The EMBO Journal vol.15 no.7 pp.1650–1657, 1996

White collar-1, a central regulator of blue light responses in *Neurospora*, is a zinc finger protein

Paola Ballario[1], Paola Vittorioso[2], Armando Magrelli[2], Claudio Talora, Andrea Cabibbo[2] and Giuseppe Macino[1,2]

Dipartimento di Genetica e Biologia Molecolare, Centro di Studio per gli Acidi Nucleici, Piazzale A. Moro 5, University of Rome 'La Sapienza', 00185 Rome and [2]Dipartimento di Biopatologia Umana, Policlinico Umberto I, Viale Regina Margherita 324, University of Rome 'La Sapienza', 00161 Rome, Italy

[1]Corresponding authors

The *Neurospora crassa* blind mutant white collar-1 (*wc-1*) is pleiotropically defective in all blue light-induced phenomena, establishing a role for the *wc-1* gene product in the signal transduction pathway. We report the cloning of the *wc-1* gene isolated by chromosome walking and mutant complementation. The elucidation of the *wc-1* gene product provides a key piece of the blue light signal transduction puzzle. The *wc-1* gene encodes a 125 kDa protein whose encoded motifs include a single class four, zinc finger DNA binding domain and a glutamine-rich putative transcription activation domain. We demonstrate that the *wc-1* zinc finger domain, expressed in *Escherichia coli*, is able to bind specifically to the promoter of a blue light-regulated gene of *Neurospora* using an *in vitro* gel retardation assay. Furthermore, we show that *wc-1* gene expression is autoregulated and is transcriptionally induced by blue light irradiation.

Keywords: blue light/*Neurospora crassa*/photomorphogenesis/zinc finger domain

Introduction

Light exerts its influence on living organisms not only as an energy donor in photosynthesis, but also as a differentiation signal involved in photomorphogenesis in plants and microorganisms such as fungi. Responses to light in these organisms are mediated by two major classes of receptors which each perceive either red or blue light. Phytochromes, the red/far-red adsorbing light receptors encoded by a family of genes, are to date the only class of photoreceptors characterized biochemically (Quail, 1991). Blue and UV light receptors, generally known as cryptochromes, are still elusive (Short and Briggs, 1994), although a putative blue light photoreceptor, the product of the HY4 locus, has been described in *Arabidopsis thaliana* (Ahmad and Cashmore, 1993). This putative receptor is a soluble flavoprotein, with a region of striking homology to light-activated microbial DNA photolyases and a C-terminal region with homology to tropomyosin A. Still unresolved is how the light-induced conformational changes of these photoreceptors are transduced to signal transduction pathways. In addition, the cellular components of these light signaling pathways are also unknown. In attempts to identify these cellular components, an effective genetic approach has been to screen for mutants that exhibit the phenotype of photoreceptor mutants, despite the fact that they possess normal receptors. This approach has been applied in plants, where it has resulted in the isolation of a number of putative steps involved in light signal transduction and in the definition of their hierarchical relationships (Chory, 1993). Well-characterized signaling molecules, such as G-proteins and calmodulin, have been suggested to be involved in light-induced cellular signaling (Neuhaus *et al.*, 1993). However, in plants, the situation seems to be extremely complex: families of photoreceptors specific for different photomorphogenetic responses control signal transduction pathways characterized by common components and specific branches (Neuhaus *et al.*, 1993; Elich and Chory, 1994).

Neurospora crassa exhibits only blue light-regulated photomorphogenesis. Under blue light irradiation, *N.crassa* responds with the induction of mycelial carotenogenesis (Harding and Turner, 1981; Nelson *et al.*, 1989; Schimdhauser *et al.*, 1990), promotion of conidia (Lauter and Russo, 1991) and protoperitecia development (Degli Innocenti and Russo, 1984), phototropism of peritecial beaks (Harding and Melles, 1984) and inhibition or shifts in the circadian rhythm of conidiation (Sargent and Briggs, 1967; Arpaia *et al.*, 1995). All of these light-induced phenomena are abolished in two *N.crassa* mutants that are 'blind' to light: white collar-1 (*wc-1*) and white collar-2 (*wc-2*) (Perkins *et al.*, 1982; Nelson *et al.*, 1989). No epistatic relationships have been described between *wc-1* and *wc-2*. These two genetic loci are indistinguishable by phenotype and map to chromosome VII and chromosome I of *Neurospora*, respectively. The pleiotropic phenotypes of these white collar mutants have suggested a role for the white collar gene products in the light signal transduction pathway. Genes under the control of blue light, such as *al-3*, *al-1*, *ccg-1* and *eas/ccg-2* have been shown to lose their light-inducible transcriptional activation in the white collar mutant backgrounds (Nelson *et al.*, 1989; Schmidhauser *et al.*, 1990; Arpaia *et al.*, 1993, 1995). Attempts to identify changes in protein content or phosphorylation in irradiated versus dark-grown cultures of *Neurospora* have not led to the identification of white collar gene products (Lauter and Russo, 1990).

Taking advantage of the apparent simplicity of the two-step light signal transduction system of *Neurospora*, we decided to clone and characterize the white collar gene products using a molecular genetic approach. Here, we present the isolation and characterization of the white collar-1 gene. *wc-1* encodes a 125 kDa protein, whose encoded motifs include a single class four, zinc finger DNA binding domain and a putative transcription activation region. The zinc finger domain of the *wc-1* product

