 6. A schematic overview of the *cg/2* promoter is shown on the left top. The numbered light grey boxes indicate position of linkers used in the linker scanning analysis. The putative transcription start point (arrow), the 5' intron (V), the start of the coding sequence (dark grey box) are also indicated (Boulianne et al., 2000; Charlton et al., 1992). n=5 for each bar in A. n=4 for constructs p742-1, -3, -4, -5, -7 and -11. n=3 for the other constructs.
The numbers indicate the fold-induction of gfp transcript levels between day 3 and day 6.
The effect of light on cg/2 transcription

Constant light inhibits the induction of cg/2 observed in constant darkness and even represses cg/2 transcription upon prolonged cultivation (Fig. 1A). To verify whether this effect was mediated by specific cg/2 promoter elements, the same constructs used for the dark induction assays were tested under constant light conditions and examined for an impairment of light repression. As for constant darkness, the minimal promoter fragment showing both normal basal activity and light repression comprised 627 bp upstream to the start codon (Fig. 3: p627). Pools of transformants containing this promoter fragment (or a longer one) incubated in constant light for 6 days displayed 40 to 70-fold lower gfp transcript levels than the same pool of transformants incubated in constant darkness over the same period. When the used promoter fragment was 566 bp or shorter, there was no significant difference in transcript levels between light- or dark-incubated transformants (Fig. 3: p566, p489 and p376). Moreover, the overall transcript levels of these transformants were around 50 to 100-fold lower than those found when using the minimal promoter or longer promoter sequences. These results confirm the conclusion already deduced from the dark induction assay, that the 566 bp long promoter fragment and shorter fragments are basically inactive (Fig. 2A).

In the analyses of linker scanning and internal deletion constructs, after 3 days of growth the basal transcription of all constructs was affected, in constant light, in the same way as observed in constant darkness (Fig. 2B and C: day 3; data for light not shown). For example, compared to the unmodified reference construct p742, linker 5 was found to significantly reduce the gfp transcript level already after 3 days of growth in the light (not shown) as already observed in the dark (Fig. 2B: p742-5). After 6 days of incubation in constant light, the gfp transcript levels from either the linker-modified or the internally deleted promoter fragments displayed a similar profile as observed after 3 days in constant darkness (Fig. 2B and C). Most importantly, in constant light no construct showed transcriptional induction from day 3 to day 6 post-inoculation (data not shown).

In summary, light repression was abolished in none of the analyzed pools of transformants containing modified versions of the 742 bp promoter fragment. We conclude that the light effect on cg/2 transcription (i.e. inhibition of dark induction and repression of cg/2 transcription upon prolonged incubation in constant light) is either
more complex (e.g. mediated by the combination of several promoter elements) or possibly not mediated by light-specific promoter elements (see discussion).

Fig. 3: *cgl2* promoter analysis in constant light. Bars indicate geometric mean (n=5, ±SD) of relative *gfp* transcript levels in cultures after 6 days of incubation in constant darkness (black) or constant light (white), respectively. Other indications are the same as for Fig. 2. For details refer to the Materials and Methods and Results sections.
5. Discussion

The aim of this work was to establish a method for gene expression analysis at transcriptional level in the homobasidiomycete *Coprinopsis cinerea* based on quantitative real-time PCR technology. The method was first used to analyze the transcriptional regulation of the *cgl2* gene expression during fruiting body formation. In subsequent experiments, a reporter plasmid containing a GFP coding sequence under the control of modified *cgl2* promoter sequences was introduced into *C. cinerea* by transformation. The activity of the various promoter fragments was measured at the transcriptional level by determining *gfp* transcript levels. In contrast to classical approaches, where usually few well-characterized individual transformants are analyzed (e.g. Carattoli et al., 1994; Dzikowska et al., 2003), the data in this study were collected from pools of large numbers of transformants, which were not characterized in detail.

In our analysis, we observed that transcripts of the endogenous *cgl2* gene and transcripts of the *gfp* gene under control of the longest promoter fragment tested (1290 bp) showed a similar regulation over an observation period of 8 days (induction in constant darkness, repression in constant light), whilst the amplitude of induction in constant darkness was approximately 10-fold higher for the endogenous gene than for the reporter construct p1290 (Fig. 1). A possible explanation for this difference is that the reporter gene is missing some upstream or downstream sequences contributing to promoter regulation. Alternatively, the difference may be due to the ectopic DNA integration, typically occurring in *C. cinerea* transformation: the place of integration, the copy number and the completeness of the integrated reporter construct will all influence the transcriptional activity of the reporter construct (Binninger et al., 1987; Mellon and Casselton, 1988; Granado et al., 1997). Finally, a different *gfp* and *cgl2* transcript stability could also account for the observed difference. Nonetheless, the same overall response obtained for the endogenous *cgl2* gene and the p1290 construct under the same conditions indicates that the real-time PCR-based method is suitable to determine relative levels of specific transcripts in *C. cinerea*.

The 5'-deletion analysis of construct p1290 allowed us to define the minimal promoter fragment sufficient to confer basal activity as well as dark induction and
light repression in both constant darkness (Fig. 2A: p627) and light (Fig. 3: p627). Dark induction and light repression were both impaired in constructs with promoter fragments equal or shorter in length than 566 bp upstream to the start codon. Additionally, these constructs displayed a much lower basal activity both in constant darkness (Fig. 2A) and in constant light (Fig. 3). This suggests that the sequence stretching from position -566 to -627 is not only necessary for induction in constant darkness and repression in constant light but also essential for basal activity of the cgI2 promoter.

