Intron mediated regulation of *BKn3*, a plant homeobox gene, as supported by BBR (GAGA) binding factor

Inaugural-Dissertation

Zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Luca SANTI

aus Genoa, Italy

Köln, 2002

Die vorliegende Arbeit wurde am Max-Planck-Institut für Züchtungsforschung, Köln-Vogelsang, in der Abteilung Pflanzenzüchtung und Ertragsphysiologie in der Arbeitsgruppe von Prof. Dr. F. Salamini angefertigt.

Berichterstatter: Prof. Dr. Francesco Salamini
Prof. Dr. Martin Hueskamp
Prof. Dr. Siegfried Roth

Tag der mündlichen Prüfung: 11. 06. 2002

To my parents Leonardo and Lia

"Dicerolti molto breve.
Questi non hanno speranza di morte,
e la lor cieca vita è tanto bassa,
che 'nvidiosi son d' ogni altra sorte.
Fama di loro il mondo esser non lassa;
misericordia e giustizia li sdegna:
non ragioniam di lor, ma guarda e passa."

Dante Alighieri
Divina Commedia, inf. III

Abbreviations

Amp	Ampicillin
a.a.	Amino acid(s)
APS	Ammonium persulphate
Bp	Base pair
BSA	Bovin Serum Albumin
CDNA	Complementary deoxyribonucleic acid
C-TAB	Cetyltrimethylamonium bromide
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
Dnase	Deoxyribonuclease
DTT	Dithiothreitol
E.coli	Escherichia coli
EDTA	Ethylenediamine tetraacetic acid
EtBr	Ethidium bromide
EtOH	Ethanol
HEPES	4-(2-hydroxiethyl)-1-piperazinethanesulfonic acid
IPTG	Isopropylthio-β-o-galactopyranoside
Kb	Kilo base (s)
Kd	Kilo Dalton (s)
Min	Minute(s)
MOPS	3-(N-morpholino)-propanesulfonic acid
mRNA	Messenger RNA
MUG	4-methylumbelliferyl-d-glucoronide
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PMSF	Phenylmethanesulfonyl fluoride
RNA	Ribonucleic acid
rpm	Rotation per minute
RT-PCR	Reverse transcription polymerase chain reaction
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
U	Unit(s) enzymatic activity

Contents

1 Introduction

1.1 Gene transcription and transcription factors

1.1.1 Generalities on transcription

(Paule and White, 2000; Pisarchik and Kartel, 2000; Lee and Young, 2000; Ansari, 2001)

The genetic information is expressed by transcription of segments of DNA into two classes of RNA: labile messenger RNA (mRNA), which is translated into proteins and stable RNAs (ribosomal, transfer, and small RNAs), which are not translated. Genetic and molecular studies have revealed a number of mechanisms regulating gene expression in bacteria and these findings have provided a foundation for the more intricate regulation of genes in higher organisms. Physiologic transcription generally reads only one strand in a particular sequence of DNA, but different segments may be transcribed on opposite strands (and consequently in the opposite direction). RNA is synthesized by a DNA dependent RNA polymerase. Purification of RNA polymerase from *E. coli* yields both a core enzyme and a holoenzyme. The process of transcription involves three main steps: initiation, chain elongation and termination.

The processes of initiation requires the holoenzyme, the proper nucleoside triphosphate and a special promoter DNA sequence. Promoters differ in their sequence and in their strength of binding the polymerase; in cases, binding requires an activator protein binding immediately upstream. With respect of initiation of transcription at +1 promoters have similar but not identical sequences in two regions, -35 and –10. The “consensus” sequence at –35 is TTGACA, while that at –10 is TATAAT, whose AT pairs (weaker than GC) promote the required melting.

In the elongation phase the core enzyme assumes a new conformation by interacting with several factors. The transcription bubble moves from one to the other end of the transcription unit, which can be several thousand bases long. Nucleotides are added to RNA at a rate of 30 to 50 bases per second. At certain sites, however, the polymerase makes pauses that can last many seconds.

At termination, the polymerase drops off and the nascent RNA chain leave the DNA template. This process is subjected to extensive modulation, which is important in regulating gene function, but it is not as well understood as initiation of transcription.

In eukaryotic cells three types of RNA polymerases exist. Ribosomal RNA is transcribed by RNA polymerase I, messanger RNA by RNA polymerase II and tRNAs (and other small RNAs) by RNA polymerase III. RNA polymerase II cannot initiate transcription itself but is absolutely dependent on

auxiliary transcription factors and together they form the basal trascription apparatus. Also in eucaryotes a TATA box is present in most promoters but located ~ 25bp upstream of the start point. It constitutes the only upstream promoter element that has a relatively fixed position. The 8bp consensus sequence consists entirely of A-T base pairs. A minority of promoters that do not contain a TATA element are called TATA-less promoters. To achieve transcription, the first step is the formation of the basal transcriptional machinery around the TATA box region. The machinery is composed by the RNA polymerase II and by at least five basal transcription factors (TFIIA, TFIIB, TFIID, TFIIE and TFIIF). Promoters do not necessarily function alone. At least in some cases, the activity is increased or decreased by the presence of an enhancer or silencer, respectively. They can be relatively distant from the promoter and can function in both orientations. How an enhancer works is not yet clear but several hypothesis, not mutually exclusive, have been proposed. An enhancer could: 1) change the overall structure of the template, for example influencing the density of supercoiling. 2) It could be responsible for locating the template at a particular place within the cell, for example attaching it to the nuclear matrix. 3) An enhancer could provide an "entry site ", a point at which RNA polymerase or some other essential proteins associate with the chromatin.
Cis-elements are short sequences present in regulatory regions, where specific transcription factors bind. The sequence between them is not important, although the extent of their separation might be.

1.1.2 Transcription factors

An abundance of highly variable transcription factors exists which can bind *cis*-elements in promoters and enhancers. These factors are active together with the basal machinery to establish the fine regulation of gene expression. The difference between constitutive expression and specific expression at a particular stage of development or in response to given stimulus, lies in the interaction of transcription factors with *cis*-acting elements. Several transcription factors have been studied structurally (examined by X-ray crystallography and NMR spectroscopy) and in their function. Constituve domains proper of a particular transcription factor protein are important to understand their transcription function.
The domains of plant transcription factors frequently defined by comparing amino acid sequences, deduced from cDNA clones, with their animal or yeast counterpart. Study of putative functional domains by mutational and functional analysis has demonstrated that typical plant transcription factors consist of a DNA-binding region, an oligomerisation site, a transcription regulation domain and a nuclear localisation signal (NLS). Although some transcription factors may lack one or more

of the above described domains, others present novel regions like the membrane spanning region in PEND (plastid envelope DNA-binding; Sato *et al.*, 1998).

The possibility of grouping transcription factors in more or less homogeneous families depends on their structural features, with groups sometimes subdivided according to the number and spacing of conserved residues in the most similar domains. For instance, according to the quantity and arrangements of cysteine (C) and hystidine (H) residues, the factors containing zinc fingers fall into five classes: C_2H_2; C_3H; C_2C_2 (GATA finger); C_3HC_4 (ring finger) and C_2HC_5 (LIM finger) (Sakamoto *et al.*, 1993). Alternatively, transcription factors of the same family are categorised using regions outside the most conserved domain (see homeobox genes further in the introduction).

The DNA binding domain of transcription factors, many of which are basic in character, contain amino acid residues that contact DNA bases at *cis*-acting elements, and these residues determine the specificity of the protein. Other residues enhance transcription factor binding by contacting the DNA less specifically through interaction with either phosphate or deoxyribose moieties. The base recognition moieties are often highly conserved. Secondary structure in DNA-binding domains seems to affect their affinity and selectivity. For example, the C-terminal DNA binding domain of the rice trihelix factor, GT-2, loses its activity when helix-breaking prolines are substituted for other amino acid residues (Ni *et al.*, 1996).

Transcription factors can interact to form hetero- and/or homooligomers, affecting DNA binding specificity and the affinity of transcription factors for promoter elements and nuclear localization. Oligomers are stabilised by hydrophobic interactions between coiled coiled and β-sheets, or by interactions between hydrophilic residues. It is known, for example, that GLABRA2-like proteins, belong to a family of homeodomain proteins that bind DNA just as dimmers (Di Cristina *et al.*, 1996).

