The GA octodinucleotide repeat binding factor BBR participates in the transcriptional regulation of the homeobox gene Bkn3

Luca Santi¹, Yamei Wang¹, Maria Rosaria Stile², Kenneth Berendzen³, Dierk Wanke¹, Cristina Roig¹, Carlo Pozzi³, Kai Müller⁴, Judith Müller⁴, Wolfgang Rohde¹ and Francesco Salamini¹,²

¹Max-Planck-Institut für Züchtungsforschung, Department of Plant Breeding and Yield Physiology, Carl-von-Linné-Weg 10, 50829 Köln, Germany, ²University of Naples ‘Federico II’, Department of Soil, Plant and Environmental Sciences, Section of Plant Genetics and Horticulture, Via Universitá 100, 80095 Portici, Italy, ³University of Milano, Department of Plant Production, Via Celoria 2, 20133 Milano, Italy, and ⁴Fraunhofer IME, Department of Applied Genome and Proteome Research, RWTH Aachen, Institute of Biology VII, Worringer Weg 1, 52056 Aachen, Germany

Received 10 February 2003; accepted 11 March 2003.
¹For correspondence (fax +49 221 5062 413; e-mail salamini@mpizkoeln.mpg.de).

Summary

In the dominant mutant Hooded (K), the barley gene Bkn3 is overexpressed as a result of a duplication of 305 bp in intron IV. When fused to a cauliflower mosaic virus 35S minimal promoter, the 305 bp element activates gene expression in tobacco, as does a 655 bp Bkn3 promoter sequence. Both DNA fragments contain a (GA)₈ repeat (GA/TC)₈. A one-hybrid screen using the 305 bp element as the DNA target led to the cloning of the barley b recombinant (BBR) protein, which binds specifically to the (GA/TC)₈ repeat. BBR is nuclear targeted and is a characterized nuclear localization signal (NLS) sequence, a DNA-binding domain extended up to 90 aa at the C-terminus and a putative N-terminal activation domain. The corresponding gene has no introns and is ubiquitously expressed in barley tissues. In co-transfection experiments, BBR activates (GA/TC)₈-containing promoters, and its overexpression in tobacco leads to a pronounced leaf shape modification. BBR has properties of a GAGA-binding factor, but the corresponding gene has no sequence homology to Tra and Psq of Drosophila, which encode functionally analogous proteins. In Arabidopsis, (GA/TC)₈ repeats occur particularly within 1500 bp upstream of gene start codons included in some homeodomain genes of different classes. The data presented suggest that expression of the barley Bkn3 is regulated, at least in part, by the binding of the transcription factor BBR to GA/TC repeats.

Keywords: homeobox genes, Kno, intron-mediated regulation, BBR, (GA)₈, PRE-like DNA elements.

Introduction

Hooded (K) of barley is a dominant mutant introduced into the western countries from the Himalayas (Badr et al., 2000; Harlan, 1931; Stebbins and Yagil, 1966). The mutation is caused by a tandem direct duplication of 305 bp in the fourth intron of the homeobox gene Bkn3, resulting in the ectopic overexpression of Bkn3 at the lemma-awn transition region: a local acquisition of meristematic activity is followed by the formation of extra floral structures (Müller et al., 1995). The presence of several barley mutants resulting in the modification of the lemma-awn transition region – among others, calcaroides (cal), Elevated Hood (K⁰), suppressor of Hooded (suK) and short awn (lk) – makes the transition zone of particular interest in the study of leaf development (Pozzi et al., 2000).

Bkn3 belongs to class 1 Kno homeobox genes, and it is the orthologue of Knotted-1 (Kn1) of maize (Vollbrecht et al., 1991). In maize and tomato, the analysis of several gain-of-function homeobox mutations – like Kn1, Rs1, Lg3, Gn-1 and TKn2/LeT6 – supports the conclusion that class 1 Kno genes are necessary for maintenance of meristems and development of leaf primordia (summarized in Pozzi et al., 1999). Their role has been supported by the study of
recessive loss-of-function mutants in maize and Arabidopsis, which are defective in meristem maintenance and inflorescence branching (Kerstetter et al., 1997; Long et al., 1996; Vollbrecht et al., 2000). Class 1 Knox genes, when overexpressed in heterologous systems, cause complex leaf phenotypes and ectopic meristem formations (Chan et al., 1998; Chuck et al., 1996; Lin and Muller, 2002; Muller et al., 1995). The class 1 Knox gene products seem to be post-translationally regulated, as supported by the finding that barley BKN3 associates with proteins encoded by classes 1 and 2 Knox genes, as well as by Bell homeobox genes (Muller et al., 2001). BELL1 of Arabidopsis is also able to functionally interact with other KNOX TALE homeodomain proteins (Bellaoui et al., 2001).

As for other homeotic genes, regulation of homeobox gene expression plays a major role in the correct development of complex organisms (Muller and Leutz, 2001). Domains of homeotic gene expression and their variation throughout development are regulated by transcriptional activation and/or repression (Kehle et al., 1998; Paro and Harte, 1996). For example, trithorax group genes (trxG) maintain an active transcriptional state of homeotic genes (Kennison, 1995), while Polycomb group genes (PcG) support their repressed state (Chan et al., 1994). In Drosophila, the gene trithorax-like (trx), required for the expression of homeotic genes, encodes the GAF factor (Bhat et al., 1996; Farkas et al., 1994), which binds to GAGAG sites, often located in homeobox regulatory regions (Busturia et al., 2001; Tsukiyama et al., 1994). Also, the Psq of Drosophila has been proved to bind to the same site (Lehmann et al., 1998).