The most interesting result of the linker scanning analysis of this study comes from construct p742-5. This construct displayed the strongest relative dark induction (113-fold) but displayed a 90-fold lower basal transcriptional activity compared to the unmodified promoter of the p742 construct (Fig. 2B). Linker 5 partially disrupted a site (position -402/-394 on the lagging strand; Fig. 4) matching the consensus binding sequence for the eukaryotic Sp1 factor (consensus sequence: KRGGCKRRK; http://www.gene-regulation.de/), the only one found on the p742 promoter fragment. Sp1 and related proteins are ubiquitously expressed transcription factors of higher eukaryotes known for their unspecific influence on transcription – basal activity in particular – of many different genes (Xu et al., 1997; Antonio et al., 2002). In basidiomycetes, functional Sp1 motifs have already been described for Cryptococcus neoformans (Zhang et al., 1999). A blast search with the Mus musculus Sp1 amino acid sequence (GenBank accession no. CAA42721) against the recently released C. cinerea genome sequence (http://www.broad.mit.edu/annotation/fungi/coprinus_cinereus/) revealed indeed the presence of various sequences potentially coding for similar Cys2His2 zinc finger motifs as the one present in the Sp1 protein sequence (data not shown).

The putative TATA box (-133 to -128 to the ATG startcodon), found at a typical position within the promoter sequence (Gurr et al., 1987), is part of a nearly perfect palindromic sequence (CTTTTtgACGTatAAAAG; Fig. 4). Replacement of promoter sequences at this site by linker 1 caused also a drop in the basal activity but here, additionally, a significant impairment in the dark induction was observed (Fig. 2B: p742-1). When an upstream 100 bp fragment including this site was deleted [Fig. 2C: p742-(3/1)], however, the transcript levels of the corresponding reporter construct were surprisingly the same as found for the reference construct (Fig. 2C: p742). No new TATA box was created by this deletion, suggesting that this modified cgI2
promoter functions without a consensus TATA box, which is not uncommon in fungal genes (Gurr et al., 1987). Moreover, an around 40 bp-long pyrimidine-rich region (74% pyrimidines) directly upstream of the putative TATA box (Fig. 4), respectively of linker 1, does not seem to be essential for promoter activity, in contrast to what has been postulated for similarly localized pyrimidine-rich stretches in other basidiomycetous promoters (Gurr et al., 1987; Yamazaki et al., 2000; Sirand-Pugnet et al., 2003).

Both *C. cinerea* galectin genes are regulated by the A mating type pathway (Boulianne et al., 2000). Construct p742-2 was generated to replace the sequence GATGX_{11}CAA (on the leading strand; Table 1 and Fig. 4) potentially involved in binding of A mating type proteins. The sequence is similar to the *hsg* (haploid specific genes) promoter motif (GATGX_{9}ACA) found in haploid specific genes of the yeast *S. cerevisiae* and known to be the target of the MATα2/MATα1 heterodimer (Goutte and Johnson, 1988; Jin et al., 1999). In the heterobasidiomycete *Ustilago maydis*, a *hsg*-like sequence (GATGX_{9}ACA) was found in the promoter of the *lga2* gene and a closely related motif (GATCX_{9}ACA) in the promoter of the *frb52* gene. In both genes, promoter sequences containing these motifs were shown to be bound by a translational fusion construct of the heterodimeric mating type transcription factor complex bE/bW (Romeis et al., 2000; Brachmann et al., 2001). A similar sequence (GATGX_{11}ACA) is present in the promoter of the A mating type regulated *clp1* gene of *C. cinerea* in close proximity of the TATA box (Inada et al., 2001). Preliminary results suggest that this sequence in the *clp1* promoter is indeed the target of the A gene products (Kamada, personal communication). However, even though this sequence is located in a larger conserved sequence identical for the *clp1*, *cgl2* and also *cgl1* promoters (GGGATGCAAX_{2}AAATX_{4}A, Kües et al., 2004), its replacement in the *cgl2* promoter by an unrelated linker sequence did not influence reporter gene transcription (Fig. 2B: p742-2).

Fig. 4 (next page): *cgl2* promoter sequence from position -743 to -113. Two identical, equal-spaced sequence stretches are outlined by light and grey boxes, respectively. Promoter elements described in the text or in Table 1 are in italics and underlined; CRE sites are indicated additionally by a heavy arrow (see text for details). Direct repeats are indicated by horizontal box-arrows. Inverted repeats are indicated by two-headed line-arrows. Sequences replaced by linkers are underlined by heavy lines. 5' promoter ends of p742, p627, p566, p489 and p376 constructs are indicated. See text for additional details.
The most informative analysis with regard to the cg/l2 promoter elements was the internal deletion analysis. Not surprisingly, deletions involving linker 5 (discussed above) showed the same drop in gfp transcript levels as caused by linker 5 alone [Fig. 2C: p742-(7/5) and p742-(5/4)]. Construct p742-(4/3), despite maintaining a relative normal dark induction, displayed a marked reduction in overall activity (around 10-fold lower) compared to the reference construct p742 (Fig. 2C). Since not displayed by the corresponding linker constructs (Fig. 2B: p742-3 and p742-4), this result suggests the presence of basal promoter elements between linkers 4 and 3. A BLAST-search with the Transfac database, however, did not indicate any particular sequence motif in this region. It is possible that deletion of this sequence-stretch is influencing transcriptional activity more generally by modifying the chromatin structure of the promoter or by altering the spacing between different regulatory sites. Alternatively the sequence contains a not yet identified motif.