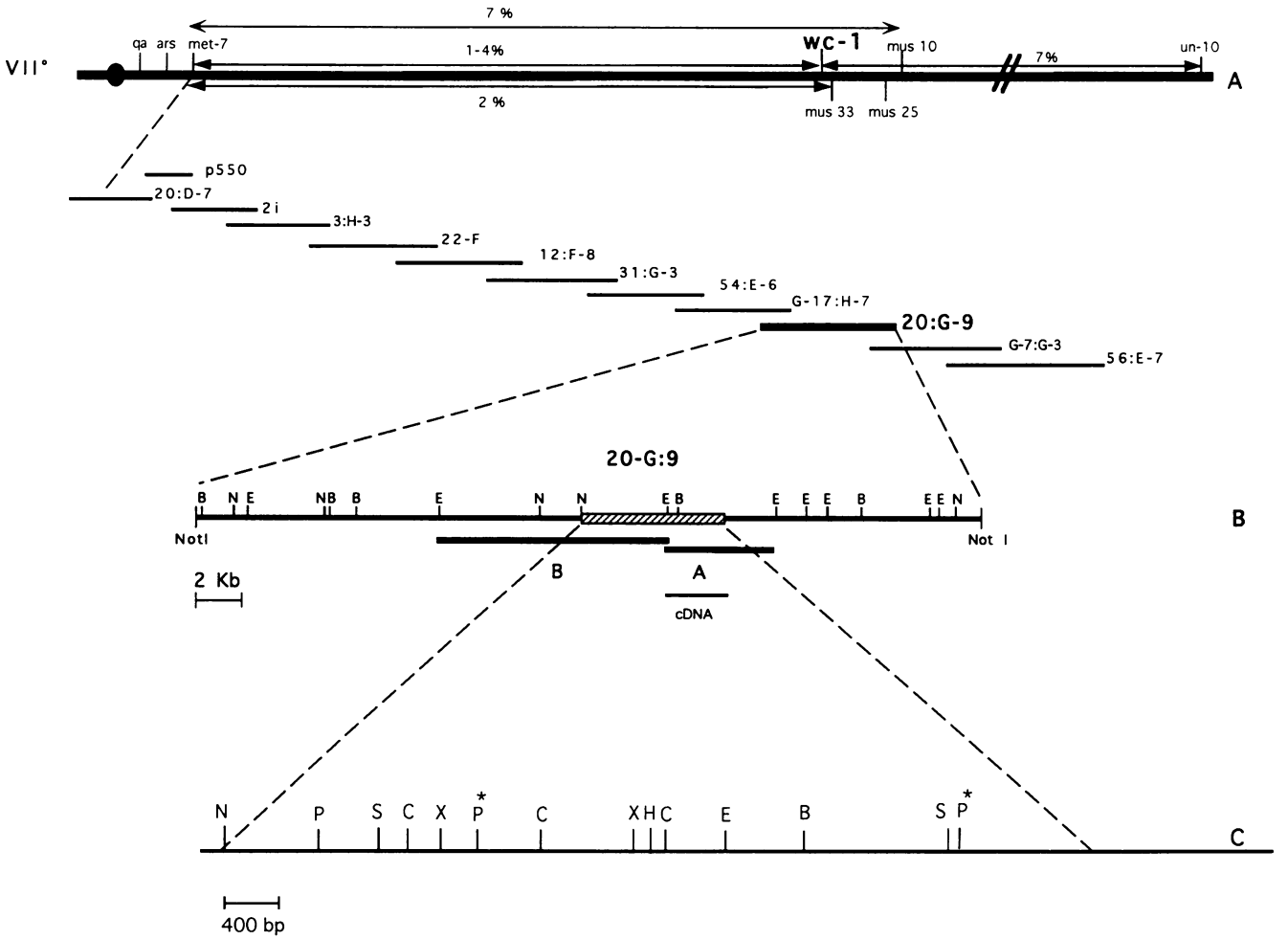


Fig. 1. Map-based cloning of the *wc-1* gene. (**A**) Genetic map of the right arm of *N.crassa* chromosome VII close to the centromere. The percentages of recombination between *wc-1* and other genetic loci are indicated. (**B**) Regional physical map obtained by chromosome walking to *wc-1*. For each walking step, only the most extended cosmid is presented and approximate overlap regions with neighboring cosmids is shown. The restriction map of cos 20G9 cosmid is illustrated in detail. B = *Bam*HI, E = *Eco*RI, N = *Nde*I. 'cDNA' indicates the Pv1 cDNA clone isolated from a λZAP *Neurospora* cDNA library. A and B indicate the two *Eco*RI fragments able to induce *wc-1* mutation by 'RIPing'. (**C**) Expansion of the *wc-1* gene region and restriction mapping. B = *Bam*HI, C = *Cla*I, E = *Eco*RI, H = *Hin*dIII, N = *Nde*I, P = *Pst*I, S = *Sal*I, X = *Xho*I.

expressed as a fusion protein in *Escherichia coli* is able to bind specifically *in vitro* to the promoter of *al-3*, a blue light-regulated gene of *Neurospora*. We demonstrate that *wc-1* gene expression is transcriptionally induced by blue light irradiation, that its transcription is autoregulated and that it is under the control of the *wc-2* gene.

Results

Cloning of the wc-1 gene by chromosome walking and complementation of the wc-1 mutant

Carotenoid biosynthesis in *Neurospora* is constitutive in conidia and light-inducible in mycelia. Therefore, the white collar phenotype is characterized by albino mycelium and orange conidia. Attempts to clone the *wc-1* gene by direct complementation of the *wc-1* mutant with *N.crassa* genomic cosmid banks and scoring for full induction of carotenoids in mycelia were unsuccessful (data not shown). Therefore, we decided to isolate the *wc-1* gene by chromosome walking and mutant complementation. The white collar-1 locus (Perkins *et al.*, 1982) was determined to map physically to the right arm of chromosome VII (Figure 1A), 1–4 units proximal to the cystathionine synthase gene (*met-7*). The *met-7* gene (Crawford *et al.*, 1992) was used to initiate the chromosome walk to *wc-1*, using two *N.crassa* genomic cosmid banks. This walk resulted in the identification of 10 overlapping cosmids spanning ~230 kb of genomic DNA. Individual cosmids isolated were tested for their correct localization on chromosome VII by restriction fragment length polymorphism (RFLP) analysis on multicent strains (Metzenberg *et al.*, 1984). In addition, these 10 cosmids were tested for their ability to restore photoinduced carotenoid biosynthesis to 'blind' *wc-1* mutant strains following transformation. A single cosmid, cos20G9 (38.5 kb), was able to complement the *wc-1* mutation.

Localization of the wc-1 coding sequence on cos20G9 by 'RIPing'

Next we attempted to identify the location of the *wc-1* gene within the cos20G9 cosmid. Restriction fragments of cos20G9 were subcloned into a *Neurospora* vector, used to transform the *wc-1* mutant, and tested for their ability to complement the *wc-1* mutation. This strategy was not successful due to the fact that the *wc-1* gene turned out to be fairly long and was cleaved by all the

restriction enzymes used (see next section). We therefore used an alternative approach to identify a DNA fragment in cos20G9 containing the *wc-1* gene. We exploited a genetic phenomenon unique to *Neurospora* called 'RIP' (Repeat-Induced Point mutation), wherein sequence duplications caused by transformation into wild-type background with extra copies of a native gene result in the silencing of both the native and introduced gene, after passage through meiosis (Selker, 1990). 'RIPing' silences duplicated genes via the induction of extensive GC→AT mutations in all duplicated sequences. We reasoned that only DNA subfragments of cos20G9 containing the *wc-1* gene could cause inactivation of the native gene and result in a 'blind' white collar-1 phenotype. The white collar-1 phenotype caused by 'RIPing' was observed only in the progeny of transformants containing the central 8 and 4.8 kb *Eco*RI fragments of cos20G9 (Figure 1B, fragments B and A, respectively) and not with the other *Eco*RI fragments. Molecular evidence that these 'RIPed' strains were 'blind' and lacking in light-induced responses was obtained by Northern analysis of light-induced genes (data not shown).

cDNA isolation and its use to obtain a white collar-1 null mutant by 'RIPing'