In the Hooded gene, as for the mutated alleles of maize Kn-1 (Greene et al., 1994; Hake et al., 1989) and of the Arabidopsis AG (Yanofsky et al., 1990) and CLF (Goodrich et al., 1997), the mutation is located in an intron. For the AG gene, moreover, it is known that cis elements responsible for spatial gene regulation are also intron located (Sieburth and Meyerowitz, 1997). Introns are known to influence the time and tissue specificity of gene expression (Fridman et al., 2000; Gidekel et al., 1996; Hwang et al., 2002; Muller et al., 1995; Plesse et al., 2001; Vollbrecht et al., 1991). In this paper, the transcriptional regulatory role of intron IV of the barley homebox gene Bkn3 has been investigated, concentrating on the 305 bp element whose duplication affects the transcription of the gene. A one-hybrid screen, based on the 305 bp element as the DNA target, allowed the isolation of a set of protein candidates for binding to the element. In this report, we characterize the first of those proteins, barley b recombinant (BBR). The protein encoded by this gene has been easily purified, allowing its capacity to interact with a double-stranded (GA)₉ sequence (hereafter indicated as (GA/TG)₉) present inside the 305 bp element to be tested. The corresponding gene BBR has no sequence homology with Trl and Psq of Drosophila, although its product has properties of GAGA-binding factor.

**Results**

**The Bkn3 305 bp intron IV element has an enhancer activity in tobacco**

The starting point of the analysis was to understand if the 305 bp element was capable of modifying gene expression when tested in an experimental system. To this scope, five promoter GUS constructs were produced (Figure 1a): in the first three, the 305 bp intron sequence was cloned upstream of the cauliflower mosaic virus 35S (CaMV 35S) minimal promoter (−46 to +9 bp; Fang et al., 1989). In the fourth construct, the GUS gene expression was under the control of only the CaMV 35S minimal promoter, and in the fifth, the same reporter gene was driven by a promoter corresponding to the 655 bp upstream of the Bkn3 ATG. Ten transgenic tobacco plants were generated via Agrobacterium-mediated transformation. Vegetative apices, shoot buds with leaf primordia, apices of inflorescences and flowers were screened for GUS activity. According to the time needed to develop intense GUS staining, the plants derived from each construct were divided into three classes. One plant for each class and for each construct was selected, and the T₁ and T₂ generations were tested for GUS activity.

In the vegetative shoot apex, GUS expression was evident for all 305 bp CaMV 35S minimal promoter constructs. During inflorescence and flower development, the 305 bp CaMV 35S minimal promoter constructs activated GUS expression in young inflorescence and flower primordia, while the CaMV 35S minimal promoter was inactive. Staining was observed at the branching of the tobacco inflorescence (Figure 1b,e) and at the base of the flower (Figure 1b,c). Based on the times of colour development, the first three constructs showed little or no difference in expression (Figure 1a).

GUS expression in the main shoot apex and in emerging lateral buds was also observed for the construct based on the 655 bp Bkn3 promoter. However, the intensity of staining was quantitatively different. Staining at branching was observed only after 12 h incubation (Figure 1a). A remarkable difference between Bkn3 and 305 bp based promoters was the absence of the reporter gene activity at the base of the flower in the former (Figure 1d). The experiment reported here led to a conclusion that the 305 bp element is a candidate for studying protein-binding-site-mediated gene transcription.

**Yeast one-hybrid screen**

The screen allowed the isolation of cDNAs encoding for proteins binding to the 305 bp element. The screening was carried out in a mutant strain of YM4271 created to knock out yeast factor(s) that could bind to the 305 bp element. The one-hybrid screening of the barley cDNA library
identified 58 positive colonies. Sequence analysis allowed the definition of four different barley candidate genes.

One clone for each of the four genes was also tested in the YM4271 lacZi non-mutated strain. Two clones from two different genes were able to overcome the repressive effect caused by the presence of the 305 bp element in the yeast non-mutated strain (see Experimental procedures). One of these genes was designated BBR, from barley b recombinant. Four different truncations of BBR were identified at a frequency of $0.75 \times 10^{-6}$ (clones BBR a, b, c, d). The four clones were, respectively, 1251, 1008, 759 and 660 bp long, encoding 325, 244, 161 and 128 aa (Figure 2a).

**BBR**

To acquire information on the BBR gene, barley genomic and cDNA inflorescence libraries were screened. A BBR genomic sequence was established from the 4946 bp clone gWBBR (EMBL accession AJ507214), while the longest cDNA sequence was from the 1380 bp long c clone cSBBR, allowing the extension of BBR by further 136 bp. Of the 136 bp added to the 5’ end, 75 encoded 25 aa and 61 bp represented part of the 5’ UTR. On inspection, the cSBBR cDNA sequence was 100% identical to the sequences of the four one-hybrid cDNAs and to the corresponding genomic sequence of gWBBR, indicating the absence of introns. The BBR ORF encodes for 350 aa (Figure 2b,c). BBR homologous proteins have been reported from rice, tomato, Arabidopsis thaliana and soybean (Figure 2d and Discussion). The homology among these sequences is particularly confined in the last 90 aa at the C-terminal end of the protein (Figure 2d). As this sequence is included in the shortest clone isolated in the one-hybrid screen (BBRa; Figure 2a), it was concluded that this region contained the DNA-binding
Luca Santi et al.