The most remarkable effect on promoter activity was observed when deleting the sequence between linkers 11 and 7 in construct p742-(11/7). This deletion completely abolished the dark induction [Fig. 2C: p742-(11/7)]. Generation of sub-deletions between linkers 11 and 7 suggests that the sequence responsible for the dark induction is probably located between the Aatll restriction site and linker 7 [Fig. 2C: p742-(Aatll/7)]. It seems however that this sequence does not act alone. In fact, construct p566, despite harboring the sequence between linker 7 and the Aatll restriction site, did not display dark induction (Fig. 2A: p566). Addition of further 61 bp upstream of the 566 promoter fragment, however, was enough to restore both basal activity and dark induction (Fig. 2A: p627). The same effect was observed when adding the sequence between position -742 and linker 11 upstream to the Aatll restriction site [Fig. 2C: p742-(11/Aatll)]. The comparison of these two alternative sequences revealed a high degree of redundancy and the presence, in both sequences, of a CRE site (cyclic AMP response element, Conkright et al., 2003) with the consensus sequence TGCGTCA (Fig. 4). The distal CRE site lies on the lagging strand at positions -739 to -733 to the start codon and the proximal one on the leading strand between -610 and -604 (Fig. 4). In fungi, cyclic AMP (cAMP) functions as a secondary messenger mediating intracellular signal transduction (D'Souza and Heitman, 2001; Lee et al., 2003).

In C. cinerea, cAMP has repeatedly been implicated in fruiting body development (Kües et al., 2004). For example, exogenously added cAMP was able to induce
fruiting in certain strains of *C. cinerea* (Uno and Ishikawa, 1982). cAMP action in *C. cinerea* is influenced by nutrients, light and the combined action of the mating type products (Uno and Ishikawa, 1982; Swamy et al., 1985) as is initiation of fruiting body development (Kües et al., 2004). Increasing levels of cAMP were observed in fruiting-competent mycelia growing on media containing 0.4% glucose whilst higher glucose concentrations had an inhibiting effect both on increasing cAMP levels and fruiting (Uno and Ishikawa, 1982; Boulianne et al., 2000). cAMP concentrations were found to be higher in secondary hyphal knots and primordia than in the mycelium and the timing of initiation of fruiting was approximately the same or slightly delayed relative to the observed peak levels of cAMP (Uno and Ishikawa, 1982; Swamy et al., 1985; Kuhad et al., 1987; Kües et al., 2004). cAMP exerts its function by activating cAMP-dependent protein kinases. Consistent with this, cAMP-dependent protein kinase activity was observed only in fruiting competent mycelia (Uno and Ishikawa, 1982; Swamy et al., 1984; Swamy et al., 1985). Taken together, these results strongly suggest a direct involvement of cAMP in the fruiting process. Galectin expression correlates in time and space with fruiting body formation and, like this developmental process, is inhibited by high nutrient content in the medium (Boulianne et al., 2000; Bottoli, 2001). Our results suggest that the putative CRE sites in the *cg/2* promoter, partially replaceable by each other, could be responsible for basal transcriptional activity (exerted by all constructs containing one or both CRE sites) and, together with the sequence between the AatII restriction site and linker 7, for induction of the *cg/2* gene in the dark. Interestingly, in the sequence between the AatII restriction site and linker 7 three direct repeats can be found on the leading strand (2x TGGAAG, 1x GGAA; Fig. 4). Consistent with a possible involvement of direct repeats, CRE-binding proteins are known to exert their action by interacting with other proteins, some of which bind DNA at direct repeats. Moreover, CRE-binding factors together with their interacting partners were shown in many cases to be responsible for both basal and cAMP-dependent gene transcription (Swanson et al., 1997; Swanson et al., 2000; Yang et al., 2001; Adachi and Lewis, 2002). Linker 8 disrupted the third GGAA repeat (Fig. 4) and, supporting a possible involvement of this site in *cg/2* transcriptional regulation in constant darkness, p742-8 displayed a slightly reduced dark induction, together with a nearly normal basal activity (Fig. 2B). In accordance with an interaction between CRE- and direct repeats-binding proteins, construct p566 displayed much lower and more variable transcript levels than constructs containing
shorter promoter fragments (Fig. 2A and Fig. 3). p566 still contains the direct repeats but is missing both CRE sites (Fig. 4). It is possible that proteins binding alone to the direct repeats without interaction with the CRE-binding proteins are somehow disturbing transcriptional activity. Accordingly, promoter fragments missing both CRE and direct repeats display higher and more stable reporter transcript levels than p566, even though they do not reach basal activity (Fig. 2A and 3: p489 and p376).

Interestingly, no CRE motif was found in the promoter of the cgl1 gene which is induced at a later stage of fruiting body development. Instead, the same direct repeats as the ones found in the cgl2 promoter (GGAA and TGGAAG) are present at a corresponding position in the cgl1 promoter. In agreement with a control by cAMP, the cgl2 gene is induced when nutrients are exhausted and primary hyphal knots form (Boulianne et al., 2000; Bottoli, 2001). The cgl1 gene, on the other hand, is induced only upon formation of secondary hyphal knots (Boulianne et al., 2000) when the cAMP peak is over (Swamy et al., 1985; Kües et al., 2004) and may therefore not need CRE elements in its promoter.