Using the 'RIPing' method described above, we determined that fragments A and B of cos20G9 were able to induce a white collar-1 phenotype in a wild-type *Neurospora* strain, and therefore each contains at least a portion of the *wc-1* gene. Next we used the *Eco*RI fragments A and B of cos20G9 to isolate putative cDNAs for *wc-1* by screening a *Neurospora* cDNA library. Only the fragment A probe (4.8 kb) allowed us repeatedly to isolate a cDNA clone (Figure 1B). The longest cDNA clone obtained (2.5 kb) was a partial cDNA truncated at the 5′ end. This clone contained an open reading frame (ORF) of 1509 bp and a poly(A) tail at the 3′ end. Northern analysis identified the cognate mRNA of this cDNA to be 5.4 kb (data not shown). As the cDNA library screened did not contain large inserts, and the *wc-1* mRNA is large, we were unable to isolate any longer cDNAs using fragment B which contains the 5′ end of the *wc-1* gene. To prove that the cDNA isolated with fragment A indeed encodes *wc-1*, we utilized it in a 'RIPing' experiment. We were able to show that this cDNA is able to induce mutants that exhibit a 'blind' white collar-1 phenotype in the progeny (data not shown).

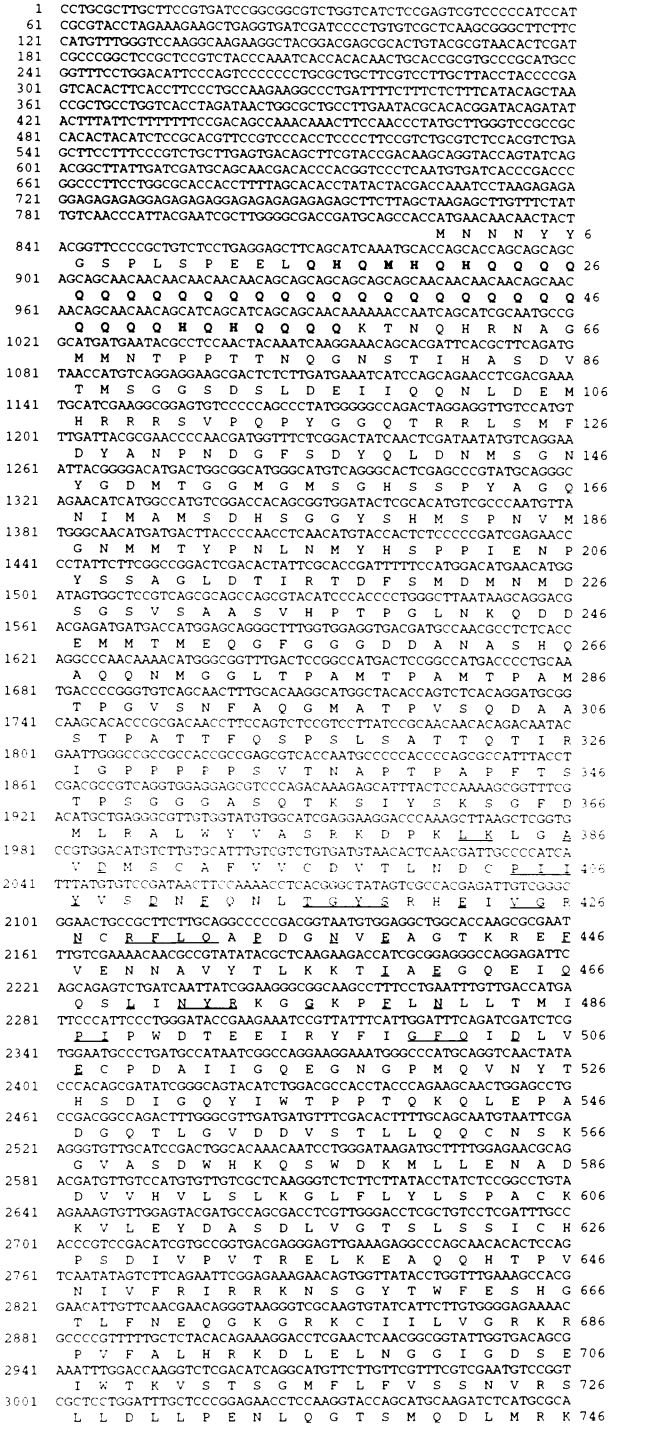
Nucleotide sequence analysis reveals that wc-1 encodes a putative transcription factor

To determine the 5′-most sequences of the *wc-1* gene, we sequenced the *Pst*I–*Pst*I genomic fragment of cos20G9 (4481 bp long) (Figure 1C). At the same time, we obtained the sequence of the entire cognate mRNA by reverse transcription PCR and used this to generate the protein sequence of the *wc-1* product (Figure 2). The identified *wc-1* gene encodes a protein of 1154 amino acids (125 kDa) (Figure 2). The *wc-1* genomic clone contains a single intron in the 3′ coding region (81 bp). The amino acid sequence encoded by the *wc-1* gene was analyzed by Genetics Computer Group programs. A hydropathy plot of the *wc-1* gene-encoded protein (WC1) indicates that it is a soluble protein with a putative nuclear targeting sequence (KKKRK). Amino acid sequence comparison revealed no overall homologies with other proteins; however, extremely informative domains were identified. A single putative class four, zinc finger DNA binding domain (C-X2-C-X18-C-X2-C) was present in the C-terminal region of WC1 (see double-dashed line in Figure 2). This type of zinc finger domain is characteristic of the vertebrate GATA factors (Orkin, 1992), and has also been identified in several fungal nitrogen-regulatory proteins (Crawford and Arst, 1993) (Figure 3). The *wc-1* gene product contains another feature common to transcriptional regulatory proteins, a polyglutamine stretch embedded in a glutamine–histidine-rich region at the N-termini. This kind of region previously observed in transcription factors such as *N.crassa* NIT4 (Yuan *et al.*, 1991) and *Saccharomyces cerevisiae* DAL81 (Bricmont *et al.*, 1991) has been shown to be involved in transcription activation. The presence of distinct domains for DNA binding and transcriptional activation suggests a transcriptional role for the *wc-1* gene product. WC1 also contains a conserved domain (spanning 115 amino acids) with 35% similarity to a bacterial regulatory gene, *bat*, which controls the light-driven ion pump of *Halobacterium halobium* (Leong *et al.*, 1988) (Figure 2, amino acids 382–507)

The WC1 zinc finger domain specifically binds to light-regulated promoter sequences of N.crassa

All of the domains contained within WC1 suggest that it is a transcription factor capable of binding to DNA. We therefore tested the ability of the zinc finger domain of WC1 to bind to specific DNA sequences contained within a light-inducible *Neurospora* promoter *in vitro*. The target *Neurospora* promoter used in these DNA binding experiments was albino-3 (*al-3*) which we have shown previously to be under blue light control, at the transcriptional level. The *al-3* promoter is a specific target for testing the WC1 protein, as we have shown previously that the photoinduction of the *al-3* gene is abolished in a *wc-1* mutant (Carattoli *et al.*, 1994). The region of the *al-3* promoter used in our DNA binding experiments (–126 to –87 bp) contains two canonical GATA sequences, located between the TATAA and CAAT boxes. The positioning of these GATA sequences in the *al-3* promoter is analogous to the GATA sequences in the promoter regions of genes regulated by other members of class four zinc finger proteins in fungi (NIT2, AREA, GLN3 and DAL80).