Transcription factors of the same family have distinct actions because of differences in their regulation domains. Regulation domains, and hence transcription factors, function as repressors or activators, depending on whether they inhibit or stimulate the transcription of target genes. Repression of gene expression may occur via exclusion of activators from target promoters by competitive binding between transcription factors for the same *cis*-element, or through processes of chromatin remodelling. Activation domains of plant transcription factors often have significant sequence divergences, although all of them are rich in either proline or glutamine. Site-directed mutagenesis of residues in the activation domain of the maize C1 Myb-like transcription factor demonstrated that only one of 11 acidic residues, namely aspartate 256, is essential. This indicates that single strategically placed residues determine activation (Sainz *et al.*, 1997).

As for proteins of other organisms that selectively enter the nucleus, plant transcription factors contain nuclear localisation sequences (NLSs) characterized by a core peptide enriched in arginine (R) and lysine (K). The basic core can be functionally influenced by flanking residues. NLSs can be single, when the basic residues are closely associated, single bipartite when residues form two functionally important groups separated by several non conserved residues and they can also be in several copies, functioning independently, dispersed within the protein. For example, in the maize homeodomain protein KNOTTED1 and in the maize bZIP factor Opaque2, two functionally independent NLSs have been found, either located within or exterior to the DNA-binding domains (Varagna *et al.*, 1994; Meisel and Lam, 1996).

A genome wide analysis of *A. thaliana* has revealed that 1533 transcription factors are present, comprising 5.9% of the predicted 26000 proteins. As a comparison, *D. melanogaster* has 635 transcription factors (TFs) in a total of 14000 proteins (4.5%), while the unicellular yeast *S. cerevisiae* has 209 TFs in 6000 proteins, (3.5%) (Tab. 1.1) (Riechmann *et al.*, 2000).

Domain type	*A.t.*	*D.m.*	Structure	Reference
Homeodomain	89	103	60 amino acids protein domain that folds in tree α-helices separated by a loop and a turn	(Schindler *et al.*, 1993)
MADS	82	2	57 amino acids domain that comprise a long α-helix and two β-strands	(Pellegrini *et al.*, 1995)
HMG-box	10	21	L-shaped domain consisting of three α-helices with an angle of about 80° between the arms	(Grasser *et al.*, 1995)
ARF	23	0	350 amino acids conserved region at the C-termini	(Ulmasov *et al.*, 1997)
Zinc finger	242	300	Finger motif(s) each maintained by cysteine and/or histidine residues organized around a zinc ion	(Sakamoto *et al.*, 1993)
trihelix	28	0	Basic, acid and proline/glutamine rich motif which forms a trihelix DNA-binding domain	(Kuhn *et al.*, 1993)
b-Zip	81	21	Basic region with a leucine-rich zipper-like motif	(Mikami *et al.*, 1994)
AP2/EREBP	144	0	68 a.a. region with a conserved domain that constitutes a putative α-helix and two β-strands	(Jofuku *et al.*, 1994)
B/HLH	139	46	A cluster of basic amino acids residues adjacent to a helix-loop-helix motif	(Sainz *et al.*, 1997)

Tab. 1-1 Transcription factors grouped based on their DNA-binding domain. Comparison between *A. thaliana* (*A.t.*) and *D. melanogaster* (*D.m*) based on genome-wide sequencing data.

1.2 Chromatin structure and regulation of transcription

Packaging of genes into higher order chromatin structure is thought to affect transcription by impeding access of transcription factors to regulatory sequences. Activation of expression requires that the chromatin packaging be disrupted, and regulation of this disruption event is a key role in gene regulation (Felsenfeld, 1996). The most basic level of eukaryotic chromatin structure, the nucleosome, is relatively well understood as a result of many years of biochemical and structural studies. Modifications, such as histone acetylation and phosphorylation, which affect nucleosome structure, have also been directly linked to transcriptional regulation and cell cycle control (Strahl and Allis, 2000). Several different models have been proposed to describe higher order chromatin packaging above the level of the nucleosome (Belmont *et al.*, 1999). Studies of the role of higher order chromatin organization in gene regulation have used a variety of different approaches. A number of studies have shown that accessibility of the DNA to nucleases and other modifying agents is increased in the region surrounding actively expressing genes (Felsenfeld, 1996) and the increased accessibility has been shown to correlate with hyperacetylation of histones (Hebbes *et al.*, 1994). The puffs that colocalize with expressing loci in *D. melanogaster* polytene chromosomes also suggest that chromatin decondensation event is associated with gene activation (Udvardy *et al.*,1985). These observations have led to the suggestion that domains of active decondensed chromatin form the functional unit of gene regulation (Dillon and Sabbattini 2000). By examining the structure of the activated domain, these studies focus mainly on the end point of transcriptional activation. However, an understanding of how large-scale changes in chromatin structure are related to gene regulation requires that the multiple events that lead to these changes be dissected in detail. In particular, it is important to gain a better understanding of how the information content, that resides in the profile of transcription factors in a given cell type, is translated into the changes in chromatin structure that lead to gene activation.

The clearest difference observed between the various types of chromatin in the nucleus, is the division between heterochromatin and euchromatin (Wallrath, 1998). Constitutive heterochromatin remains condensed during interphase, replicates late, is rich in repetitive sequences and contains relatively few transcribed genes. Heterochromatic DNA tends to be highly methylated. Eucromatin is more decondensed in interphase cells, replicates early and has a higher density of functional genes. The inhibitory effect of heterochromatin on gene expression is well documented. In addition to constitutive heterochromatin, a second class of heterochromatin, termed facultative heterochromatin, is found for cases like the inactive X chromosome in mammals and the homeotic loci are stably silenced by the action of the polycomb group (PcG) proteins (Brockdorff, 1998).

Translocations that place euchromatic genes close to pericentromeric heterochromatin may lead to silencing or position effect variegation (PEV) (Henikoff, 1990). Genetic studies in Drosophila have identified a number of modifier genes that affect PEV, observations have recently been extended to variegation in mice (Milot *et al.*, 1996). The products of modifier genes have been hypothesized to bind to heterochromatin, forming specific DNA-protein complexes containing multiple components. According to one model, PEV would be caused by variable spreading of the complexes along the chromatin fiber (Locke *et al.*, 1988). An alternative model proposes that the protein components of heterochromatin bind to multiple binding sites along the fiber with variegation occurring as a result of direct contact between euchromatic genes and heterochromatin through looping (Henikoff, 2000). While these models are usefull for considering the effects of modifiers, they do not address the issue of how variegation and formation of heterochromatin protein complexes are related to the chromatin condensation observed by cytogenetic approaches. The link between PEV and silencing by the PcG proteins involves the transmission of the repressed state through mitosis. The hypothesis that heterochromatization could be involved in establishing cellular memory of gene expression patterns in different cell lineages has received further support from experiments showing that silencing of genes located on the long arm of the chromosomes is frequently associated with colocalization of the silenced gene with pericentromeric heterochromatin in interphase nuclei (Cockell, 1999). Progressive heterochromatinization of most of the genome is also associated with terminal differentiation of diverse cell types which include glial cell, antibody producing plasma cells and reticulocytes.

Different chromatin regulators may be broadly conserved is supported by the discovery of the SET domain genes, an emerging, well conserved gene family encoding proteins with diverse, chromatin-based biological functions. The SET domain, a motif of about 130 amino acids, was first recognized as a common element present in Drosophila genes: *SU(VAR)3-9*, *E(Z)* and *TRX*. The domain was subsequently found in a number of eucaryotic proteins from yeast to mammals and to plants. This family of proteins was divided into four subgroups according to the sequence conservation within their SET domain. The first subgroup is represent by E(Z), which is a member of the polycomb group proteins involved in the maintenance of repressive transcriptional state of *HOM-C*. The second subgroup includes TRX, which belongs to the trithorax group and functions as an antagonist of polycomb group proteins in activating *HOM-C* expression, and the saccharomyces cerevisiae SET1, which affects mating-type switching and telomeric silencing. The third group comprises the *D. melanogaster* trithorax protein ASH1, and S. *cerevisiae* SET2, which appear to ensure a basal repression of genes at euchromatic positions. The mammalian nuclear receptor binding protein NDS1 also belongs to this third subgroup. SU(VAR)3-9, together with the human ankyrin-repeat-

containing protein G9a and *S. pombe* protein CLR4, constitute the forth group. SU(VAR)3-9 is the strongest (PEV) suppressor in *D. melanogaster*, and CLR4 is involved in mating-type silencing and centromere function in *S .pombe.* In yeast, mutations in the SET domain of CLR4 and SET1 disrupt centromeric silencing in *S. pombe* and telomeric silencing in *S. cerevisiae*, respectively, indicating a function for the SET domain in silencing. In mammals, it has been shown recently that the human hortolog of SU(VAR)3-9 is differentially phosphorylated during the cell cycle and it has been suggested that the SET domain may function as a gatekeeper motif in integrating upstream signalling pathways in epigenetic regulation and growth control.