During the fruiting process in C. cinerea, cAMP metabolism is closely linked to mating type and light regulation (Swamy et al., 1985; Kües et al., 2004). Striking in this study is that we were neither able to define any motif on the cgl2 promoter responsible for A mating type gene regulation nor for light regulation. Therefore, regulation of transcription by A mating type products and by light might not be direct but may work by modulating the activity of transcription factors, as those responsible for cgl2 induction in the dark. Crosstalk between mating type and cAMP/MAPK regulated development was shown recently in U. maydis to act via a unique factor, Prf1, which, depending on its phosphorylation state, is able to activate b or a mating type genes (Kaffarnik et al., 2003). Similar mechanisms might be also responsible for the transcriptional regulation of the cgl2 gene in C. cinerea.
Acknowledgments

We thank T. Kamada for communication of unpublished results, F. Bootz (Institute of Laboratory Animal Science, University of Zurich) and R. Gay (Department of Rheumatology, University Hospital Zurich) for kindly making available the ABI Prism 7700 Sequence Detectors and M. Gut of Applied Biosystems (Rotkreuz, Switzerland) for continuous support in real-time PCR technology. This work was funded by the ETH Zurich. U.K. is supported by the Deutsche Bundesstiftung Umwelt (DBU).
6. References


7. Annex

7.1. Validation of RNA quality

As reported in materials and methods, RNA quality was verified according to the 260/280 nm absorption ratio and, visually, by size-separation of RNA on formaldehyde-agarose gel. In Fig. 5 a typical RNA formaldehyde-agarose gel is shown to illustrate RNA acceptance limits.

Fig. 5: Typical formaldehyde-agarose RNA gel. The picture shows total RNA isolated from mycelial C. cinerea samples after 6 days of growth in the dark and separated according to size in a formaldehyde-agarose gel (stained with ethidium bromide). 10 μg RNA per lane were loaded. The upper band corresponds to 28S, the lower to 18S ribosomal RNA, respectively. Samples showing partial (A: 18S rRNA band appears stronger than 28S rRNA band) or complete (B) RNA degradation were discarded. Intact samples (C and D: stronger stained 28S than 18S rRNA band) were used for cDNA synthesis.
7.2. Validation of real-time PCR assays

In real-time PCR assays, standard curves are prepared to verify amplification efficiency and detection limit of the assay for every particular target sequence. Efficiency is derived from the idealized function for the amount of PCR product formed: \( N = N_0 \times E^c \), where \( N \) is number of amplified molecules, \( N_0 \) is the initial number of molecules, \( c \) is the number of amplification cycles and \( E \) is the exponential amplification value. At an efficiency of 100% (\( E=2 \)), the template doubles after each cycle during exponential amplification. By plotting the log of the concentration of a prepared dilution series of the desired target DNA versus the corresponding \( C_T \) values, the data points will result in a linear correlation with function \( c = -(1/\log E) \times \log N_0 + (\log N / \log E) \). Therefore the slope of the standard curve can be used to determine the exponential amplification value and efficiency of the PCR reaction by the following equations:

\[
\text{Exponential amplification value (E) } = 10^{-1/\text{slope}}
\]

\[
\text{Efficiency } = E - 1
\]

Validation curves were prepared using cDNA in a 1:10 dilution series. Two measurements were made per dilution point. The results are reported in the following graph where linear regression curves are shown for every gene analyzed (\textit{benA, cgl2, gfp}).
**Chapter II**

### Standard curves

![Standard curves graph](image)

**Efficiency parameters:**

**benA:**
- Correlation coefficient ($r^2$): 0.999
- Slope: -3.80
- $E$: 1.83
- Efficiency: 0.83

**gfp:**
- Correlation coefficient ($r^2$): 0.998
- Slope: -3.823
- $E$: 1.83
- Efficiency: 0.83

**cgl2:**
- Correlation coefficient ($r^2$): 0.999
- Slope: -3.57
- $E$: 1.91
- Efficiency: 0.91
7.3. Addendum to graphs presented in the “results” section

*gfp* transcript levels of transformant pools incubated in constant darkness and shown in Fig. 2 in the “results” section were transformed according to the *cg/l*2 value in the respective sample (see “material and methods” section for details). Original data (normalized only to *benA*) for the endogenous *cg/l*2 and for the *gfp* reporter gene transcripts are given in the following (Fig. 6 and 7). For the 5'-deletion analysis, in addition to the data obtained for days 3 and 6, measurements made after 8 days of growth in constant darkness are included in the graphs (Fig. 6 and 7: dark-grey bars in panel A).

Further to the values obtained after incubation in constant darkness, data gained after incubation of transformants in constant light are reported in Fig. 8 for *cg/l*2 and Fig. 9 for *gfp*, respectively.
Fig. 6: *cgl2* transcript levels in constant darkness (normalized to *benA*).
Fig. 7: gfp transcript levels in constant darkness (normalized to benA).
Fig. 8 cgl2 transcript levels in constant light (normalized to benA).
Fig. 9: gfp transcript levels in constant light (normalized to benA)
CHAPTER III

Transcriptional regulation of cg/l2 and of gfp reporter gene constructs in the monokaryotic C. cinerea strain FA2222

1. Summary

In previous analyses, transcriptional regulation at the promoter level of the cg/l2 gene was investigated in vivo in a self-compatible homokaryon (chapter II). Here, first data on the regulation of cg/l2 transcription in a monokaryotic background are presented and discussed.