It has been demonstrated that GATA represents the core DNA binding sequence necessary for promoter recognition in vertebrates (Evans *et al.*, 1988) and in fungal class four zinc finger proteins (Crawford and Arst, 1993). The presence of these GATA sequences in the light-regulated *al-3* gene is reminiscent of the involvement of GATA repeats found in the promoters of plant light-regulated genes such as *Cab-E* and other light-regulated plant promoters (Schindler and Cashmore, 1990). To test the ability of the zinc finger domain of the WC1 protein to interact with the GATA sequences of the *al-3* promoter, the region of *wc-1* encoding the 111 amino acid zinc finger domain (see arrows in Figure 2), was expressed as a bacterial fusion protein in *E.coli*, under the direction of an inducible *tac* promoter. We tested the ability of this WC1 zinc finger fusion protein to bind to the *al-3* promoter



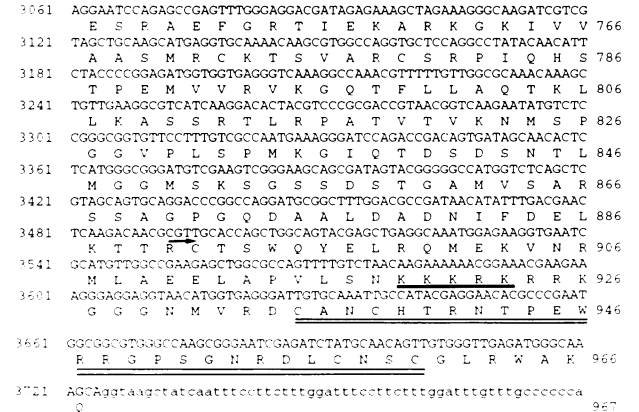


Fig. 2. Nucleotide and derived amino acid sequence of white collar-1. Nucleotides 1–5148 are derived from the genomic sequence of *wc-1*. Nucleotides denoted in lower case correspond to the 81 bp intron sequence. The partial *wc-1* cDNA clone isolated from the λZAP library starts at nucleotide 2776 (*Eco*RI site). The deduced amino acid sequence of *wc-1* was analyzed by the Genetics Computer Group programs (FASTA and PROSITE). Several structural motifs identified are indicated as follows. The glutamine-rich region is indicated in bold letters. The amino acids identical to residues 160–274 of the bacterial *bat* gene product are underlined. The double underline indicates the zinc finger motif. The two arrows indicate the zinc finger-containing region which was cloned and expressed as a fusion protein in *E.coli*. The bold underlined sequence identifies a putative nuclear targeting sequence. Several putative phosphorylation sites have been identified, but are not indicated in the figure for the sake of clarity. These putative phosphorylation sites include: (i) four putative sites for cAMP- and cGMP-dependent protein kinase (R,K)2x(S,T), (ii) 22 putative sites for protein kinase C (S,T)x(R,K) and (iii) 16 putative sites for casein kinase II (S,T)x2(D,E).

region *in vitro*, using an electrophoretic mobility shift assay (EMSA) (Figure 4A). For these gel-shift experiments, we incubated lysates from *E.coli* cells which express the WC1–glutathione transferase (WC1–GST) fusion protein with the double-stranded oligonucleotide (41mer) and observed distinct gel-shifts using EMSA (Figure 4A, lane 1). To demonstrate that this gel-shift resulted specifically from association with the WC1 fusion protein, we showed that no shift was obtained using extracts from uninduced cells (Figure 4A, lane 2). To demonstrate that the binding of the WC1 fusion protein to the *al-3* promoter sequence is sequence-specific, we performed competition experiments and showed that the binding was competed specifically by increasing amounts of an unlabeled 41mer (Figure 4B, lanes 3–6), but not by increasing amounts of an unrelated 41mer double-stranded oligo (Figure 4C, lanes 3–6). To test whether divalent cations such as zinc are required for specific binding of the WC1 fusion protein to the labeled 41mer, we carried out the experiments in the presence of EDTA. The marked reduction of band shift observed (Figure 4C, lane 7) indicates that the interaction of the WC1 fusion protein with the *al-3* promoter region is dependent on the presence of divalent cations. To explore the involvement of the GATA motif in the binding of the

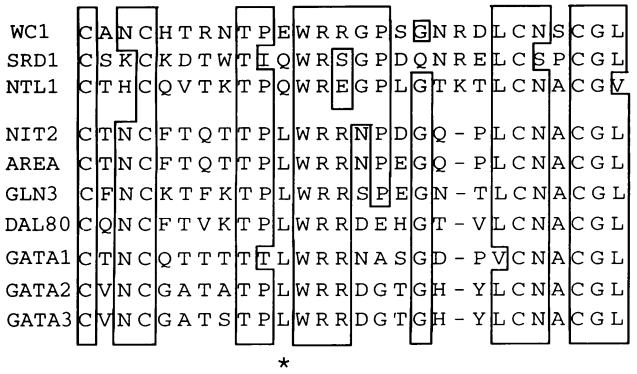


Fig. 3. Sequence identities between a region of WC1 and the zinc finger domains of other regulatory proteins. The residues between amino acids 935 and 962 of WC1 are shown. Identities between WC1 and other zinc finger regulatory proteins are boxed. The asterisk indicates a leucine residue that is conserved in all the class four zinc finger proteins with a 17 residue loop, but is absent in all three proteins containing an 18 residue loop. All the proteins shown in the comparison are cited in the text except human GATA-1, GATA-2 and GATA-3 whose accession numbers in the Swiss protein bank are P15976, P23769 and P23771, respectively.

zinc finger region of the WC1 protein to *al-3* DNA, we performed a binding competition experiment using increasing amounts of a mutated 41mer oligonucleotide, in which both the GATA sequences of *al-3* have been mutagenized (for sequence, see legend to Figure 4). Using the mutated 41mer double-stranded oligo, we observed no competition with wild-type *al-3* at low concentrations (1- to 50-fold excess of competitor) (Figure 4B, lanes 7–9). At 100-fold excess of mutant competitor, we observe some competition with the *al-3* wild-type GATA, indicating a reduced affinity of the WC1 protein for the GATA mutant probe (Figure 4B, lane 10). These results clearly point to an involvement of the promoter GATA motif in the WC1 protein binding. The low level but evident displacement of the wild-type sequence by a high molar excess of the mutated oligonucleotide suggests that the target region of the zinc finger domain may also involve other bases.