© Blackwell Publishing Ltd, The Plant Journal, (2003), 34, 813-826
domain. A second feature of BBR is a long stretch of 14 repeated amino acid residues (from aa 63–123; glutamine and histidine, QH, in Figure 2c).

To define the BBR transcription initiation site, a primer extension analysis was carried out. Transcription initiates at a position located 18 bp downstream from a putative CAAT box, and the corresponding base is numbered +1 in Figure 2(b). Based on primer extension site, the complete transcriptional unit of BBR covers 1401 bp (Figure 2b).

A. thaliana BBR homologues form a small gene family of seven members: At1g688120 (chromosome 1), At2g01930 (chromosome 2), At1g14665 (chromosome 1), At5g42520 (chromosome 5), At2g21240 (chromosome 2), At4g38910 (chromosome 4) and At2g35550 (chromosome 2). The sequences of these genes indicate that their ORFs are not interrupted by introns. All At BBR genes with known 5′ UTRs contain GA/CT repeats.

**BBR encodes a nuclear targeted protein**

The subcellular localization of BBR was established by transiently expressing BBR::GFP and BBR::RFP fusions in cultured tobacco BY2 cells *in vitro*. Nuclear targeted proteins contain stretches of positively charged amino acids (Chelsky *et al.*, 1989). In BBR, such stretches are present in the amino acid sequence PVKKRQQGRQPVKKPKKK at aa position 198–219 (Nakai and Kanehisa, 1992; Figure 3a). This conforms to a consensus nuclear localization signal (NLS) sequence (Robbins *et al.*, 1991). A BBR::GFP fusion was assembled in the vector pGFPJS and a BBR::RFP in pGJ1425. As a control, BY2 cells were transformed with the empty vectors. Cells transformed with pGFPJS-BBR::GFP and with pGJ1425-BBR::RFP fusion constructs showed nuclei with green or red fluorescence, respectively (Figure 3b,c).
The BBR protein binds in vitro to a GA/CT repeat present in the 305 bp sequence

In vitro DNA-binding studies were carried out to address BBR-binding properties. Two strategies were adopted to produce recombinant BBR: in vitro transcription and translation and BBR expression as an IPTG-inducible glutathione-S-transferase (GST) fusion protein in *Escherichia coli*. In both cases, the longest yeast one-hybrid BBR clone, lacking the last 25 aa at the N-terminal end, was used. The GST::BBR fusion protein was purified via affinity chromatography on glutathione-agarose beads (Figure 4e). The radioactively labelled 305 bp element was the first target sequence tested. Recombinant BBR bound the 305 bp element in vitro (Figure 4c). The binding was competed by a 200-fold excess of unlabelled DNA probe. In a second set of experiments, several deletions of the 305 bp element were generated (Figure 4a). Three overlapping fragments were produced by PCR, and only the fragment covering the last 125 bp (from 180 to 305 bp) was retarded by BBR (data not shown). This was subsequently reduced to four shorter fragments. From this point on, synthetic oligonucleotides having 10 bp overlaps were used in the assay. One fragment, from bp 238 to 276, was retarded by BBR, defining a 39 bp binding sequence. This sequence was further shortened to 26 bp (from bp 243 to 268), and finally, fragments TCTCTGCTCTCTCTCTC (GA/TC)_n and TCTCTCTCTCTCTCT (GA/TC)_n were tested (the first one is a mutated version of the second; Figure 4d). BBR was able to bind specifically to the latter double-stranded oligonucleotide. This allowed to define a DNA sequence of the type (GA/TC)_n, with n equal to 8 but not excluding 6 and 7, as the precise binding site inside the 305 bp sequence (Figure 4b). All other DNA sequences present in the 305 bp element were not capable of binding BBR.

**BBR transactivates promoters containing the 305 bp element in planta**

Co-transfection of SR1 tobacco protoplasts was carried out with (i) effector constructs (either FullBBR or ADFullBBR, which contains the herpes simplex virus

---

**Figure 4.** The BBR protein binds to the (GA/TC)n sequence present in the 305 bp element.
(a) Schematic representation of the strategy used in retardation experiments to define the precise site of BBR binding.
(b) Sequence of the Bx3 intron IV 305 bp element showing in black the GA/TC octadecanucleotide element.
(c, d) Band shift experiments. The 305 bp element was used as a probe in the experiment illustrated in (c). Similar results were obtained with other probes shown in black (a). In (d), the probes were either (left) GAGAGAGAGAGAGA, or (right) GAGAGAGAGAGAGA. 50, 100, 200 x refer to molar excess of cold probe included in the reaction. GST: glutathione S transferase protein.
(e) PAGE of the BBR::GST fusion (62 kDa, open arrow), purified by affinity chromatography. Lane at right: molecular markers.

© Blackwell Publishing Ltd, *The Plant Journal*, (2003), 34, 813-826
VP16 activation domain) used alone or in combination with (ii) the reporter construct Min1, where the 305 bp element is positioned upstream of the CaMV 35S minimal promoter and (iii) a K373 construct carrying the β-glucoronidase (UidA) gene under the control of a CaMV 35S minimal promoter (Figure 5a). In a second set of experiments, co-transfections included the LUC vector necessary for GUS data normalization (see Experimental procedures).