The first SET domain described for plants was found in *CURLY LEAF (CLF)* and is required for repression of transcription of the floral homeotic gene *AGAMOUS* in *Arabidopsis* leaf, inflorescence, stem and flower (Goodrich *et al.*, 1997, see next session of the introduction). Another SET-domain containing gene described for plants is MEDEA/MEA, a gene whose maternally derived allele is required for *Arabidopsis* embryogenesis. Both CLF and MEDEA belong to the first subgroup of the SET domain protein family and show highest homologies with E(Z).

The *medea* mutant shows maternal embryo lethality. In flowrering plants embryo development is affected by both the female gametophyte and the sporophytic tissue of the parent plant, so the survival of the resultant embryo depends on the presence of a wild-type MEA allele in the genome of themother plant. These facts, together with its high homology to the SET domain, suggests that MEA controls cell proliferation by regulating gene expression through modulation of higher chromatin structure (Grossniklaus *et al.*, 1998).

Apart from SET domain containing proteins, additional evidence suggests that in plants, the chromatin structure influences transcription. New regions of DNaseI hypersensitivity are formed *in vivo*, upstream of the maize *Adh1* and *Adh2* genes, as well as in the maize *Shrunken* gene; the maize *P* gene; the pea *RbcS* genes and the *Arabidopsis Adh* gene, as the expression of the gene increases. These induced hypersensitivity sites suggest that the nucleosome arrays in these regions are disrupted upon transcription and that they are likely to be the binding sites of transcription factors. In wheat, DNaseI preferentially digests transcriptionally active sequences, suggesting that these sequences assume open chromatin structures (Spiker *et al.*,1983). Moreover, the nucleosome present on the β-phaseolin promoter is removed upon gene activation in transgenic tobacco and nucleosomes are partly responsible for the closed conformation of a silent methylated *A1* transgene in petunia. Therefore, nucleosome positioning and arrangement play a role in plant gene transcriptional regulation.

In recent work the group of Y.L. Chua has demonstrated that transcriptional regulation of the pea plastocyanin gene (PetE) is established by hyperacetylation of both histones H3 and H4 in the

enhancer/promoter region of the gene in green shoots indicating that just specific nucleosomes along the gene were modified (Chua 2001). Acetylation of histones involves the transfer of acetyl groups from acetyl-CoA to the ε-amino groups of K9 or K14 of histone H3, or K5, K8, K12, or K16 of histone H4 by histone acetyltransferases (HATs). HATs have been classified into type-A a or type-B enzymes. Type B HATs are cytoplasmatically localized, whereas type-A are nuclear localized and are involved in regulating transcription, although their involvement in other processes such as DNA repair and replication is also likely. Both histone acetylation and deacetylation occurs in plants and it is likely that these processes regulate similar functions to those identified in other eucaryotes. The most extensively acetylated histone in plants is H3, in contrast to the situation in other eukaryotes where H4 is most highly acetylated (Waterborg, 1990). In plants, genes involved in the process of histone deacetylation have been cloned and analysed for their activity. The phenomenon of histone acetylation has been mainly addressed via biochemical studies of purified complexes and the analysis of histone acetylation patterns is based mainly on isolated chromatin components.

1.3 Homeobox genes

Homeodomain containing proteins are transcription factors that control the expression of target genes through specific binding to DNA elements. They are responsible for important developmental processes of multicellular organisms ranging from animals to plants.

Historically, the first studies on homeobox genes came from the fruit fly *D. melanogaster*. In particular, mutants for abnormal segmentation identity like *Antennapedia* (*Antp*), in which a pair of legs develops instead of the antenna, or mutants affected in the number and polarity of embryonic segments like *fushi tarazu* (*ftz*. In Japanese: not enough segments) or *engrailed* (*en*. Every segment lacks the posterior compartment), have helped in defining the role, importance and function of homeobox genes. Characterising the *Antp* gene, the groups of Gehring W.J. and Scott M.P. independently discovered that the *Antp* cDNA hybridised with both the *Antp* gene and with the *ftz* gene revealing that they share a common DNA sequence (Kuroiwa *et al.*, 1984; Scott *et al.*, 1984). This observation led to the discovery that many of these sequences are present in the *D. melanogaster* genome, particularly in homeotic gene complexes like *ANT-C* and *BX-C*. Other hybridisation studies, using homeobox genes as probes, led to the discovery that these kind of genes are present in many segmented organisms belonging to the animal kingdom, from annelids, to vertebrates like *Xenopus*, mouse and human. Later their presence was extended to the plant kingdom. In some of these sequences the degree of identity is very high, for example, the *Antennapedia* gene from *D. melanogaster* and the *MM3* gene from *X. laevis* encode polypeptides

that have 59 out of 60 amino acids in common. The fact that vertebrates and invertebrates diverged more than 600 million years ago indicates that the genic products of the homeobox domain have an essential function. In this domain, around 30% of the amino acids are basic (Arg + Lys) and this suggested from the beginning, a putative DNA binding property. Also sequence comparison and NMR studies have suggested nucleic acid binding capabilities, as these domain fold into helix-turn-helix structures that resemble motifs present in regulatory genes of procaryots, like the *trp* repressor from *E. coli* or the Cro protein of the λ phage. The empirical proof of DNA binding arrived with the discovery that the genic product of the *engrailed* (*en*) gene was able to bind the regulatory region of *en* and of *ftz* (Ish-Horowicz *et al.*, 1989). These observations prompted the idea that homeobox containing genes encode transcription factors and that these developmental genes act as regulators in the control of the expression of other genes.

In recent years, homeobox genes have been well characterized, even in mouse, where there are more than 30 genes divided in four groups. Each group is present on a different chromosome and their members have been named *Hox* genes. *Hox* genes, like the homeobox genes of *D. melanogaster* follow precise patterns of expression in particular stages during embryogenesis. They are responsible for homeotic mutations, as demonstrated by the constitutive overexpression in transgenic mice of the *Hox-1.1* gene, which led to a strong morphogenetic effect, transforming one body part into a second similarly to what has been observed in *D. melanogaster* (Mahon *et al.*, 1988). Humans have four groups of Hox genes as well, with nine or more genes in each group (Acampora *et al.*, 1989).

The crystal structure of the homeodomain has been resolved (Kissinger *et al.*, 1990). The DNA sequence consists of 180bp and encodes a 60 amino acid protein domain, that folds into three α-helices separated by a loop and a turn. The two helices separated by the turn give the characteristic helix-turn-helix (HTH) conformation that shows similarities with other procaryotic regulators before (Fig.1.1). DNA recognition is established by helix III, which lies in the major groove of DNA, and by the N-terminal flexible arm and the loop between helices I and II. Most DNA sequences that are bound efficiently by homeodomains contain the ATTA (TAAT in the complementary strand) core, which interacts with the highly conserved amino acids.

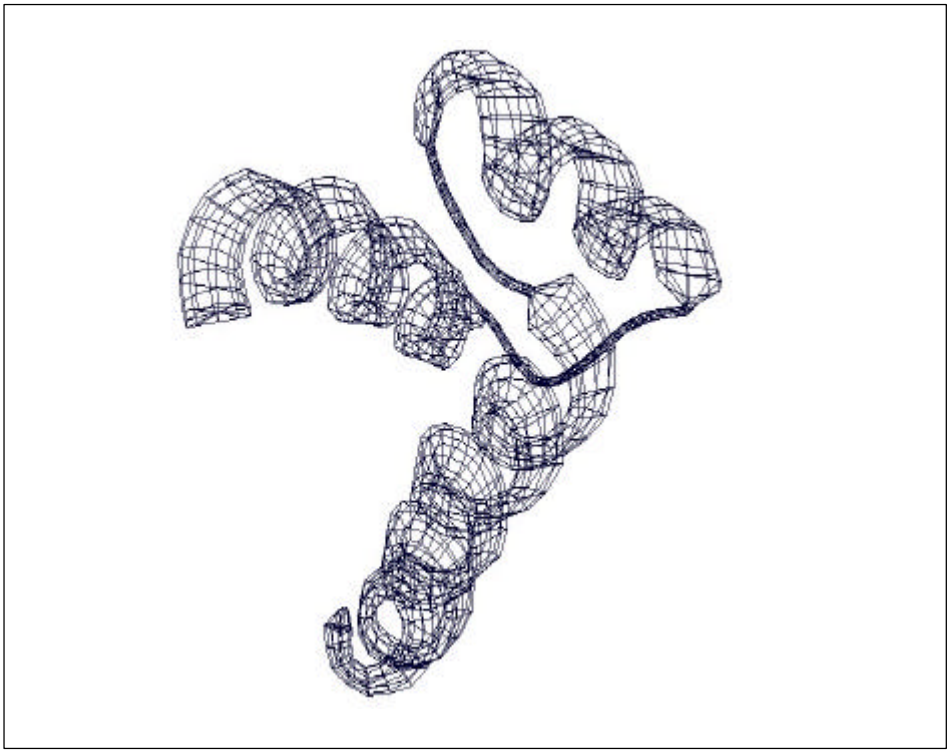


Fig. 1.1 Representation of the helix-turn-helix-loop-helix structure of the homeodomain of the Antennapedia (AHD1) homeobox gene from *D. melanogaster melanogaster* (Fraenkel and Pabo, 1998).