The wc-1 gene is autoregulated at the transcriptional level

Next we examined the levels of *wc-1* mRNA in dark-grown or light-induced mycelia, and observed that steady-state levels of *wc-1* mRNA increased upon blue light irradiation (Figure 5, compare lanes 1 and 2). In addition, we have shown that the light activation of *wc-1* is under the genetic control of *wc-1* and *wc-2* genes, as no light regulation is observed in the *wc-1* and *wc-2* mutant backgrounds (Figure 5, lanes 3–6). We and others have shown previously that light regulation of other target genes is also abolished in the white collar mutant backgrounds (Sommer *et al.*, 1989; Carattoli *et al.*, 1994). This result clearly indicates that *wc-1* is autoregulated, either by a light-induced increase in the rate of transcription or by stabilization of the transcript. Furthermore, the lack of light regulation of *wc-1* mRNA in a *wc-2* mutant suggests that the *wc-2* gene is epistatic to *wc-1*.

Discussion

Herein, we report the isolation of the *wc-1* gene, the characterization of its product, the definition of its role in light signal transduction and its light-dependent transcriptional regulation. The deduced amino acid sequence of WC1 suggests a role for this protein as a transcription factor involved in light regulation. Structural characteristics of WC1 include the presence of a zinc finger domain similar to that previously described for vertebrate GATA factors and also for general regulators of nitrogen metabolism in fungi like the *Aspergillus nidulans areA* gene, the *N.crassa nit-2* gene and the *S.cerevisiae* GLN3 and DAL80 genes (Cunningham and Cooper, 1991; Crawford and Arst, 1993). The vertebrate GATA factors have two zinc fingers, while the single zinc finger present in the fungal proteins is similar to the carboxy-proximal zinc finger of vertebrates. A comparison of the zinc finger domain of WC1 with those of other members of the class four zinc fingers is shown in Figure 3. All proteins with this class four zinc finger motif (17 aa finger loop), have been demonstrated to be transcriptional regulators capable of binding to promoter GATA sequences. Among the class four zinc finger proteins, WC1 has the unusual characteristic of possessing an 18 aa loop in the zinc finger and lacks a leucine residue in the loop that is normally conserved (indicated by the asterisk in Figure 3). These unusual features of WC1 are shared with other recently isolated putative transcription factors such as the tobacco putative nitrogen-regulatory protein NTL1 (Vedele and Caboche, 1993) and with the *S.cerevisiae* protein SRD1 (Hesse *et al.*, 1994). The leucine residue conserved in the majority of class four zinc finger proteins has been shown to be involved in the binding of the chicken GATA-1 protein to the GATA site of its cognate promoter (Omichinski *et al.*, 1993) and has also been proposed to be responsible for GATA binding in the AREA protein (Kudla *et al.*, 1990). The N-terminal region of WC1 contains a stretch of 28 glutamine residues, embedded in a region rich in glutamine and histidine residues (Figure 2, see underlined residues). Similar glutamine-rich regions are present in many other transcription factors and are required for their regulatory function *in vivo*, as demonstrated for regulatory factors such as *N.crassa* NIT4 (Yuan *et al.*, 1991) and *S.cerevisiae* DAL81 (Bricmont *et al.*, 1991). The importance of these domains is verified by the fact that deletions in these glutamine-rich regions in NIT4 or DAL81 result in loss of transcriptional activity. The region between amino acids 382 and 507 shows a 35% of similarity with the *bat* gene product, the putative activator of the bacterio-opsin gene *bop*. Interestingly, *bat* is itself regulated by a light-induced gene *brp* (bacterio-opsin-related protein) which is thought to be a membrane-bound light sensor or transmitter (Shand and Betlanch, 1991).

The idea that WC1 is a transcription factor is reinforced by the band-shift experiments which demonstrated that WC1 is able to bind to a specific *Neurospora* promoter sequence. The *al-3* promoter had been characterized by deletion and site-directed mutagenesis (Carattoli *et al.*, 1994). These results indicate that the *al-3* promoter region between positions –226 and –55 nt contains all the necessary information for blue light photoinduction. In this region, it is possible to identify two direct repeated GATA sequences separated by 17 nt. A 41mer oligonucleotide containing the *al-3* GATA sequence region has been shown to be retarded in EMSA experiments. In order to investigate the importance of the GATA sequences for

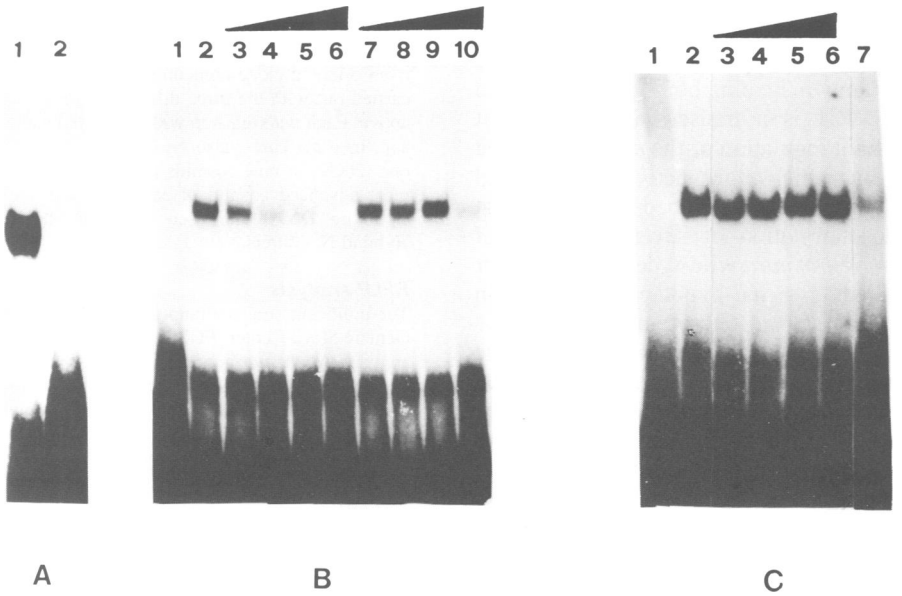


Fig. 4. The zinc finger domain of WC1 binds to the light-regulated *al-3* promoter. EMSA performed with the WC1–GST fusion protein and a 41mer double-stranded oligonucleotide (GCGGTATTATCGTCATAGCGTGCGGGTATCGAATATTGCC) derived from the *al-3* promoter region. The GATA motifs in the *al-3* promoter sequence are underlined. (**A**) One ng of terminally labeled oligonucleotide with 5 μg of WC1–GST-induced protein extract, lane 1; or 5 μg of WC1–GST uninduced extract, lane 2. (**B**) One ng of terminally labeled oligonucleotide incubated with: lane 1, no extract; lane 2, with 5 μg of WC1–GST-induced-extract; lanes 3–6, same as lane 2, but with the *al-3* 41mer oligonucleotide as unlabeled competitor (1, 10, 50 and 100 ng); lanes 7–10, as lane 2, but with a unlabeled *al-3* mutated oligonucleotide (GCGGTATGAGAGTCATAGCGTGCGGGAGGAGAATATTGCCC) as competitor (1, 10, 50 and 100 ng). The mutated GATA motifs are underlined. (**C**) Lanes 1 and 2, same as lanes 1 and 2 in (B); lanes 3–6, same as lane 2, but with an unlabeled unrelated 41mer oligonucleotide as competitor (1, 10, 50 and 100 ng); lane 7, same as lane 2, but in the presence of 50 mM EDTA.