2. Introduction

Mating type genes are considered to be key regulators of sexual events in mushrooms (Casselton and Olesnicky, 1998; Moore, 1998; Kües and Liu, 2000). Fruiting bodies contain the sexually produced basidiospores and develop usually only on the dikaryotic mycelium. Since the mating type genes regulate dikaryon formation, these genes ultimately also control fruiting body initiation (Kües et al., 1998; Kües and Liu, 2000; Kües et al., 2002). Even in case other key parameters (i.e. nutrients and light) are favorable, no fruiting body development is observed on wild-type monokaryons, in contrast to strains carrying specific defects in the mating type loci and being therefore self-compatible (Swamy et al., 1984).

In the self-compatible homokaryotic strain Am/Bmut of Coprinopsis cinerea, the β-galactoside binding lectin (galectin) CGL2 is expressed at the onset of fruiting body development (primary hyphal knot formation). Like primary hyphal knots, cg/l2 expression is inhibited by light and induced in the dark upon nutrient depletion (Boulianne et al., 2000). Moreover, as observed in monokaryotic strains transformed with A mating type genes of compatible specificity, activation of the A mating type pathway induces cg/l2 expression as well as primary hyphal knot formation. Some
strains, such as monokaryon 218, show very low expression of CGL2 and low production of primary hyphal knots independently of activation of the A mating type pathway. Others, such as monokaryon FA2222, produce CGL2 and primary hyphal knots only when transformed with compatible A mating type genes (Kües et al., 1998; Boulianne et al., 2000).

To identify promoter elements responsible for the transcriptional regulation of the cgl2 gene in a dikaryotic background, gfp reporter constructs containing modified versions of the cgl2 promoter have previously been transformed into homokaryon AmutBmut and gfp transcription been followed by real-time PCR (see chapter II). The analysis defined a minimal promoter fragment of 627 bp containing putative binding sites for a Sp1-type transcription factor as well as a CRE-like motif. To further investigate transcriptional regulation of cgl2 in a monokaryotic background, in this chapter we analyzed the transcription of both the endogenous cgl2 gene and of gfp reporter constructs in monokaryon FA2222 using the formerly developed detection method based on real-time PCR. In spite of the high detection power of the method, endogenous cgl2 transcription was not at all detectable, confirming former less sensitive Western blot results (Boulianne et al., 2000). In all groups of analyzed gfp transformants, low to very low gfp transcript levels were detected regardless of deletions in the cgl2 promoter sequence. Further preliminary data on gfp transcript levels suggest that, in the monokaryon, transcriptional repression of the cgl2 gene could be also mediated by discrete promoter sequences. On the other hand, deletion of promoter stretches shown to be crucial for transcriptional activity in the homokaryotic background AmutBmut, caused in monokaryon FA2222 a stronger impairment on transcription than in the homokaryon.

3. Materials and methods

The overall method for transcription analysis in the monokaryotic strain FA2222 (A5, B6, acu-1, trp-1.1, 1.6 Kües et al., 1998) was the same as described in chapter II. Plasmids used for transformation containing the gfp coding sequence under the control of integral or modified cgl2 promoter sequences were described in chapter II. To select positive transformants from monokaryon FA2222, reporter gene plasmids were co-transformed into C. cinerea protoplasts together with plasmid pCc1001, that contains the tryptophan synthetase gene (trp-1) of C. cinerea (Binninger et al., 1987).
Co-transformation and cultivation procedures were the same as previously described (chapter II). A total of 120 different FA2222 transformants per transformed construct were analyzed. Mycelia from 4 different plates with 10 different transformants each were harvested and pooled together, giving a total of 3 pools a 40 transformants each. Because of the highly invasive growth of monokaryon FA2222, a sterile cellophane disc (8 cm Ø) was placed before inoculation on the medium in the Petri-dishes (YMG/T agar as described by Granado et al., 1997). This prevented the mycelium to grow into the substrate and reduced mycelial loss when harvesting. Due to the slow and irregular growth of FA2222 transformants on YMG/T complete medium, mycelia were collected from the plates after 4 days of growth at 37°C. This growth stage roughly corresponded to 3-days-old AmutBmut mycelia analyzed in chapter II (Fig. 1).

RNA extraction, real-time PCR assay and statistical analyses were done exactly as previously described (chapter II).

It was observed that, although maintaining the same physical conditions (light and temperature), identical transformed pools displayed consistent differences in cg/I2 and gfp transcript levels between different experimental sessions (e.g. chapter II, Fig. 6: compare cg/I2 transcript levels in panel B with levels in panel C). Thus, the direct comparison of gfp transcript levels obtained from different experimental sessions would lead to misinterpretations. A formula that took the observed differences between different experimental sessions was used to transform the gfp values when comparing the gfp data (Fig. 2B). In short, holding the same conditions between two experimental sessions (same transformants pool, same physical conditions), to obtain the corrected gfp value (gxc), the gfp value of the experiment to be compared (A) is multiplied by a factor obtained by dividing the average cg/I2 transcript level in the reference experiment (B) by the average cg/I2 transcript level in the experiment A. This is expressed by the formula:

\[ g_{xc} = g_{xA} \times \frac{C_B}{C_A} \]

where:
- \( g_{xA} = \) gfp transcript levels of pool x in experiment A;
- \( C_A = \) average cg/I2 transcript level in experiment to be compared (A);
- \( C_B = \) average cg/I2 transcript level in reference experiment (B);
- \( g_{xc} = \) corrected gfp transcript level of pool x.
4. Results and discussion

Previous results indicated $cg/l2$ expression – like primary hyphal knot formation – to be under the influence of the $A$ mating type genes. Accordingly, significant amounts of $CGL2$ are detected in the homokaryotic strain AmutBmut (self-compatible for mutations in the mating type loci) and in strains in which the $A$ mating type pathway was activated by transformation with mating type genes of compatible specificity (Boulianne et al., 2000). Transcriptional regulation of $cg/l2$ and $gfp$ under control of $cg/l2$ promoter sequences was analyzed in the monokaryotic strain FA2222 making use of the experimental set up and promoter deletion constructs described in chapter II. Transcripts levels were determined after 4 days of incubation in constant darkness.