1.3.1 Plant homeobox genes

Since 1991 when the first plant homeobox gene was discovered by M. Freeling and colleagues in maize (Vollbrecht *et al.*, 1991), several others have been cloned in many plant species using different approaches, such as trasposon tagging in known mutants, comapping strategies and yeast one- and two-hybrid systems.

In the model system *A. thaliana*, the complete sequence of the genome reveals the presence of 89 genes coding for homeodomain proteins. In the attempt to understand more about the role of homeobox gene in plants they have been classified in to different classes of homologies for their homeodomain and for other characteristic regions outside the homeodomain. It is generally accepted that plant homeobox genes can be divided in 4 different families; some of these genes can be grouped into super families and sub grouped in sub families.

TALE superfamily

TALE means **t**hree **a**mino acid **l**oop **e**xtension; these three extra amino acids are positioned between helix1 and helix2. In plants, the *kn1* and *bell* family belong to this superfamily. TALE classes have been also identified in animals and fungi.

Kn1 family

Typical examples for monocot plants are *Kn1* from maize and *Bkn3* from barley, both affecting leaf development (Vollbrecht *et al.*, 1991; Mueller, *et al.*, 1995). KN 1 family members from dicots include the STM gene from *A. thaliana* and the LE T6 gene from tomato (Long *et al.*, 1996; Chen *et al.*, 1997), the former being involved in meristem maintenance, while the latter is again responsible for an abnormal leaf phenotype.

KN 1 family proteins are around 400 amino acids long and the homeodomain is mainly located at the C-terminal end of the protein. External to the homeodomain at the N-terminal, there is another region of homology, the so-called ELK domain, which follows after the first three amino acids. According to sequence homology and expression pattern, KN 1 the family has been further sub grouped in classI and classII (Reiser *et al.*, 2000).

Bell family

In BELL proteins, the homodomain is also located at the C-terminus end of the protein. Outside the homodomain, no distinct features exist, exept for the presence of regions rich in serine/threonine and proline which may function as transcription-activating domains. In dicots, BELL1 gene has been cloned from *A thaliana*, which have been shown to affect ovule development. From barley, JUBEL1 and 2 genes have been cloned based on their properties of binding the KNOX protein (Mueller *et al.*, 2001).

HD-Zip family

In these proteins, a leucine zipper domain is located at the C-terminal side of the homeodomain, forming an HD-Zip domain that lies approximately in the central region of the protein. Several members belong to this family with display very different functions ranging from stress response, to epidermal cell differentiation (Masucci *et al.*, 1996). They have been further divided into four sub groups.

PHD finger family

The cysteine rich region N-terminal of the homeodomain is what defines this family; the real function of this domain is still unclear, but many hypothesis have been proposed and it may function as a protein-protein interaction domain; as a regulatory domain, or as a DNA binding domain. This last hypothesis seems not correct, at least in the case of PRHA, as deleted forms of this protein lacking the PHD finger still specifically bind DNA (Plesch *et al.*, 1997).

GLABRA2-like proteins are not considered by some authors a sub-group of the HD-Zip family, but as an independent family (Chan *et al.*,.1998).

The identification of two proteins from *Poplus*, has suggested the presence of a new family: PALE (**p**enta **l**oop **e**xtention) after five extra amino acids present between helix 1 and 2 of the homeodomain (Hertzberg and Olsson, 1998).

1.4 The dominant *Hooded* mutant and the role of KNOX3 in leaf development

The floret of barley (*Hordeum vulgare*) is protected by two leafy organs, the lemma and the palea. The lemma together with subtending glumes and carpels are organs homologus to leaves that can develop or be repressed in specific regions of the plant. The homology of these organs is based on morphology, position and histogenesis and has also been revealed by the effect of several mutants. For example, *leafy lemma* (*lel*) induces the transformation of this organ into a rudimentary leaf with a reduced, but clearly discernable sheath and blade. The upper part of the lemma expands into a long distal appendage called the awn. Awns are found in all grass inflorescences and are usually determinate structures (Fig. 1.2a).

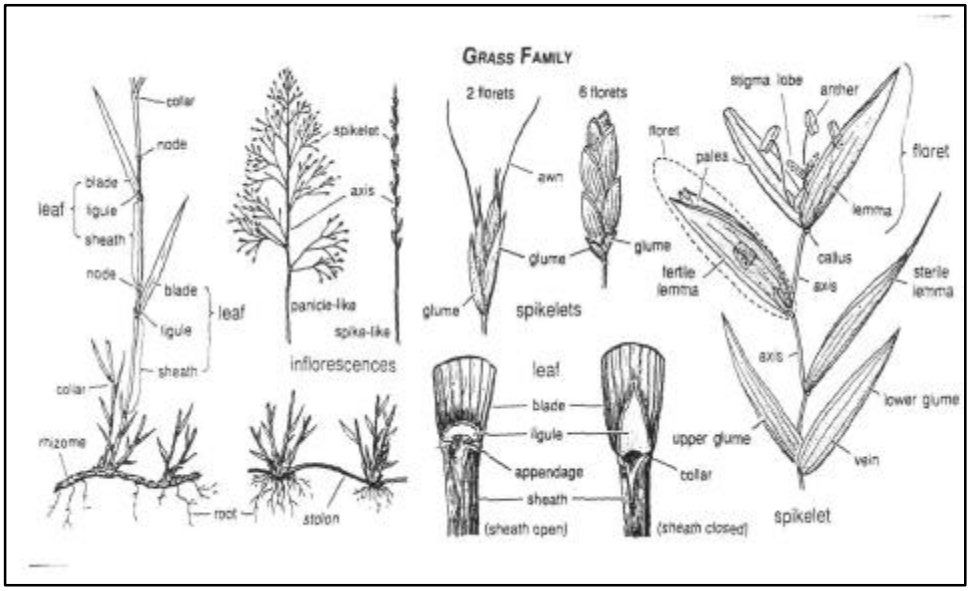


Fig. 1.2a Morfology of organs and related terminology discussed in the text

Hooded is a dominant mutation that causes the ectopic growth of an extra floret on the lemma. This second floret develops in place of the awn on the distal part of the lemma, with an inverse polarity compared to the first one. In this region, the mutant differs from the WT in the cell size of the adaxial epidermis and in the direction of epidermal cell division (Fig. 1.2b). This phenotype is governed by the single dominant locus *K*, that maps on chromosome 4. This locus has been associated with *Bkn3*, a homeobox gene that belongs to the kn1 family, class 1. It has been demonstrated that the difference at the molecular level, between the WT and the mutant allele, is caused by the duplication of 305bp fragment in the large fourth intron of *Bkn3* (Mueller *et al.*,1995). This mutation modifies the expression of *Bkn3; in situ* hybridisation in the mutant shows that the transcript is particularly abundant in the cushion, a region of reacquired meristematic activity, from which the extra flower develops (Fig. 1.3). *Bkn3* is highly homologus with *Kn1* from maize, the first plant homeobox gene identified. In the *Knotted* phenotype the mis-expression of *Bkn3* interferes with the normal differentiation of cells around the lateral veins of the leaf. Affected cells proliferate to form sporadic outgrowths –knots- of tissue which assumes more basal characteristic of the leaf, like cells of auricle or sheath. In *Kn 1*, as in *Hooded*, the gain of function is produced by a mutation in an intron sequence, in this case due to a transposition of a movable element.