FullBBR activated the expression of the Min1 reporter gene 5.7-fold, supporting the binding in planta of BBR to the 305 bp element (Figure 5b). Induction was also achieved when the ADFullBBR effector construct was used. Min1 reporter alone supported a basal level of GUS expression. This was much lower compared to Min1 + effectors, but three times higher than the level of GUS recorded for K373 alone (Figure 5b). Compared to K373, Min1 differed for the presence of the 305 bp DNA element. In tobacco protoplasts, the basal level of expression of Min1 alone can be explained by endogenous activating factors binding to the 305 bp element. In the second set of experiments, just the FullBBR effector and the Min1 reporter constructs were co-transfected, leading to an average of 6.7-fold activation of the reporter construct mediated by BBR (Figure 5b).

The question as to whether BBR is able to transactivate promoters different in sequence from Min1, but still containing a (GA/TC)₈ repeat, was addressed by co-transfecting FullBBR together with PBkn3 (construct V in Figure 1a). The experiment was performed four times, and the comparison between the results of the reporter construct alone, or co-transfected with FullBBR effector construct (data normalized based on luciferase values), showed 3.59-fold activation. These data are statistically significant and support the conclusion that, in co-transfection experiments, the activation by BBR of reporter constructs is mediated through the (GA/TC)₈ repeat.

**BBR expression in barley and BBR overexpression in tobacco**

Tissue specificity of expression of BBR in barley was addressed by Northern blot and quantitative RT-PCR. Northern blot mRNAs were extracted from different tissues of the wild type (wt) cv. Atlas and from K Atlas and were probed with cSBBR. An Antirrhinum actin cDNA fragment was used as a reference in each lane. The 1.5 kb long BBR transcript was detected in all barley tissues (Figure 6a). The highest level was observed in mature embryos and the lowest in seedling leaves. Similar results were obtained by RT-PCR analysis (Figure 6b).

Three versions of BBR, driven by the CaMV 35S promoter (~941 to +208 bp; Odell et al., 1985) were introduced into SR1 tobacco (Figure 7a). Construct I (ΔBBR) included the BBRd gene truncation encoding a 25 aa defective protein. Construct II (ADΔBBR) included the yeast GAL4 activation domain fused 5’ upstream to the BBR gene. Construct III (FullBBR) contained the full-length BBR sequence. T₀ tobacco plants transformed with ΔBBR did not show evident phenotypic changes. Nine out of 10 T₀ tobacco plants transformed with the GAL4AD fusion construct, 6 out of 10 ADΔBBR plants and six out of seven plants transformed with FullBBR had a significantly altered phenotype that was maintained in T₁ and T₂ progenies. Compared to wt SR1, the phenotype of these plants had narrow and more elongated leaves (Figure 7b,c). Also, the cotyledons were elongated with rolled lateral margins and a curved shape (Figure 7e). This phenotype persisted during development. Flowers had elongated sepals and petals, with pistils and stamens positioned inside the corolla (Figure 7d).
the mutant compared to wt excluded this possibility (data not shown).

**Bioinformatics**

Distribution of the GA/TC repeats was verified in the available sequences of the Arabidopsis genome. Chromosomes were scanned for the (GA/TC)$_6$ element with a 1.5 kb window, and the number of loci hosting (GA/TC)$_6$ repeats (not considering repeat length) was 547 (Table 1). The frequency observed was about 5.6 $\times$ 10$^4$ times higher than the frequency that would have been expected if its occurrence in the genome was random. The last row of Table 1 provides an estimate of the number of base pairs between two (GA/TC)$_6$ repeats (for $n$-values between 6 and 11). As expected, the distance in base pairs between two repeats was the lowest in promoters. This type of analysis was also carried out for a fraction of the rice genome, and it provided comparable results (data not shown).

**Discussion**

The data presented provide evidence that a 305 bp fragment in the intron IV and a 655 bp in the promoter of the homeobox gene Bkn3 can activate gene expression in specific developmental domains in planta. Both sequences contain a (GA/TC)$_6$ element, although in reverse orientation. The (GA/TC)$_6$ element is sufficient to direct gene expression in yeast mediated by BBR. The BBR protein is nuclear targeted and is characterized by an NLS sequence, a DNA-binding domain residing at the C-terminal (and extended up to 90 aa), a QH domain and a putative N-terminal activation domain. The corresponding gene is ubiquitously expressed in barley tissues and, particularly, in embryos. At the time when this manuscript was ready for submission, Sangwan and O’Brien (2002) published data on the identification of the plant protein GBP that interacts with AG/TC dinucleotides. The new soybean protein was from a one-hybrid screen carried out with a DNA bait corresponding to a (GA/TC)$_{67}$ sequence. This motif is present as a (GA/TC)$_6$ element in the promoter of the soybean

### Table 1 Number and frequency of (GA/TC)$_6$ repeats in the Arabidopsis genome

<table>
<thead>
<tr>
<th></th>
<th>Total number</th>
<th>In exons</th>
<th>In introns</th>
<th>In promoters (from –1500 to ATG)</th>
<th>In intergenic regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>(GA/TC)$_6$</td>
<td>547</td>
<td>5</td>
<td>72</td>
<td>344</td>
<td>126</td>
</tr>
<tr>
<td>Number of base pairs between two (GA/TC)$_n$ repeats</td>
<td>69068</td>
<td>775000</td>
<td>60476</td>
<td>38943</td>
<td>99068</td>
</tr>
</tbody>
</table>

The distribution of the (GA/TC)$_6$ sequences in coding and non-coding regions, as well as the percentage of repeats on their total number, is also reported. The second row is an estimate of the number of base pairs between two (GA/TC)$_n$ repeats, with $n$ between 6 and 11. The complete set of data for $n = 6 \to 11$ is presented in Table S2.
Gsa1 gene that is expressed in leaves and in root nodules. 