specific binding of WC1, we performed competition experiments with unlabeled GATA mutated oligonucleotides. As expected, the mutated oligonucleotides do not displace the binding at low and high molar ratios. However, at very high molar ratios the mutant GATA begins to compete, suggesting the existence of other bases involved in the WC1 DNA binding. This effect has not been described for the other one-finger regulatory proteins (NIT2, AREA and GLN3); however, WC1 has unusual structural characteristics such as the 18 amino acid loop and the absence of the otherwise conserved leucine that could justify this difference. Further experiments on WC1 DNA binding ability and the characterization of the binding properties of the other 18 amino acid loop proteins will demonstrate whether these proteins belong to a new class of GATA-like proteins.

Northern analyses of *wc-1* mRNA levels reveals that *wc-1* is itself a light-regulated gene. The steady-state level of *wc-1* mRNA present in dark-grown cultures increases about five times following 20 min of light induction. This enhancement is abolished in a *wc-1* mutant, indicating that the *wc-1* product is required for the light induction of its cognate gene. We observed no light induction of *wc-1* transcription in a *wc-2* genetic background, suggesting that the *wc-2* product is necessary for white collar-1 light induction. This result may mean that *wc-1* and *wc-2* work together at the transcriptional level or that *wc-2* is epistatic to *wc-1*.

Collectively, the data presented indicate that the *wc-1* gene encodes an autoregulated transcription factor, able to bind to a light-inducible target promoter such as *al-3*. Even though *wc-1* gene expression is itself activated by light, there are two pieces of data which suggest that the

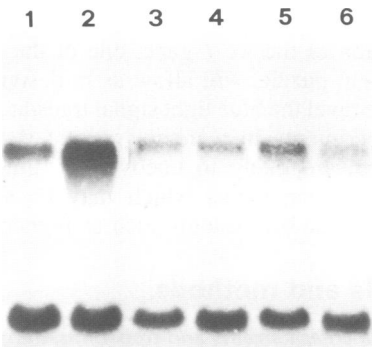


Fig. 5. Northern analysis of *wc-1* mRNA. Lanes 1 and 2, *Neurospora* wild-type strain (74A); lanes 3 and 4, *wc-1* mutant strain; lanes 5 and 6, *wc-2* mutant strain. *Neurospora* were grown either in the dark (lanes 1, 3 and 5) or exposed to blue light irradiation for 20 min (lanes 2, 4 and 6). Upper panel is probed with *wc-1* cDNA, the PV1 insert (upper panel). To normalize the amount of RNA in each lane, we used as the probe the light-independent gene IF2 (lower panel). The *wc-1* mRNA detected is ~5.4 kb long, as estimated from molecular size standards.

wc-1 gene product, present at low levels in dark-grown mycelia, can induce transcription of target genes in response to light. First, the light induction of the *al-3* transcript and other light-induced transcripts of *Neurospora* is rapid, and induction is detectable after only a few minutes of light induction (Sommer *et al.*, 1989; Baima *et al.*, 1991). Secondly, the light induction of *al-3* is not impaired by cycloheximide pre-treatment (Baima *et al.*, 1991). Therefore, the light-induced transcriptional

activation of the *al-3* gene and other genes probably results from a light-triggered post-translational modification of a protein. We have shown herein, both genetically and biochemically, that WC1 is a transcription factor that binds to and affects light regulation of the *al-3* gene. What is the mechanism by which light activates the rapid responses via WC1? It has been described previously that light induces dephosphorylation of several proteins of unknown function in *Neurospora* wild-type strains (Lauter and Russo, 1990), leading to the hypothesis that protein (de)phosphorylation is a step in phototransduction. WC1 described herein is characterized by several putative phosphorylation sites that are non-uniformly distributed within the polypeptide. This pattern is reminiscent of the organization of putative phosphorylation sites in other transcription factors (Crawford and Arst, 1993). The presence of potential post-translational modification sites in WC1 suggests the possibility that WC1, always present in the cells, becomes activated by specific light-induced phosphorylation or dephosphorylation. Finally, in *Neurospora* there are only two genetic loci known to be involved in blue light signal transduction: *wc-1* and *wc-2*. This is curious as all signal transduction pathways investigated so far include several molecular steps. We present data indicating that *wc-1* encodes a light-induced transcriptional activator. We also expect to have in this signal transduction pathway, at least, additional components including a blue light photoreceptor and a phosphorylation–dephosphorylation system. This expectation means that either additional white collar mutants will be isolated in the future or that the *wc-2* product is a photoreceptor with a kinase domain. The isolation of the *wc-1* gene, one of the key pieces of the blue light puzzle, will allow us to design new experiments to unravel the blue light signal transduction pathway. These experiments in a model genetic system, such as *Neurospora*, are likely to uncover the blue light signal transduction components which may be shared in less genetically tractable systems such as higher plants.

Materials and methods

Neurospora crassa strains and transformation

The strain *wc-1, qa-2, aro-9* was obtained by crossing strain FGSC 4398 (*wc-1a*) obtained from the Fungal Genetic Stock Center to strain FGSC 3957 (*qa-2, aro-9A*). This strain was used as the recipient in transformation experiments (carried out as described in Nelson *et al.*, 1989) involving CBM libraries, and the transformants were selected for growth in Vogel's minimal medium (Davis and de Serres, 1970). The strain FGSC 4403 (*wc-1*) was used for pMOcosx library transformation and these transformants were selected on minimal medium in the presence of hygromycin. Transformants were transferred routinely to slants in order to observe the mycelium carotenogenetic response to light irradiation. The strains 74A (wild-type), FGCS 4397 (*wc-1*) and FGCS 4408 (*wc-2*) were used for RNA preparations.

Neurospora crassa libraries and plasmids

Three *N.crassa* genomic libraries were used: (i) library CBM1, an ordered cosmid library containing 6048 clones (nomenclature: number:letter–number; Cabibbo *et al.*, 1991); (ii) library CBM2, a cosmid library randomly plated containing 15 000 individual clones (nomenclature: number–letter) and (iii) library pMOcosx, an ordered library represented by 4800 individual clones (nomenclature: letter–number:letter–number). The *N.crassa* cDNA library used was cloned in λZAP and was obtained from the Fungal Genetic Stock Center (Orbach *et al.*, 1990). Plasmid SKqa, an SK-plasmid (Stratagene) containing the *Neurospora qa-2* gene inserted between *Sal*I and *Pst*I sites, was used for subcloning and 'RIPing' experiments. Plasmid and cosmid DNA was isolated according to Birnboim and Doly (1979).