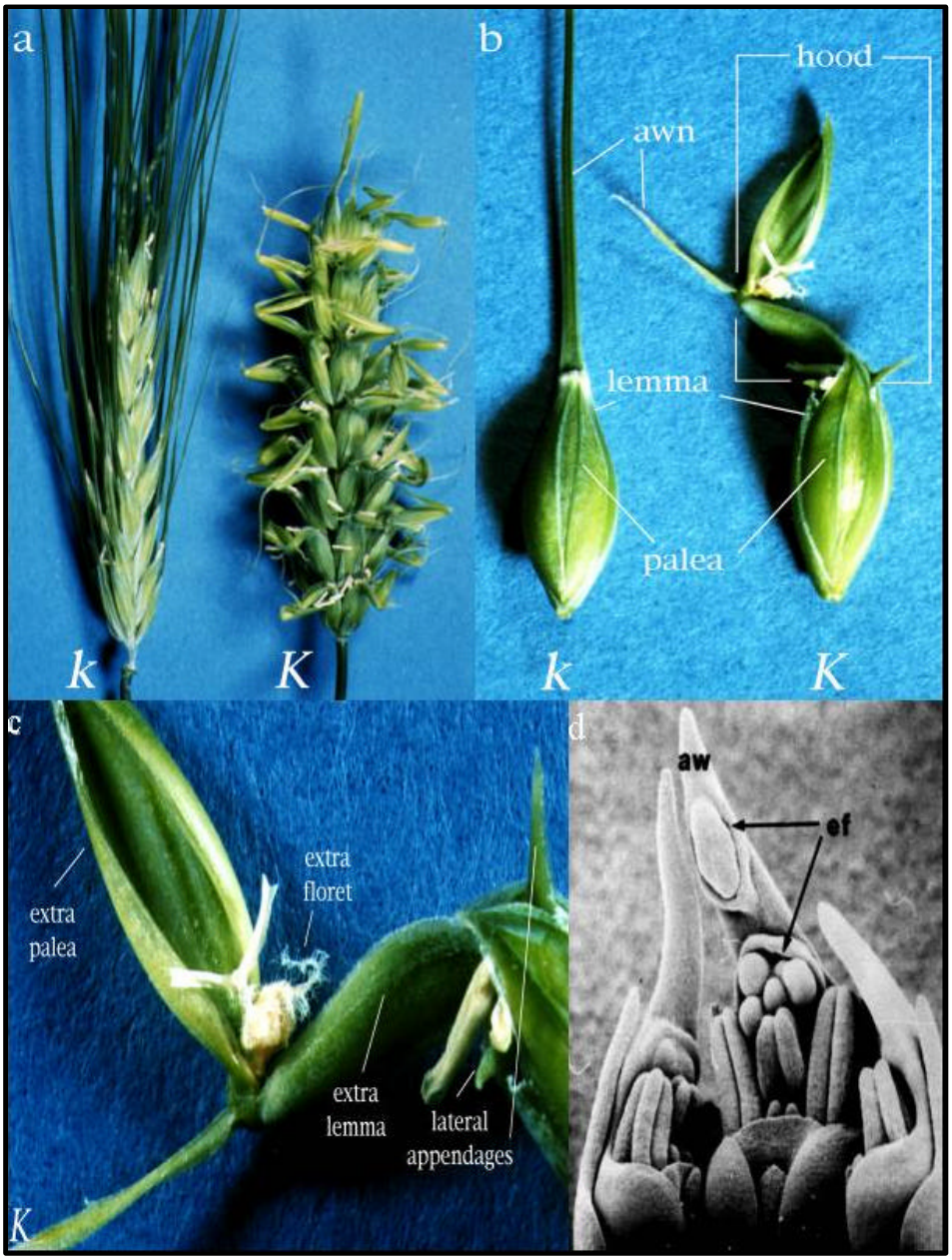


Fig 1.2b *Hodded* (K) mutation a) at spike and b) at spikelet level. The ectopic structure is composed of c) an extra lemma, palea and floret, d) it generates from a region of reacquired meristematic activity on the lemma/awn (aw) transition zone.

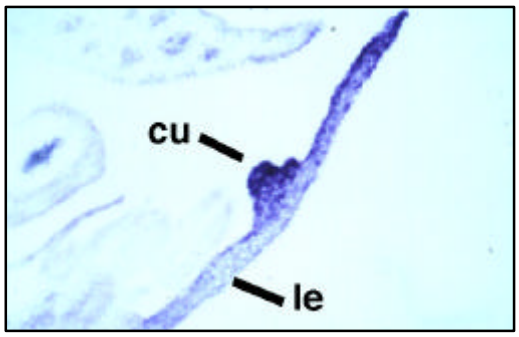


Fig. 1.3 Expression pattern of *Bkn3* mRNA detected by *in situ* hybridisation. The *Bkn3* transcript localise in the meristematic cushion (cu) on the lemma (le), from which the ectopic extra flower develops (Muller *et al.*, 1995).

The intron modified in *Kn1* is the third which corresponds to the forth intron of *Bkn3*, pointing out similarity not only on the structure of these two genes but also in the regulation that lies at the base of both mutant phenotypes. Moreover, the constitutive expression of *Kn1* in barley reproduces the *Hooded* mutant (Williams-Carrier *et al.*, 1997).

Other KNOX genes are associated with dominant mutants affected in leaf development in maize: *rough sheath 1* (*rs 1*) that shows a ligule displacement into the leaf blade (Becraft and Freeling, 1994); *liguless 3* (*lg 3*) that is missing the ligule (Fowler *et al.*, 1996) and *gnarley 1* (*gn 1*) where regions with blade identity appear in the leaf sheath (Foster *et al.*, 1999). Also in tomato, *tkn2* has been associated with the two mutants *Curl* (*Cu*) and *mouse-ear* (*Me*) (Parnis *et al.*, 1997). Both mutants affect the typical compound character of the tomato leaf increasing the level of ramification, as described for the overexpression of *kn1* in tomato.

Heterologous over expression of KNOX genes cloned from different species has been widely used. The class 1 subfamily of genes shows a drastic morphogenetic effect. For example, the over expression of *Bkn3* in transgenic tobacco plants, induces severely dwarfed plants with short leaf blades, and an early divergence of lateral veins from the midrib. In addition, the flowers petals are fused just at the base and two kinds of epiphilluous structures were found: shoots on the basal leaf laminae and epiphyllous flowers on the upper stem leaves (Muller *et al.*, 1995) (Fig. 1.4).

Both dominant mutants and over expression studies support a role of class 1 KNOX genes in meristem maintenance. The direct evidence arrived when the first loss of function mutants were identified, *shoot-meristemless* (*stm)* from *Arabodopsis* and later *kn1* from maize (Long *et al,*. 1996; Kerstetter *et al.*, 1997).

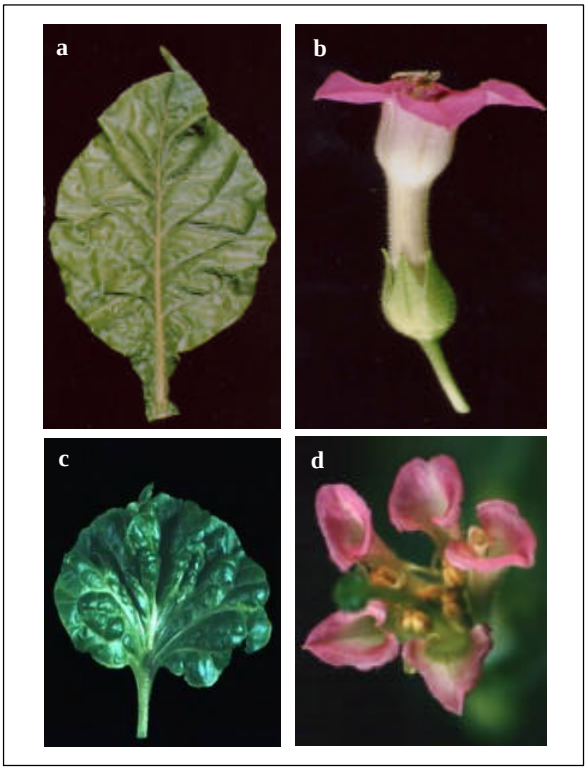


Fig. 1.4 Overexpression of *Bkn3*, a typical example of the class I Knox genes, in transgenic tobacco plants. c) Leaves are short and rounded with early divergence of lateral vein from the midrib d) in flowers the petals are fused only at the base. In addition, the overexpressing lines present severe dwarfism and different epiphilluos structures. (Muller *et al.*, 1995). In (a) and (b) a WT leaf and flower respectively are shown.