In vitro, GBP binds to the Gsa1 promoter. The published data support our isolation of BBR as a (GA/TC)$_n$-binding protein and reinforce the conclusion that BBR-like proteins are widespread in plants. BBR is 58.6% identical to the soybean GBP amino acid sequence, a percentage that increases to 81.9% in the putative DNA-binding domain. In co-transfection experiments carried out in tobacco protoplasts, BBR activates (GA/TC)$_n$ containing promoters, and overexpression of the gene in tobacco leads to a pronounced leaf shape modification. This was a candidate phenotype because BBR is now credited to interact with homeobox partners and, as it is well known, the modification of the level of expression of homeobox class I genes leads to new leaf phenotypes in plants (Chan et al., 1998; Lin et al., 2002; Muller et al., 1995).

All evidence suggests that BBR acts transcriptionally to control gene expression mediated by (GA/TC)$_n$ repeats, with $n$ equal to 8 but not excluding 6 and 7. These repeats are well represented in plant genomes, particularly in gene promoters. GA/TC-repeat-binding proteins are known in Drosophila as GAGA or GAF proteins (Soeller et al., 1993). GAF is encoded by the trithorax-like gene (trl; Farkas et al., 1994), is required for the expression of some homeotic genes (Bhat et al., 1996), and is associated with heat shock and histone gene promoters (Biggin and Tjian, 1988; Gilmour et al., 1989; Soeller et al., 1988). The natural target promoters for GAF contain multiple GAGAG-binding elements (Katsani et al., 1999). Polycomb response elements (PREs), which participate in the maintenance of silencing of homeotic loci in Drosophila, also contain GAGAG elements to which GAF (Busturia et al., 2001; Hodgson et al., 2001; Wilkins and Lis, 1999) and Psq, a second Drosophila protein (Lehman et al., 1998), bind. Neither GAF nor Psq has sequence homology to BBR. Extensive BLAST searches indicate that BBR is present only in plants and Trl and Psq only in insects. This is surprising because the system of gene activation-repression that interferes with chromatin re-modelling – based in Drosophila on gene groups PcG, trxG and trl – is represented by PcG-like (Goodrich et al., 1997; Grossniklaus et al., 1998) and trxG-like (Alvarez-Venegas and Avramova, 2001) genes in plants, but apparently not by genes encoding GAF or Psq-like proteins. The possibility exists that BBR is a functional plant substitute of GAF or Psq. If so, BBR should be present in protein complexes binding plant promoters with single or multiple (GA/TC)$_n$-binding sites, as well as in plant PRE-like Polycomb response elements. These silencing elements contain binding sites for PHO – a protein recruiting PcG-encoded molecules – and GAF proteins (Busturia et al., 2001; Hodgson et al., 2001).

When the Mu transposable element inserts into the large intron III of the maize homebox gene Knotted1, dominant mutations caused by the overexpression of the gene are observed (Greene et al., 1994). This situation is quite similar to the case of the intron IV Knox-3 modification leading to the K mutant phenotype. All nine Mu insertions recorded for intron III of maize Knotted1 cluster in a 310 bp (out of the 404 published so far) DNA fragment (A in Figure 8), a region which shows 76% identity to the corresponding A region of the barley BKn3 intron IV. Four out of the nine Mu insertions separate a (GA/TC)$_n$ element from a PHO site located 173 bp 5’ upstream. Further sequencing data derived from maize intron III will reveal if the distribution of GA/TC and PHO elements holds true for the remaining part of the region A. Several hypotheses have been discussed concerning the mechanisms responsible for the Knotted1 overexpression, including the knock-out of silencers present in intron III. Regions A and B, in barley and maize, have in common a (GA/TC)$_7$-repeat and PHO-binding sites (ATGGC). More precisely, the investigation of the barley 305 bp DNA sequence revealed the presence of a PHO-binding site positioned 5 bp downstream of (GA/TC)$_8$ (Figure 8). The intron region which, in Hooded, is affected by the duplication, consists of a more 3’-located DNA sequence (B in Figure 8) and shows, in the mutant, a duplication of the

**Figure 8.** Schematic representation of intron IV of barley BKn3 and intron III of maize Kn1.

Presence of putative binding sites and insertions are highlighted. Region A is the only region for which the DNA sequence is available for both introns.

© Blackwell Publishing Ltd, *The Plant Journal*, (2003), 34, 813 826
two (GA/TC)_8 and PHO sites. The data presented open two intriguing possibilities: (i) the DNA sequence of the two introns contain PRE-like components involved in gene silencing; in this case, silencing should be mediated by the BBR protein; (ii) as proposed for the GAF gene in insects (Katsani et al., 1999), BBR, besides participating in protein complexes, could also act in promoter activation, as supported by its transcriptional activity observed in tobacco (Figure 1). When the role of BBR is sorted out, it should become possible to propose a model explaining both the mutagenic role of insertions in introns and promoters and the origin of dominant mutations as a result of gene derepression.