Chromosome walking

Chromosome walking to the *wc-1* locus began with clone p550 which is an oriented clone containing part of *met-7* and *met-9* genes and was carried out with the three different *N.crassa* cosmid libraries described above. Each walking step was conducted using the two ordered libraries and, in some cases, also on library CBM2. For each step, we isolated one (20G9) to nine cosmids (54E-6 region). The terminal region of the most advanced cosmid of each step was nick-translated (sp. act. 10^8 c.p.m./μg DNA) and used in colony hybridization experiments on Hybond N+ filters.

RFLP analysis

The multicent strains, comprising 38 individual strains from the Fungal Genetic Stock Center (FGSC 4450–FGSC 4488) were used for the RFLP analysis. *Neurospora crassa* DNA was prepared as described by Morelli *et al.* (1993). Restriction polymorphisms were searched for using either full cosmid inserts or single restriction fragments from the walking cosmids, and the results were compared with the annually published updates of RFLP of *Neurospora* in *Fungal Genetics Newsletter.*

'RIPing'

*Eco*RI subfragments of cos20G9 cloned in SKqa were used to transform *Neurospora* spheroplasts (*qa-2, aro-9*). For each DNA subfragment of cos20G9, two independent transformants were isolated and independently crossed to wild-type *Neurospora* (strain 74a). For each cross, we examined ~250 individual ascospores.

Nucleotide sequence analysis

The nucleotide sequence of *wc-1* was determined for both strands by the dideoxy method (Sequenase system USB) on the *wc-1* cDNA, PV1, and on the *Pst*I–*Pst*I genomic region of *wc-1* contained in cos20G9 (4481 bp long). The DNA sequence of the entire mRNA was obtained by reverse transcription PCR (Kawasaki, 1990). The most upstream primer for the RT–PCR starts at 385 nt of the sequence. The entire nucleotide sequence has been sent to the EMBL data bank (accession No. X94300). The derived amino acid sequence of WC1 was analyzed by Genetics Computer Group programs.

WC1 fusion protein preparation and EMSA

To construct the WC1–GST fusion protein, a PCR amplification product from the WC1 zinc finger region (Figure 2) was cloned in-frame into the *Eco*RI restriction site of the *E.coli* expression vector pGEX-2T (Smith and Johnson, 1988). The protein extracts were prepared essentially as previously described (Sessa *et al.*, 1993). All DNA binding assays were carried out by incubating probe plus extract at 4°C for 30 min in the presence of 1 μg poly(dI–dC) in a final volume of 20 μl. Electrophoresis was performed on 6% polyacrylamide gels in 0.5× TBE, run for 3 h at 4°C, then dried and exposed to X-ray film at –80°C for autoradiography. The specific radioactivity of the labeled oligonucleotide was ~10 000 c.p.m./ng.

Northern blot analysis

Total RNA was prepared from *Neurospora* mycelia grown in liquid as described (Baima *et al.*, 1991). For light induction studies, liquid cultures grown for 18 h at 25°C were irradiated with saturating light for 20 min, then immediately filtered and frozen. RNA was electrophoresed under denaturing conditions, transferred to Hybond-N and hybridized in the presence of formamide with radioactive probes (sp. act. 10^8 c.p.m./μg).

Acknowledgements

We thank Professor R.Geever for the gift of the p550 plasmid DNA, Dr I.Ruberti for advice on *E.coli* protein expression, Drs A.Folgori and L.Massimi for RFLP analysis and initial DNA binding assays, M.Helmer-Citterich and R.Gargamelli for skilfull technical assistance and Drs P.Filetici and G.Coruzzi for comments on the manuscript. This work was supported by the National Research Council of Italy, Special Project RAISA (subproject No. 2, paper No. 2477) and by Fondazione Pasteur-Cenci Bolognetti.

References

Arpaia,G., Loros,J., Dunlap,J.C., Morelli,G. and Macino,G. (1993) The interplay of light and the circadian clock: independent dual regulation of clock-controlled gene *ccg-2 (eas). Plant Physiol.*, **102**, 1299–1305.

Arpaia,G., Loros,J., Dunlap,J.C., Morelli,G. and Macino,G. (1995) The

circadian clock-controlled gene *ccg-1* is induced by light. *Mol. Gen. Genet.*, **247**, 157–163.

Ahmad,M. and Cashmore,A.R. (1993) The HY4 gene of *A.thaliana* encodes a protein with characteristics of a blue-light photoreceptor. *Nature*, **366**, 162–166.

Baima,S., Macino,G. and Morelli,G. (1991) Photoregulation of albino-3 gene of *Neurospora crassa*. *J. Photochem. Photobiol.*, **11**, 107–115.

Birnboim,H.C. and Doly,J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.*, **7**, 1513–1523.

Bricmont,P.A., Daugherty,J.R. and Cooper,T.G. (1991) The DAL 81 gene product is required for induced expression of differentially regulated nitrogen catabolic genes in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **11**, 1161–1166.

Cabibbo,A., Sporeno,E., Macino,G. and Ballario,P. (1991) CBM1, an *N.crassa* genomic library in pAC3 and its use for walking on chromosome VII right arm. *Fungal Genet. Newslett.*, **38**, 68–70.

Carattoli,A., Cogoni,C., Morelli,G. and Macino,G. (1994) Molecular characterization of upstream regulatory sequences controlling the photoinduced expression of the albino-3 gene of *Neurospora crassa*. *Mol. Microbiol.*, **13**, 787–795.

Chory,J. (1993) Out of darkness: mutants reveal pathways controlling light regulated development in plants. *Trend Genet.*, **9**, 167–172.

Crawford,J.M., Geever,R.F., Asch,D.K. and Case,M.E. (1992) Sequence and characterization of the *met-7* gene of *Neurospora crassa*. *Gene*, **111**, 265–266.

Crawford,N.M. and Arst,H.N. (1993) The molecular genetics of nitrate assimilation in fungi and plants. *Annu. Rev. Genet.*, **27**, 115–146.

Cunningham,T.S. and Cooper,T.G. (1991) Expression of the DAL80 gene, whose product is homologous to the GATA factors and is a negative regulator of multiple nitrogen catabolic genes in *S.cerevisiae* is sensitive to nitrogen catabolite repression. *Mol. Cell. Biol.*, **12**, 6205–6215.

Davis,R.H. and de Serres,F.J. (1970) Genetic and microbial research techniques for *Neurospora crassa*. *Methods Enzymol.*, **17A**, 79–143.