STM is required for shoot apical meristem formation during embryogenesis and is required postembryonically for continuous shoot and floral meristem function; preventing differentiation of cells into more determinate organ primordia. Other mutations like in the WUSCHEL (WUS) or ZWILLE (ZLL) genes result in defective organisation and premature termination of the shoot meristem. Genetic interactions the three mutants cited show that STM acts upstream of WUS and ZLL (Endrizzi *et al.*, 1996). *Kn 1* is expressed in the vegetative SAM, as well as in the axillary and in terminal and lateral inflorescence meristems. The loss of function mutation of this gene, causes

inflorescence and floral defects, reduced branching and ears often absent. When they are present are small with few spikelets; less frequently, extra leaves form in the axils of vegetative leaves.

In a screening for recessive mutations that confer phenotypes similar to the KNOX dominant mutation, *rough sheat 2 (rs2)* was identified. Its phenotype includes dwarfism, leaf twisting, disorganised differentiation of the blade-sheath boundary, aberrant vascular patterning and generation of semi-bladeless leaves. The ectopic expression of *Kn1*, *Lg 3* and *Rs 1* in the developing primordia of the *Rs 2* mutant revealed the inhibitory action of this gene on the expression of KNOX genes, confining them at least in maize, to the central zone of the shoot apex. Sequence analysis revealed that the RS2 gene encodes a protein that shares sequence similarity (62.9% identity, 76.7% similarity) with the Myb-like protein encoded by the *PHANTASTICA* (*PHAN*) gene of *Antirrhinum*. The RS2 and PHAN gene products are more similar to one another (62.9% identity) than to any other *myb* gene product. In addition, RS2 is more similar to PHAN than to a second PHAN-like gene from *Antirrhinum*. The sequence conservation, similarity in expression patterns and ectopic accumulation of KNOX gene products in mutant leaves suggest that *RS2* and *PHAN* are functional orthologs. At least one KNOX-like gene (*AmSTM1*) is also ectopically expressed in *phan* mutant leaves. This concurs with the finding that the expression domains of *PHAN* and *AmSTM1* are mutually exclusive during wild-type *Antirrhinum* development (Schneeberger *et al.*, 1998; Tsiantis *et al.*, 1999).

All results cited support the idea that genes like *RS2* are involved in specifying or maintaining a developmental state through the exclusion of genes which promote the elaboration of a different developmental state. A conceptually similar role has been proposed for the *CURLY LEAF* gene in *Arabidopsis* which acts to exclude *AGAMOUS* gene expression from leaves (Goodrich *et al.*, 1997). In this case the assumption are based on the phenotypical and genetic analysis of these mutants and the expression patterns of the respective genes, but almost nothing is known about their molecular mechanism of action.

On the molecular side, more it is known about the protein interactions of the KNOX genes.

They interact with members of their own family from both classes and at least two cases are known of interaction with a homeobox gene of the Bell family, suggesting a role of action of homeobox heterodimers in developmental processes (Mueller *et al.*, 2001; Bellaoui *et al.*, 2001). In *Arabidopsis*, the BELL1 gene, that is crucial to the production of lateral primordia within developing ovules, has been used as the bait in a two-hybrid screen and has been shown to interact with the KNOX proteins from: KNAT1, KNAT2, KNAT5 and STM.

Using the opposite strategy, in barley, BKN3 was used as bait and showed to interact with two BELL-like proteins JUBEL1 and JUBEL2.

In barley, contradictory to the common model that the down-transcriptional regulation of class 1 *knox* genes is the prerequisite for organ differentiation, transcripts of this family of genes were detected in developed organs. A similar pattern of KNOX gene expression was detected in tomato, in incipient and immature leaves as well as in meristematic tissues (Avivi *et al.*, 2000). This complicates the understanding of this process even more, and highlights the role of post-transcriptional regulation in KNOX gene action.

The availability of a set of mutants affecting leaf development is the feature that makes barley an attractive model system to study developmental problems. For instance knowledge of genes responsible for the *calcaroides* phenotype should help in understanding the regulation of the *Hooded* phenotype (Pozzi *et al.*, 2000). *Calcaroides's* name is derived from the similarity of its mutated lemma to a heel ('*calcar*' in latin). In those phenotypes at the tip of the lemma in a position corresponding to the transition between WT lemma and awn a well organized ectopic structure, the sac appears. In contrast to the hooded phenotype, the sac does not develop into an epiphyllous flower and the sac bear a distal awn. Extensive genetic studies have been conducted on this mutant, revealing that five different loci are responsible: *cal a, b, C15, d* and *23*. For loci *cal a, b and d* more than one allele is available. All five loci have been mapped on the (Proctor X Nudinka) barley map (Castiglioni *et al.*, 1998). Mapping homeotic genes that are available, should allow the identification of candidate genes for a particular mutant.

Also associating *Hooded* suppressors with known genes would be very informative. *Hooded* is homozygous viable and this allowed the generation of second site mutants at loci coding for putative factors participating in the regulation of *Bkn3*. Five loci responsible for the suppression of the *Hooded* syndrome have been identified and mapped (see appendix). The suppressors condition the substitution of the ectopic flower, present on the *K* lemma, with a WT-like awn reduced by 50% in its length.

For a deeper understanding of the real role and action of homeobox genes in plant development the identify target genes, of modulator genes and of signal cascades responsible for pattern formation at the molecular level is now a permanent necessity.

1.5 Transcriptional regulation of homeobox and homeotic genes

1.5.1 *D. melanogaster* and other animals

Regulation of the action and the expression of homeobox genes is crucial to establish the correct development of a complex organism and has been particular well studied in *D. melanogaster*.

Domains of homeotic gene expression are regulated through both transcriptional activation and repression (Paro and Harte, 1996). The topological domain of activation or repression must be precisely established and maintained during development, since misexpression of homeotic genes leads to severe homeotic transformations and lethality. Establishment of the activation and repression domains of homeobox genes takes place early in embryonic development and depends on the function of the segmentation genes. Maintenance throughout development involves transition to a different set of proteins (Kehle *et al.*, 1998). The *trithorax* group of genes (*trx*G) is required for the maintenance of the transcriptional state of the homeotic genes (Kennison, 1995). In contrast, the genes for the maintenance of the repression are included in the *Polycomb* group (PcG). Some pairs of PcG proteins interact *in vitro* (Kyba and Brock, 1998) and *in vivo* and are associated with large protein complexes (Tie *et al.* 2001). Two large PcG complexes have been described to be present in *D. melanogaster* embryonic extracts. One, called PRC1, has a molecular mass greater than 2 MDa and contains at least four PcG proteins: PC, PH, PSC, and SCM (Shao *et al.*, 1999). The other complex of 600 kDa includes at least two PcG proteins, ESC and E(Z) (Tie *et al.* 2001). The cloning of the *Pho* gene has shown that it encodes a zinc finger protein related to the mammalian transcription factor YY1 and that it recognizes specific DNA sequences. This makes PHO the only PcG protein so far characterized that has specific DNA binding properties. The PHO consensusus DNA binding site has been defined, ATGGC (Busturia *et al.*, 2001). This finding has led to the proposal that PHO protein binds to DNA and recruits the PcG silencing complexes to specific DNA sequences (Brown *et al.*, 1998).

PcG proteins act on *cis*-regulatory elements of the homeotic genes to maintain the silenced state. These elements are known as Polycomb Response elements, or PREs (Chan *et al.*, 1994). Several PREs from the bithorax complex as well as from other genes regulated by PcG proteins have been characterized. They are segments of several hundred base pairs that, when fused to a reporter, are able to silence target promoters in cis (Pirrotta 1999). They can mediate pairing-sensitive silencing. Specific sequences recognized by PHO have been shown to play a role in the silencing activity of PREs (Shimel *et al.*, 2000). The same studies showed, however, that the PHO recognition sequences by themselves are not sufficient to serve as a PRE, suggesting the existence of additional protein(s) that bind PREs.