How do our new data increase the knowledge about barley homeobox gene regulation? The finding that part of the Bkn3 intron is transcriptionally active is important. This was predicted according to the data derived from an extensive chemical mutagenesis programme of a barley K line, which indicated that five K suppressor genes map away from Bkn3 on other chromosomes (unpublished data of author’s lab). Thus, such proteins should exist that bind to the Bkn3 intron IV. Regulatory control elements residing within a gene are not unusual, as evident from the studies in Drosophila, humans and mouse (Gremke et al., 1993; Hinz et al., 1992; Seul and Beyer, 2000; Shamsher et al., 2000). In plants, introns reported to enhance gene transcription are found in the Adh1-S (Luehrsén and Walbot, 1991), Sh1 (Clancy et al., 1994) and Lg3 (Jackson et al., 1994) maize genes, in Plena of Antirrhinum (Bradley et al., 1993), Agamous of Arabidopsis (Sieburth and Meyerowitz, 1997) and Lin5 of tomato (Fridman et al., 2000). However, the role of the barley Bkn3 intron IV seems particular in that a duplication-insertion of a 305 bp sequence leads to the dominant mutation Hooded.

Experimental procedures

Standard molecular protocols were followed in all construct assemblings (Sambrook et al., 1989), and all PCR based cloning were verified by sequencing. Sequences of the oligonucleotides mentioned in the paper are reported in Table S1.

Yeast one hybrid

The 305 bp fragment was amplified by PCR with primer pair 305f and 305r from genomic DNA of K Atlas and was inserted into the Clontech vectors pHIS1, pHIS3 (Clontech, Palo Alto, CA, USA) and pLacZ1. In YM4271 yeast strain, when the 305 bp were fused to the reporter gene, the basal activity of the leaky minimal promoters Phis3, in the cases of pHIS1 and pHIS3, and Pcycl of pLacZ1, was abolished. Nevertheless, pHISI allowed cloning of the 305 bp element 5’ upstream of the his3 reporter gene, giving rise to pHISI305. Because of the ura3 gene in pHISI305, the selection of YM4271 transformants become possible on minimal medium (MM) without uracil (Clontech manual, Palo Alto, CA, USA). The transformants were not able to grow on MM without histidine. One clone was sequenced, and its progeny was subjected to ethyl methane sulfonate (EMS) mutagenesis by treating cells with 3% EMS for 1 h. Twenty four mutant clones prototrophic for uracil and histidine were generated and grown on MM (uracil and histidine) at different concentrations of 3 AT. Growth was abolished step wise at 5 mM 3 AT, and the strain YM4271pHISI305 5, already blocked at 5 mM 3 AT, was chosen to screen the barley expression library. The 305 bp element cloned 5’ to the Pcycl minimal promoter in the vector pLacZ1, when inserted at the ura3 locus of S. cerevisiae, gave rise to the YM4271 1X305lacZ strain.

Barley library screening

The pADGAL4 plasmid (Stratagene, Amersham, NL, USA) was the acceptor of barley cDNAs. The barley yeast expression library was from mRNA purified from florets of K Atlas barley, as described by Müller et al. (2001). Approximately 3 million pADGAL4 trans formed derivatives of YM4271pHISI305 5 were screened on MM without uracil, without histidine and without leucine supplemented with 5 mM 3 AT.

Library screening and primer extension analysis

A cDNA library made from inflorescences of K Atlas and a genomic library from the mutant CalC15, described in Müller (1997), were screened (Benton and Davis, 1977). The longest cDNA BBrd (Figure 2a) was used as a probe. The BBrd primer extension experiment was performed according to Ausubel et al. (1994); 10 μg of total RNA isolated from young barley inflorescences (wt and K Atlas) were used.

Recombinant protein production and purification

A fragment encoding aa 25-350 of BBR, excised from clone BBrd, corresponding to bp positions from 2137 to 3114 of the gWBBR clone, was inserted in the pGEX 5 X 1 (Pharmacia, Freiburg, Germany) in frame with the ORF for glutathione-S-transferase (GST). The construct was introduced into E. coli BL21DE (Ausubel et al., 1994). After 4 h of induction with 1 mM IPTG at 37 °C, BBR was purified via affinity chromatography on glutathione agarose beads (supplier’s manual, Stratagene, Amsterdam, NL, USA). The same cDNA fragment was introduced in pBlueScript KS (Strata gene, Amsterdam, NL, USA) and was used as a template for in vitro transcription with T3 RNA polymerase (Roche, Mannheim, Germany). In vitro translation followed using wheat germ extract (Promega, Madison, WI, USA) in the presence of 35S methionine.