Degli Innocenti,F. and Russo,V.E. (1984) Isolation of new white collar mutants of *Neurospora crassa* and studies on their behaviour in the blue-light-induced formation of protoperitecia. *J. Bacteriol.*, **159**, 757–761.

Elich,T.D. and Chory,J. (1994) Initial events in phytochrome signalling, still in the dark. *Plant Mol. Biol.*, **26**, 1315–1327.

Evans,T., Reitman,M. and Fenselfeld,G. (1988) An erythrocyte-specific DNA-binding factor recognizes a regulatory sequences common to all chicken globin genes. *Proc. Natl Acad. Sci. USA*, **85**, 5976–5980.

Harding,R.W. and Melles,S. (1984) Genetic analysis of the phototrophism of *Neurospora crassa* perithecial beaks using white collar and albino mutants. *Plant Physiol.*, **72**, 996–1000.

Harding,R.W. and Turner,R.V. (1981) Photoregulation of the carotenoid biosynthetic pathways in albino and white collar mutants of *Neurospora crassa*. *Plant Physiol.*, **68**, 745–749.

Hesse,S.M., Stanford,D.R. and Hopper,A.K. (1994) SRD1, an *S.cerevisiae* gene affecting pre-rRNA processing contains a C2/C2 zinc-finger motif. *Nucleic Acids Res.*, **7**, 1265–1271.

Kawasaki,E.S. (1990) *Amplification of RNA in PCR Protocols: A Guide to Methods and Applications*. Academic Press Inc., pp. 21–27.

Kudla,B., Caddick,M.X., Langdon,T., Martinez-Rossi,N.M., Bennet,C.F., Sisley,S., Davies,R.W. and Arst,H.N. (1990) The regulatory gene *areA* mediating nitrogen metabolite repression in *Aspergillus nidulans*. Mutations affecting specificity of gene activation alter a loop residue of a putative Zn-finger. *EMBO J.*, **9**, 1355–1364.

Lauter,F.R. and Russo,V.A. (1990) Light-induced dephosphorylation of a 33kDa protein in the wild-type strain of *Neurospora crassa*. The regulatory mutants *wc-1* and *wc-2* are abnormal. *J. Photochem. Photobiol.*, **5**, 95–103.

Lauter,F.R. and Russo,V.A. (1991) Blue light induction of conidiation specific genes in *Neurospora crassa*. *Nucleic Acids Res.*, **19**, 6883–6886.

Leong,D., Boyer,H. and Betlach,M. (1988) Transcription of genes involved in bacterio-opsin gene expression in mutants of a halophilic Archeabacterium. *J. Bacteriol.*, **170**, 4910–4915.

Metzenberg,R.L., Stevens,J.N., Selker,E.U. and Morzycka-Wroblewska,E. (1984) A method for finding the genetic map positioning of cloned DNA fragments. *Fungal Genet. Newslett.*, **31**, 35–39.

Morelli,G., Nelson,M.A., Ballario,P. and Macino,G. (1993) The photoregulated carotenoid, biosynthetic genes of *N.crassa*. *Methods Enzymol.*, **214**, 412–424.

Nelson,M.A., Morelli,G., Carattoli,A., Romano,N. and Macino,G. (1989) Molecular cloning of *Neurospora crassa* carotenoid biosynthetic gene (albino-3) regulated by blue light and the product of the white collar-1 gene. *Mol. Cell. Biol.*, **9**, 1271–1276.

Neuhaus,G., Bowler,C., Kern,R. and Chua,N.H. (1993) Calcium/calmodulin-dependent and independent phytochrome signal transduction pathways. *Cell*, **73**, 937–952.

Omichinski,J., Clore,G., Schaad,O., Fenselfeld,G., Trainor,C., Appella,E., Stahl,S. and Gronenborn,A. (1993) NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science*, **261**, 438–446.

Orbach,M.J., Sachs,M.S. and Yanofsky,C. (1990) *Neurospora crassa arg-2* locus. Structure and expression of the gene encoding the small subunit of arginine specific carbamoyl-phosphate synthetase. *J. Biol. Chem.*, **265**, 10981–10987.

Orkin,S.H. (1992) GATA-binding transcription factors in hematopoietic cells. *Blood*, **80**, 575–581.

Perkins,D.D., Radford,A., Newmeyer,D. and Bjorkman,M. (1982) Chromosomal loci of *Neurospora crassa*. *Microbiol. Rev.*, **46**, 426–570.

Quail,P.H. (1991) Phytochrome; a light-activated molecular switch that regulates plant gene expression. *Annu. Rev. Genet.*, **25**, 389–409.

Sargent,M.L. and Briggs,W.R. (1967) The effects of light on a circadian rhythm of conidiation in *Neurospora*. *Plant Physiol.*, **42**, 1504–1510.

Schindler,U. and Cashmore,A.R. (1990) Photoregulated gene expression may involve ubiquitous DNA binding proteins. *EMBO J.*, **9**, 3415–3427.

Schmidhauser,T.J., Lauter,F.R., Russo,E.A. and Yanofsky,C. (1990) Cloning, sequence and photoregulation of *al-1*, a carotenoid biosynthetic gene of *Neurospora crassa*. *Mol. Cell. Biol.*, **10**, 5064–5070.

Selker,E.U. (1990) Premeiotic instability of repeated sequences in *Neurospora crassa*. *Annu. Rev. Genet.*, **24**, 579–631.

Sessa,G., Morelli,G. and Ruberti,I. (1993) The Athb-1 and Athb-2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. *EMBO J.*, **12**, 3507–3517.

Shand,R.F. and Betlanch,M.C. (1991) Expression of the *bop* gene cluster of *Halobacterium halobium* is induced by low oxygen tension and by light. *J. Bacteriol.*, **173**, 4692–4699.

Short,T.W. and Briggs,W.R. (1994) The transduction of blue light signals in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **45**, 143–171.

Smith,D.B. and Johnson,K.S. (1988) Single-step purification of polypeptides expressed in *E.coli* as fusions with glutathione S-transferase. *Gene*, **67**, 31–40.

Sommer,G.J., Chambers,A.A., Eberle,J., Lauter,F.R. and Russo,V.E. (1989) Fast light-regulated genes of *N.crassa*. *Nucleic Acids Res.*, **17**, 5717–5724.

Vedele,F. and Caboche,M. (1993) A tobacco cDNA clone encoding a GATA-1 zinc-finger protein homologous to regulators of nitrogen metabolism in fungi. *Mol. Gen. Genet.*, **240**, 365–373.

Yuan,G., Fu,Y. and Marzluf,G.A. (1991) *nit-2*, the major nitrogen regulatory gene of *N.crassa* encodes a protein with a putative Zn-finger DNA-binding domain. *Mol. Cell. Biol.*, **11**, 5735–5745.

Received on November 10, 1995; revised on December 8, 1995