Extensive work has been conducted on the so called MCP element (Busturia 2001). This DNA fragment is a 822 bp sequence from the *iab-5* regulatory region of the *Abd-B* gene and corresponds to the genomic region deleted in *Mcp* mutant alleles, which cause *Abd-B* to be expressed outside its normal domain (Karch *et al.*, 1994). The MPC element is central to homeobox genes regulation. Several of its properties have been defined: (1) it silences a reporter gene when upstream or

downstream of the associated enhancer (Busturia *et al.*, 1997); (2) it functions in both orientations (Bustura *et al.*, 1997), (3) it behaves as a PRE since PcG functions are necessary for its silencing activity (Bustura *et al.*, 1997); and (4) it participates to long distance silencing (Muller *et al.*, 1999). The element shortened to 138bp contains four PHO binding sites and two GAGA factor (GAF) binding sites. Both PHO and GAF are required for the silencing activity of MCP element through the action of *pleiohomeotic* and the GAF encoding gene *Trithorax-like* (Bustura 2001). The authors propose two models of action: 1) GAF and PHO bind to the MCP element in a sequential order, with GAF binding absolutely required for binding or activity of PHO. GAF, for example, may open up the chromatin at MCP, allowing binding of PHO. Upon binding, PHO may recruit PcG silencing complexes. GAF has been shown to induce Dnase I hypersensitive sites, or nucleosome-free regions (Lu *et al.*, 1993) and this may create a prerequisite condition for PHO to bind to its recognition sites; PHO may act alternatively as a facilitator of GAF binding by creating preconditions; 2) GAF and PHO bind to MCP independently of each other. Each protein may induce a unique chromatin modification that, together, can have a positive synergistic effect on the recruitment of PcG silencing complexes.

Although the importance of GAGA binding sites in regulating various types of genes, homeobox genes in particular, has been widely demonstrated, even in members of very different species like *D. melanogaster* (Granok *et al.*, 1995; Wilkins and Lis, 1997; Busturia *et al.*, 2001; Hodgson *et al.*, 2001), *Xenopus* (Li *et al.*, 1998), mouse (Bevilacqua *et al.*, 2000) and rat (Rabadan-Diehl, *et al.*, 2000) only in *D. melanogaster* have the factors been cloned and molecularly characterised. Extensive studies have been conducted on TRL/GAGA factor (GAF) encoded by the *trithorax-like* gene, originally it has been identified by the ability to stimulate transcription from the *engrailed* (*en*) and *ultrabithorax (Ubx)* promoters *in vitro*, GAF binds to GA-rich sequences, $(GA)_n$.

GAF is associated with euchromatic and heterochromatic chromosome regions (Raff, *et al*, 1994). Mutated alleles of the *Trl*/GAF locus enhance position-effect variegation (PEV; Farkas, *et al.*, 1994), while GAF colocalizes with Polycomb proteins on homeotic genes (Strutt *et al.*, 1997), and is involved in gene silencing (Pirrotta, 1997). This supports the role GAGA in correct body development. GAF does not appear to affect the transcriptional machinery directly, but acts to relive the repressive effects of histones (Croston *et al.*, 1991). Mutations in GA-rich sequences in the promoter of *hsp70* have been shown *in vivo* to specifically affect the formation of promoter RNA polymerase II (Pol II) complexes, the accessibility of the heat shock factor (HSF) to target sequences, and transcriptional activation (Shopland *et al*, 1995). Recombinant GAF can disrupt nucleosomes with the help of a complex termed NURF which facilitates the “remodelling” of

chromatin in an ATP-dependent fashion (Tsukiyama *et al.*, 1994). Thus GAF stimulate gene expression by opening chromatin and maintaining the affected promoter in an open conformation.

GAF has been studied in the context of transcriptional activation, and its effects on promoter architecture have been examined both *in vitro* and *in vivo*, but the contribution of the individual domains remains poorly understood. GAF has a POZ/BTB domain (Zollman, *et al.*, 1994), a zinc finger DNA-binding domain (Nardelli, *et al.*, 1991) and a glutamine-rich region (Q domain; Wilkins and Lis, 1999). little evidence exists for the functions of the POZ/BTB and of the Q-domain.

The POZ/BTB domain defines a growing family in *D. melanogaster* (Zollman *et al.*, 1994), and has been shown to function as a protein interaction domain (Chen *et al.*, 1995). However, the POZ/BTB domain does not appear to mediate homodimer formation *in vitro*, yet it serves to inhibit the DNA binding ability of GAF, presumably through intramolecular interactions (Bardwell and Treisman, 1994). While GAF may not dimerize with itself *in vitro via* a POZ-POZ interaction, different isoforms of GAF can indeed form heteromultimers *in vivo* (Benyajati *et al.*, 1997). Whether GAF multimerization involves heteromeric interactions with the POZ domain is not clear. The POZ/BTB-containing factors that might potentially interact with GAF remain to be determined. It has been proposed, neverthless that GAF might assist communication between distal enhancers and promoters (Katsani *et al.*, 1999). In their paper EM images of GAGA-*ubx* promoter complexes are shown, where GAGA oligomers bind two DNA molecules simultaneously. This suggests that POZ-POZ interactions may lead to the association of enhancers- or PREs-bound GAGA with promoter-bound GAGA. The GAGA complex might binds target promoters by contacting multiple binding sites that are spread out over a region of few hundred base pairs and induce bending of the promoter DNA (Fig. 1.5). The Q domain of GAF also has the potential to mediate protein-protein interactions. The glutamine rich domains of SP1 is critic for transcriptional activation and multimerization (Pascal and Tjian, 1991) and necessary interactions with TBP (Emili *et al.*, 1994) and heteromultimerization with Oct1 (Strom *et al.*, 1996). The products of the Groucho-related gene (*grg*) family can also dimerize through their amino-terminal Q domains (Pinto and Lobe, 1996). Molecular modelling suggests polar zippers of antiparallel β strands linked by hydrogen bonds between amide groups are formed (Stott, et al., 1995). This allows assuming that the Q domain of GAF, and/or flanking regions, may have some function in protein-protein interaction. Other functions have been assigned to this domain. Permanganate reactivity of DNA sequences and Dnase I protection in the presence of different deletions and fusions of GAF demonstrated that the Q domain is responsible of promoter distortion and extended protection. The Q domain is able to bind single stranded DNA (Wilkins and Lis, 1999).

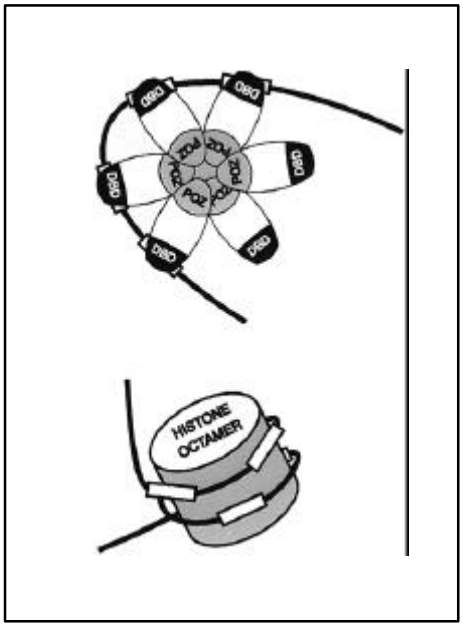


Fig. 1.5 Possible mechanism of action of the transcriptional regulation mediated by (GAGA) elements. The DNA binding domains (DBD) of the GAGA binding factors (GAF) from *D. melanogaster* contact the (GAGA) elements, while the POZ domains mediate oligomerisation. The reduced accessibility of the DNA in between the GAGA sites and the bending of the promoter DNA is the result of the GAGA complex binding.
As a comparison, the promoter DNA wrapped around a histone octamer is shown (Katsani *et al.*, 1999).