Electrophoretic mobility shift assay

Probes from the 305 bp element were amplified by PCR or synthe sized as oligonucleotides. Probes were labelled with the Klenow fragment (Roche, Mannheim, Germany) using 32P labelled nucleo tides or with T4 polynucleotide kinase (Roche, Mannheim, Germany) using 32P 5' P ATP. G1 buffer (0.5 mM EDTA, 10 mM Tris HCl, pH 7.5, 1 mM MgCl2, 50 mM NaCl, 4% glycerol, 50 ng 1-1 Poly (dI dC) and 0.5 mM DTT) was used for the protein DNA binding reactions. The DNA binding reaction in 1× G1 buffer included 100 ng of labelled DNA, or 1 50 pmol of labelled oligonucleotides, and 10 50 ng of purified BBR protein. The reaction took place at room temperature for 20 min, and products were separated by PAGE on 4% native polyacrylamide gel in 0.2× TAE at 4 °C for 8 h.

At first, EMSA experiments were based on the complete 305 bp element. Subsequently, several deletions were generated. Three partially (20 bp) overlapping fragments were amplified using PCR primers S1f, S1r, S2f, S2r, S3f, S3r. The fragment containing bp 180 305 of the 305 bp element was divided into four shorter
fragments having 10 bp of overlaps. Four pairs of complementary annealed oligonucleotides, labelled by \textsuperscript{32}P P-ATP, were used: M1\textprime{}, M1\textprime{}r; M2\textprime{}, M2\textprime{}r; M3\textprime{}, M3\textprime{}r; M4\textprime{}, M4\textprime{}r. The DNA fragment from bp 238 to 276 was further reduced using two 26 bp long complementary oligonucleotides M5\textprime{}f and M5\textprime{}r corresponding to bp positions 243 268. In the last step, the oligonucleotides M6\textprime{}f and M7\textprime{}, each one annealed together with its complementary sequence, were used.

Expression studies

Total RNA from wt barley, cv. Atlas, mature embryos, seedling leaves, seedling roots, internodes, nodes and inflorescences and from K Atlas inflorescences was extracted by the RNAsin plant minikit (Quiagen, Hilden, Germany). Poly(A)+ mRNA was prepared using the Oligotex Direct mRNA minikit (Quiagen, Hilden, Germany). One to two micrograms of poly(A)+ mRNA was used in Northern hybridization with a probe derived from the BBR\textalpha{} sequence. The blot was re-probed with an Antirrhinum actin cDNA fragment to control RNA loading. Quantitative RT PCR was carried out according to Varotto et al. (2002). The tissues analysed were root, leaf and inflorescence in wt and K Atlas. Three micrograms of total RNA was transcribed with 200 U of Superscript reverse transcriptase (Life Technologies, Karlsruhe, Germany). The pro duct of the reverse transcription was amplified by primers RTBBRf and RTBBRr.

BBR cellular localization

The full length BBR sequence, corresponding to bp position 2062 3114 of gWBBR, was PCR amplified from cSBBR and was used to produce fusion constructs with red fluorescence protein (RFP) in vector pGJ1425 (Jach et al., 2001) and with green fluorescence protein (GFP) in pGFP JS (Sheen et al., 1995). Following the pro tocol of Negrutiu et al. (1987), the constructs were transformed in tobacco protoplasts prepared from BY2 cells grown in liquid medium. Flourescent microscopy was performed with a Zeiss Axiophot microscope (Zeiss, Jena, Germany), and pictures were taken using a video image system consisting of a Hitachi CCD video camera (Hitachi, Tokyo, Japan).

GUS transgenics and GUS tissue staining

One, two or three copies of the 305 bp element subcloned into pUC19 were cloned 5' to the EcoRV NcoI CaMV 35S minimal promoter and were positioned upstream of the GUS reporter gene, generating constructs I, II and III as shown in Figure 2(a). In construct V, a 655 bp promoter fragment of Bkn\textprime{} was inserted in front of the U\textalpha{}d\textalpha{} gene. Construct IV, which was used as control, contained the CaMV 35S minimal promoter positioned upstream of U\textalpha{}d\textalpha{} gene (Figure 2a). All constructs were transferred in pBIN19 (Frisch et al., 1995) and were introduced into Agrobacterium tumefaciens LBA4404. Transgenic tobacco plants were generated by leaf transformation (Matsouka and Sanada, 1991). For histochem ical analysis, plant tissues were incubated at 37 °C in the presence of 1 mM X GLUC for 1, 3 and 12 h (see Results), fixed, de stained and stored in 50% glycerol at 4 °C (Jefferson et al., 1987).

Transient expression experiments

Two effector constructs were assembled, p101BBR (FullBBR) and pVPADBRR (ADFulBBR; Figure 5a). For FullBBR, the complete BBR ORF sequence was obtained by PCR amplification from cSBBR and inserted as a Smal Xba fragment in the pRT101 vector (Topfer et al., 1987). For ADFulBBR, the complete BBR ORF was PCR amplified and cloned in the Bspel site at the 3' end of the VP16AD sequence of the pBT4 35S NLS::VP16 expression vector (Sprenger and Weisshaar, 2000). The use of the herpes simplex virus VP16 activation domain is summarized for plants by Wilde et al. (1994). In both effector constructs, BBR was under control of the full length CaMV 35S promoter.