$cgl2$ transcripts are not detectable in the monokaryon

Fig. 2A compares the $cg/l2$ transcript levels in pools of AmutBmut transformants grown for 3 days in the dark with the corresponding situation (growth of 4 days in the dark) in monokaryon FA2222. In homokaryon AmutBmut, $cg/l2$ transcripts were detected in all the analyzed pools at comparable levels, irrespectively of which $gfp$ construct had been transformed. In contrast, in the monokaryotic background, $cg/l2$ transcripts were not at all detectable within the limits of detection of the method (relative transcription value of $10^{-6}$ compared to the transcript levels of the house-
keeping gene β-tubulin arbitrarily set at 1). This suggests that, in the absence of active mating type gene products, the cgl2 gene is not at all transcribed, as found previously through Western blot analyses (Boulianne et al., 2000). This result stresses the central role played by the mating type genes as master regulators in fruiting events: unless their activation upon compatible matings, the transcription of downstream genes such as cgl2 is completely silent, at least in monokaryon FA2222.

The reporter gene is not completely silent in the monokaryotic background

Different from cgl2, transcription of the gfp reporter gene in pools of transformants of monokaryon FA2222 was always detectable. gfp transcript levels were however only low to very low and displayed mostly a high variability between values from different pools transformed with the same construct. Noteworthy gfp transcript levels were only registered for constructs p742 and p742-(5/4) that, however, did not reach the corresponding levels of homokaryon AmutBmut (Fig. 2B). In the monokaryotic background, the only statistically significant difference was observed when comparing gfp transcript levels of the reference construct p742 either with construct p742-(7/5) or with p742-(11/7), whose respective transformants produced the lowest gfp transcript levels observed. In the case of p742-(11/7), however, the difference was at the limit of the statistical significance (P=0.059). The differences of transcript levels in constructs p742-(7/5) and p742-(11/7) to construct p742-(5/4) were not statistically relevant, albeit Fig. 2B may have suggested such a distinction.

From an overall comparison with homokaryon AmutBmut, on the whole it has to be concluded that the cgl2 promoter sequences in the gfp constructs tested in monokaryon FA2222 are rather silent in transcription than active. Nevertheless, despite the high variability observed in some pools of FA2222 transformants, some cautious thoughts on the regulation of the cgl2 promoter in a monokaryotic background might be made. Indeed, the high variability in transcript levels observed for some pools could be even caused by modification of defined promoter elements.

1. Emphasising chromatin as the principal responsible for the observed repression of the cgl2 gene in the monokaryon, detection of gfp transcripts in the monokaryotic background - in contrast to the endogenous cgl2 gene - might be explained by the ectopic integration of the reporter construct in the genome of the
transformants: definite conditions for complete silencing are likely not meet at most of the ectopic integration sites. If this would be the sole accountable mechanism, the same \textit{gfp} levels were to be expected for all transformant pools, irrespectively of the construct they contain. Whenever present, chromatin silencing alone can not explain the statistically differences observed in \textit{gfp} transcript levels in the monokaryotic background. Why are, for instance, \textit{gfp} transcript levels for constructs p742-(11/7) and p742-(7/5) significantly lower than for p742?

2. Since construct p1290 gave much lower transcript levels in monokaryon FA2222 than the shorter construct p742 (Fig. 2B), important silencing sites could reside in the promoter sequence between positions -1290 and -742. Another silencing site could be located between linkers 5 and 4, since deletion of the promoter sequence between these two linkers seems partly to remove the inhibition of \textit{gfp} transcription [Fig. 2B: p742-(5/4)]. If existent, this silencer could also account for some of the higher variability observed in p742 transformant pools (Fig. 2B: p742). Giving credit to silencers as the responsible for part of the inhibition observed in \textit{cg/2} transcription in the monokaryotic background, the presence of such elements at the original \textit{cg/2} genomic site but missing in the \textit{gfp} constructs used for the analyses, could partly explain the higher reporter gene transcripts observed.

3. \textit{gfp} transcript levels for constructs p742-(11/7) and p742-(7/5) were significantly lower than for the reference construct p742. In the homokaryotic background AmutBmut, key elements responsible for \textit{gfp} transcription were identified in the sequence between linkers 11 and 7 (chapter II). The same region, and therefore possibly the same elements, seems to play a relevant role also in the monokaryotic background: its elimination from the 742 bp promoter fragment markedly reduced transcription in the corresponding construct [Fig. 2B: p742-(11/7)]. Likewise, important promoter elements might be expected in the promoter sequence between linkers 7 and 5, as p742-(7/5) produced similar \textit{gfp} transcript levels as p742-(11/7). At present, however, there is no reasonable speculation regarding their nature.
Fig. 2: cgl2 promoter analysis in monokaryon FA2222. The picture shows the comparison of transcriptional activities of the endogenous gene cgl2 (A) and the reporter construct gfp (B) in pools of transformants of monokaryon FA2222 (grey bars) and homokaryon AmutBmut (black bars) grown for 3 days in constant darkness. A schematic overview of the cgl2 promoter is shown at the top on the left side. The numbered light grey boxes indicate positions of linkers used in the linker scanning analysis (for details see Chapter II). The putative transcription start point (arrow), the 5' intron (V) and the start of the coding sequence (dark grey box) are also indicated (Boulianne et al., 2000; Charlton et al., 1992). Constructs with modified or unmodified cgl2 promoter sequences that were tested in this analysis are shown at the left side of panels A and B. Transcript levels (geometric mean ±SD) are indicated by black (homokaryon AmutBmut) and grey (monokaryon FA2222) bars, respectively. n=5 for p1290; n=3 for all other constructs.
5. References