In the reporter vector Min1 (Figure 5a), the 305 bp element was cloned using two HindIII SacI restriction sites artificially introduced in the K373 plasmid 5' to the CaMV 35S minimal promoter upstream of the coding region of the U\textalpha{}d\textalpha{} gene. Further constructs were: K373, which does not contain BBR binding sites; the empty vector pRT101; and the vector ms129 pbt8 ubi-lucm3 (pLUC), which was co transfected to allow data normalization. The Photin pyrus pyralis luciferase gene (LUC) is under the control of the constitutive UB\textalpha{}I4/2 promoter (Kawai leck et al., 1993). Transient transformations of mesophyll protoplasts derived from Nicotiana tabacum cv. SR1 were performed according to Negrutiu (1987). GUS activity was determined fluorimetrically (Jefferson et al., 1987). Data were normalized based on the protein content (Brad ford, 1976) and on the luciferase values, as internal control (Luehrs en et al., 1992), because of co transfection with the ms129 pbt8 ubi-lucm3 vector (Sprenger and Weisshaar, 2000). Protoplasts were incubated for 48 h, collected and lysed by freezing in liquid nitrogen. Aliquots of the supernatant were used to determine GUS activity and protein content. GUS activity was estimated at 30, 60 and 120 min as $\Delta E_{420}$ increments (Jefferson et al., 1987) and was expressed as pmol 4 methylumbelliferone (MU) mg$^{-1}$ min$^{-1}$ when referred to the protein content, and as pmol MU x 10 000 RLU$^{-1}$ (Relative Light Unit) when referred to the luciferase values. Each experiment was repeated three times and four independent measures, in condition of linear increase of $E_{420}$, were taken for each time point. Average values were expressed as fold of activa tion based on the K373 values.

BBR overexpression in tobacco

Three different versions of the BBR gene were moved to tobacco. The first and second encoded 325 out of the 350 BBR aa. In the second, the yeast GAL4 activation domain sequence was added. Both versions were obtained by restriction digests from the BBRd clone and were introduced in the pRT101 vector (Topfer et al., 1987) generating constructs ADBBR and ADJABBR, respectively. The third BBR gene version encoded the complete BBR protein; the fragment (from bp 2062 to 3114 of gWBBR clone) was amplified by PCR and was cloned in the unique BamHI restriction site of pPCV91 binary vector (Strizhov et al., 1996). The expression cas settes of both pRT101 based constructs were transferred by HindIII restriction digest to pBl19 (Frisch et al., 1995) and were intro duced into A. tumefaciens LBA4404 (Hoeckema et al., 1983). Strain GV3101PMP90RR (Koncz et al., 1994) was used for the full BBR sequence cloned in pPCV91 (Strizhov et al., 1996). All constructs were used in SR1 tobacco leaf disc transformation (Matsouka and Sanada, 1991).

Mapping the BBR genetic locus

The BBR promoter region was amplified by PCR from genomic DNA of Proutc and Nudinka barley cultivars. The primers PBR1 and PBBR2 were used. The amplified PCR fragment differed in length in the two varieties, allowing the typing of 100 Prot cor x Nudinka double haploid (DH) lines (Heun et al., 1991). Typ ing data were added to the database available for the same segregating population. Mapping was performed using MAP MAKER/EXP3.0 (Lander et al., 1987), and the map position of BBR was entered in the AFLP/RFLP/ISTR linkage map of Castiglioni...
Bioinformatics
To calculate the number of (GA/TC)n repeats, the contiguous sequences for the A. thaliana chromosomes were retrieved from GenBank (NC 003070, NC 003071, NC 003074, NC 003075 and NC 003076) and were converted into continuous text file strings. Each chromosome sequence was scanned for motifs within a 1500 bp window. Dinucleotide (GA/TC)n repeats in exons and introns were identified using Patmatch from TAIR (http://www.arabidopsis.org). To obtain a data set 1500 bp before the ATG of each gene, 25 540 AGI sequences were retrieved by choosing upstream output from −1 to −1500 available at Regulatory Sequence Analysis Tools (RSA; http://embnet.cifn.unam.mx/rsa tools). All data sets were investigated for repeats of (GA/TC)n-12. Single loci of (GA/TC)n were estimated by subtraction of (GA/TC)n-1 repeats. The probability of the (GA/TC)n repeat is based on the Arabidopsis genomic frequency weights for G/C (0.18) and A/T (0.32) nucleotides. The observed frequencies (base pair occurrences) were calculated by dividing the set length in base pairs by the number of observed loci. The length of each set is: total genome, 234.9 × 10^6 bp; exons, 37.2 × 10^6 bp; introns, 25.4 × 10^6 bp; promoters (1.5 kb per gene), 67.6 × 10^6 bp; and intergenic sequences, 95.7 × 10^6 bp. Gene homology comparisons were performed using the BLASTP and TBLASTN algorithms available from NCBI (http://www.ncbi.nlm.nih.gov) and by selecting individual kingdom data sets.

Acknowledgements
We would like to thank Paul J. Ruashon for his suggestions and corrections on the manuscript, Kurt Stueber for sharing his unpublished bioinformatics results and Margrit Pasemann for helping in the editing of the paper.

Supplementary Material
The following material is available from http://www.blackwellpublishing.com/products/journals/supportmat/TPJ/TPJ1767/TPJ1767sm.htm
Table S1 Oligonucleotides
Table S2 Complete set of data of number and frequency of GA/TC repeats in the Arabidopsis genome

References

© Blackwell Publishing Ltd, The Plant Journal, (2003), 34, 813 826


© Blackwell Publishing Ltd, The Plant Journal, (2003), 34, 813–826


EMBL accession number gWBBR: AJ507214 (BBR genomic sequence).