CHAPTER IV

A model for cgl2 transcriptional regulation

The analysis of the cgl2 promoter in a self-compatible Amut Bmut homokaryon (chapter II) identified two promoter stretches as major regulators of cgl2 transcription in the situation of activated mating type pathways. The first stretch – about 60 bp in length – contains a CRE-like motif (TGCGTCA), the second, directly downstream of the first stretch, harbors three direct repeats of the sequence GGAA (respectively, two of them are longer with the sequence TGGAAG; chapter II, Fig. 4). Their role, as known so far, can be summarized as follows: in the situation of activated mating type pathways,

1. The CRE-containing sequence appears indispensable for adequate basal transcription of a reporter gene construct.

2. The promoter region encompassing the direct repeats is necessary for transcriptional induction of the reporter gene from day 3 to day 6.

A combined contribution of the CRE motif together with direct repeats as described here in the transcriptional regulation of the cgl2 gene of C. cinerea, is reminiscent of some non-fungal systems (of mainly animal origin). In the non-fungal systems, CRE-binding proteins have repeatedly been shown to control both constitutive and cAMP-mediated transcription by interacting with proteins binding at adjacent sites. Such interaction with other promoter-binding proteins seems to be required to restrict the unspecific and more general cAMP stimulation to specific tissues or cells. A few examples of many more available are presented in the following: a CRE-like site (TGACGTAG) and a TACGAC direct repeat were shown to be essential for basal and induced transcription, respectively, of the human dihydrolipoamide dehydrogenase gene (DLD). Despite CREB seems to bind to the CRE-like site, neither this nor the direct repeat site appeared able to respond to cAMP-stimulation, whose regulation is believed to function via transcription factors interacting with CREB and binding at other sites in the DLD promoter (Yang et al., 2001). In the promoter sequence of the rat tyrosine aminotransferase gene (TAT),
the presence of a CRE site is not enough for cAMP-mediated response. This requires an additional HNF-4 (hepatic nuclear factor 4)-responding site. As HNF-4 expression is restricted to liver cells, the presumed function of the HNF-4-responding site on the TAT promoter is to confine the more general hormone-induced cAMP-stimulation to hepatic cells. Indeed, mutation in the HNF-4 site eliminated the ability of cAMP to stimulate TAT expression in liver cells (Nitsch et al., 1993). As a last example, in the promoter of the dopamine β-hydroxylase (DBH) gene, a distinct DNA region of around 30 base pairs (termed DB1) mediates the response to basal, cAMP and cell-type specific signals. A CRE-like motif identified in the DB1 sequence is necessary for PKA-dependent response of the DBH gene. Cell-type specific signals are mediated by the homeodomain protein Arix which, on the DBH promoter, binds at three different ATTA motifs, two of which localize in the DB1 promoter region (Swanson et al., 1997). Although all three ATTA sites are needed for basal response of the DBH gene, different roles for the single ATTA motifs were recognized. Interaction between the two ATTA sites in the DB1 region (immediately downstream to the CRE-like motif) and Arix is necessary for maximal response to PKA-stimulation, whereas the third ATTA site, located more downstream in the promoter, is implicated in general promoter activation (Swanson et al., 2000). Arix was additionally shown to interact with the transcriptional co-activator CBP (CREB binding protein), that enhances transcription from the DBH promoter in a PKA-dependent manner (Swanson et al., 2000). At the basal level, Arix is constitutively phosphorylated in vivo. PKA-dependent dephosphorylation of Arix appears to be necessary for DNA-binding, indicating the ability of Arix to modulate DBH transcription by integrating signals through post-translational modification (Adachi and Lewis, 2002). Arix is thus believed to be the tissue-specific activator of the DBH gene: without binding of Arix to the DBH promoter (after PKA-dependent dephosphorylation), cAMP stimulation is prevented to fully activate DBH transcription.

Considering the few and, partly, preliminary results obtained in this work on cg/l2, it might seem premature and possibly presumptuous to propose a model for the transcriptional regulation of the cg/l2 gene. Nevertheless, to provide a starting point for further studies on the transcriptional regulation of galectin genes in C. cinerea, the formulation of a working model is attempted. Inspired by the above-cited studies on animal promoters, in the proposed model few elements are considered to be essential for transcriptional regulation of the C. cinerea cg/l2 gene. These are the
CRE motif, the GGAA direct repeats (respectively TGGAAG repeats) and the Sp1-like motif (see chapter II). The latter is considered to act as enhancer, since its presence in the cg/l2 promoter seems to enhance the effects on transcription produced by other, more specific cis-acting sites (see results in chapter II). In a dikaryotic background, the activation of the A mating type pathway sets the prerequisite for the induction of the cg/l2 gene by a yet unknown mechanism, that probably does not require specific (A mating type-responding) promoter elements. In the situation of activated mating type pathways, when sufficient nutrients are available, CRE, the GGAA repeats and the Sp1-like motif are all required to provide a respectable basal transcription, no matter whether the mycelium is growing in the dark or in the light. When nutrients become depleted, cAMP concentration in the mycelium increases (Uno and Ishikawa, 1982; Kües et al., 2004) and this may in consequence activate proteins binding at the direct repeats. Protein binding enables them to interact with other proteins binding at the CRE motif. This interaction would provide the mechanism by which cg/l2 transcription is increased in the dark. Swamy and co-workers (1985) showed that a detectable increase in cAMP concentration in the mycelium is dependent both on activation of mating type pathways and on light stimulation. In constant light, one might expect therefore best cAMP-mediated stimulation. In contrast, some unknown mechanism hinders the increase of cg/l2 transcription. This might be mediated via (further) modification (e.g. phosphorylation or dephosphorylation) of the same DNA-binding proteins responsible for the activation in the dark. Light could for instance cause the phosphorylation of proteins binding at the GGAA direct repeats or/and at the CRE site, preventing thus their mutual interaction, otherwise required for transcriptional induction in the dark.

Explaining the regulation of cg/l2 transcription in a monokaryotic background represents pure speculation. As indicated already above, the missing activation of cg/l2 transcription in monokaryons might be explained by two mechanisms. The first assumes that the genomic environment (chromatin packaging) is the cause for the silencing of the cg/l2 gene in the monokaryon. The second connects the missing activation in a monokaryotic background to the same proteins responsible for induction in the dikaryotic background: they may be either not present at all or, if present, rather in an inactive form. Of course, both mechanisms are not mutually exclusive. Indeed, as discussed in chapter III, the preliminary data available can be explained either by the one or the other system or both.
The cg/2 coding sequence starts 1367 bp after the stop codon of the cg/1 coding sequence and the high identity found between the genes hints for a gene duplication happened during evolution (Boulianne et al., 2000). Given the conservation of large promoter stretches, it is reasonable to expect similar expression patterns for both galectins, either in time or in space. Indeed, expression of both genes in C. cinerea is restricted to the first fruiting stages, although at slightly different times (Boulianne et al., 2000). It is interesting to note that nowhere in the cg/1 promoter sequence (analyzed were 1720 bp upstream to the start codon) CRE-like motifs were found, unlike to the cg/2 promoter that has two such sites (chapter II). On the other hand, in the cg/1 promoter and not in the cg/2 promoter, motifs resembling known light-responding elements from other fungi are present (see chapter I, Ballario and Macino, 1997; Froehlich et al., 2002). The most important of these are LRE-like motifs aligned in a similar fashion as found in light-regulated N. crassa genes (Froehlich et al., 2002). In the genomic cg/1 sequence (Boulianne et al., 2000) two tandemly arranged pairs of LRE-like motifs can be found on the lagging strand. The first LRE-like motif of the most distal pair (referred to the translational start codon ATG) begins with the sequence CGATTGGC and starts at position -1047 from the transcriptional start codon (ATG); the immediately following LRE-like motif starts at position -1069 and begins with sequence CGATTATCCG. The first LRE-like motif of the proximal LRE pair has the sequence CGATGCGCA and begins at position -987 and the second LRE-like motif starts with the sequence CGATCACC and begins at position -1003. Additionally, as found in the cg/2 promoter, different groups of GGAA (partly TGGAAG) direct repeats are also present in the cg/1 promoter. The first group is located at position -1405 to -1387 from the ATG start codon and contains two GGAA motifs. The second (-542 to -472), which corresponds in position to the three repeats found in the cg/2 promoter (position -552 to -520), contains five such elements. cg/2 transcripts appear early in development, when primary hyphal knots are formed, apparently upon nutrient depletion. cg/1 expression starts later, possibly as a consequence of light stimulation (Boulianne et al., 2000). Nevertheless, both proteins may be expressed in the same tissues within the developing primordia (Walser et al., 2004b). Taking in account these data, one can imagine the CRE element in cg/2 to cause its expression upon nutrient depletion and the potential light-responding motifs in cg/1 to be responsible for its activation upon light stimulation. This gives an explanation for the different temporal expression of the two galectins.
The GGAA direct repeats, on the other hand, being present in both promoters, might explain the similar spatial expression confined to specific cells (in particular outer stipe and outer cap tissues) of the developing primordia (Boulianne et al., 2000; Walser et al., 2004a).
References


Seite Leer / Blank leaf
Curriculum vitae

Name: Bertossa
First name(s): Rinaldo, Camillo
Date of birth, place: February 22, 1972, Lugano (Switzerland)

Address: Cäcilienstrasse 5 Tel: +41 1 252 85 15
8032 Zurich Fax: +41 1 252 85 57
Switzerland e-mail: bertossa@micro.biol.ethz.ch

Education

1999 – 2004 Ph.D. student at the Institute of Microbiology, Swiss Federal Institute of Technology (ETH), Zurich.
Supervisors: Prof. Dr. M. Aebi, Prof. Dr. U. Kües, Dr. M. Künzler.


Title of Diploma thesis: “Induction of volatile emissions in maize by three herbivore species with different feeding habits”.
Supervisors: Prof. Dr. S. Dorn and Dr. T. C. J. Turlings.

1997 – 1991 High school “Liceo Lugano 1”, Lugano (Switzerland).

1983 – 1987 Secondary school, Barbengo (Switzerland).

1978 – 1983 Primary school, Pazzallo (Switzerland).

Publications


Scientific conference presentations


R. C. Bertossa, M. Künzler, U. Kües, M. Aebi: Promoter analysis in the basidiomycete Coprinus cinereus by Real Time PCR. 4th D-Biol Symposium, Davos, Switzerland, 05.08.-05.10.2002. [Poster]