1.5.1 Plants

Plant homeotic genes are regulated by *polycomb-trithorax* group of genes as well (Goodrich *et al.*, 1997). Experiments based on the analysis of cell lineages and on the effects of ablating cells, suggest that in many cases the fate of plant cells is not predetermined, but rather is position-dependent and maintained through interactions between neighbouring cells (Weigel and Doerner 1996, Spena and Salamini 1995). Around the time of floral induction, shoot meristems may become stably committed to form inflorescence or flowers. Based on the genetic and morphological analysis of floral homeotic mutants, a model has been proposed to account for the specification of organ identity in the different whorls based on combinatorial action of homeotic genes (Coen and Meyerowitz, 1991). The isolation and characterization of the *CURLY LEAF (CLF)* gene in *Arabidopsis* helped in extending the analogies in *D. melanogaster* gene regulation with plants. *clf* mutants produce leaf blades curled upwards from the two sides to the centre. The mutant plant has alterations in the identity of outer floral organs of the sepals and petals. These phenotypes

resemble those observed in plants ectopically expressing the floral homeotic gene *AGAMOUS (AG)* (Mizukami and Ma, 1992). This suggests that *clf* mutants express *AG* ectopically. *AG* from *Arabidopsis* is required to specify stamen and carpel identity in whorls 3 and 4 respectively. Molecular isolation of the *AG* gene indicates that it encodes a protein belonging to the MADS box family of transcription factors with striking similarities to the sequences of transcription factors from both humans (*SRF*) and yeast (*MCM1*), as well as the deduced amino acid sequence of an homeotic gene (DEF A) from the flowering plant *Anthirrinum majus.* Indeed, in a *clf* mutant, AG is precociously expressed in most of the vegetative organs and although it is expressed normally in early flowers, it is expressed ectopically in late petals and the inflorescence stem. Therefore CLF is required to repress AG in several organs during vegetative and reproductive development. In particular, CLF is not involved in the establishment of correct *AG* expression pattern, but rather it is required to keep *AG* turned off, where it should not be expressed. The ORF encodes a potential protein, CLF, of 902 amino acids. Comparison of the amino acids sequence of CLF to protein sequence databases using the BLAST program reveals extensive homology between CLF and the product of the *D. melanogaster* gene *Enhancer of zeste* (E(*z*)), a member of the Polycomb group of genes. There are three regions that are conserved between the two proteins. First the carboxy terminus of CLF contains a 115 amino acid region, the SET domain, previously recognised as a conserved region in the products of three *D. melanogaster* genes, E(Z), TRITHORAX (TRX) and SU(VAR)3-9, all involved in transcriptional regulation (Jones and Gelbart, 1993). The SET domains of CLF and E(Z) are more similar to each other (65% identity) than either is to the SET domains of TRX or SU(VAR)3-9.

Recently two genes from *Arabidopsis* homologous to the animal *tritorax* genes: *ATX-1* and *ATX-2* have been characterized (Alvarez-Venegas and Zoya Avramova, 2001). These two genes are highly similar but display different tissue and developmental expression patterns. While *ATX-1* is ubiquitously expressed, *ATX-2* display a more specialized pattern with the highest level apparently in roots, both of them contain the SET domain. The finding of plant gene members of the *Trithorax* family does not necessarily imply that these genes function necessarily within a complex of synergistically acting factors similar to the Trx-G complex of *D. melanogaster*, just as the presence of a SET domain belonging to the *Trithorax* superfamily does not imply that such genes would have activating function. Thus, despite the fact that the human homolog of the *D. melanogaster trithorax* contains a SET domain belonging to the *Trithorax* family, the human gene might be involved in a gene-repression function, in contrast to the fly. It also has to be noted, that activating functions have not yet been reported for most of the currently known animal homologs of *thrithorax.* As already The SET domain is present in both activators and repressors and is regarded as a dual function

motif. Its role appears highly specific and dependent upon the remaining components of the protein. The SET domain of the *Su(var)* family genes (but of neither *E(z)* nor *trithorax*) encodes a histone H3-specific methyltransferase. All three members have a SET domain, but a specific function requires the presence and collaboration of other elements within a particular molecule.

2 Materials and Methods

2.1 Materials

2.1.1 Plant materials

Hordeum vulgare. Cultivars: Atlas, *Hooded Atlas*, Proctor, Nudinka, Golden promise.
Nicotiana tabacum. Cultivar: Petit Havana line (SR1).

2.1.2 Bacterial strains, yeast strains

E.coli strains

BL21(DE): *hsdS gal* (cI*ts*857*ind*1 *S*am7 *nin*5 *lac* UV5-T7 gene*1*.
DH10B: F$^-$,mcrAΔ(mrr-hsdRMS-mcrBC)Φ80dlacZΔM15,ΔlacX74,deoR, recA1, endA1, araD139,Δ(ara,leu)7607, galU, galK, λ$^-$rpsl,nupG (GIBCO-BRL)
Xl1-Blue MRF´: Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 *supE44 thi-1 recA1 gyrA96 relA1 lac*[F´*proAB lac*^{q}*ZDM15* Tn*10* (Tet^r)]
POP 13:

S. cerevisiae strains

YRG-2: Mata *ura3-52 his3-200 ade2-101 lys2-801 trp1-901 leu2-3 112 gal4-542 gal80-538* LYS2::UAS$_{GAL1}$-TATA $_{GAL1}$-HIS3 URA3::UAS$_{GAL\ 4\ 17mers(x3)}$-TATA$_{CYC1}$-*lacZ*.

A. tumafaciens strains

LBA 4404: Sm^r, Rif^r.
GV 3101 PMP 90 RK: Kan^r, Rif^r

2.1.3 Chemicals

Laboratory reagents were obtained from Boehringer-Manheim/Roche, Gibco, Merck GmbH and Sigma-Aldrich unless otherwise state. All were reagent grade, if not otherwise stated. Filter paper was obtained from Whatman, Hybond N-membranes and radioisotopes were from Amersham Buchler, XR films were from Eastman Kodak Co. Tissue culture chemicals were from DUCHEFA, Merck and Boehringer. Plant solidifying agents, Gelrite, were purchased from Carl ROTH GmbH Co.

2.1.4 Cloning vectors

Name	organism	description	origin
pBlue KS	*E. coli*	Basic cloning ve tor.	Stratagene
K373	*E. coli*	Small reporter GUS vector with a l 35S minimal. Used in protoplast transient ex pression .	Thompson R.
pBT10GUS	*E. coli*	Small reporter GUS vector with a l 35S minimal. Used in protoplast transient e pression.	Weisshaar B.
pBT4	*E. coli*	Effector vector, allows to produce fusion with the VP16 AD. The expression is driven by a 5S full promoter. Used in protoplast transient xpression.	Weisshaar B.
ms129-pbt8-ubi-lucm3	*E. coli*	It carries the LUC reporter gene under the ubiquitin promoter, used as internal control n protoplast transient expression.	Weisshaar B.
pRT101	*E. coli*	PUC derivative, allows to clone i side a 35S complete cassette.	Topfer *et al*., 1987
pENTR4	*E. coli*	Cloning vector, compatible with t e Invitrogen getway technology	Invitrogen
pAD Gal4	*E. coli* *S. Cerevisiae*	One- and two-hybrid vector, allows o produce fusions with the Gal4 A.D	Stratagene
pBD Gal4	*E. coli* *S. Cerevisiae*	Two-hybrid vector, allows to produ e fusions with the Gal4 B.D.	Stratagene
pGEX 5X	*E. coli*	Protein overexpression and p rification vector. GST tag at the N-te minus.	Pharmacia
pBin19	*E. coli* *A. tumefaciens* *N. tabacum*	Binary overexpression vector wit Kanamycin as plant selectable mar er.	Frish *et al.*, 1995
PGPTV-KAN	*E. coli* *A. tumefaciens* *N. tabacum*	Binary reporter GUS vector with Kanamycin as plant selectable mar er.	Becker *et al.*, 1992
pPCV91	*E. coli* *A. tumefaciens* *N. tabacum*	Binary overexpression vector wit Hygromycin as plant selectable mar er.	Koncz C

2.1.5 Media

Antibiotics

name	stock	storage	Final concentration	organism
Ampicillin	water	-20°C	100 µg/mL	*E. coli*
Carbenicillin	water	-20°C	100 µg/mL	*E. coli/A. tumefaciens*
Gentamycin	water	-20°C	25 µg/mL	*A. tumefaciens*
Hygromycin	Phophate buffer saline	+ 4°C	20 µg/mL	*N. tabacum*
Kanamycin	water	-20°C	25 µg/mL (1) 100 µg/mL (2)	*A.tumefaciens* (1) *E. coli* (1) *N. tabacum* (2)
Rifampicin	methanol	fresh	100 µg/mL	*A. tumefaciens*
Streptomycin	water	-20°C	300. µg/mL	*A. tumefaciens*
Tetracycline	ethanol	-20°C	100 µg/mL	*E. coli*

Luria Bertani (LB) medium: 1% tryptone, 0.5% yeast extract, 1% NaCl, and 2% of agar in the case of the solid medium the medium is autoclaved and cooled to 50°C before adding the antibiotic.

YEP medium: 0.5% saccharose, 1% yeast extract, 1% peptone and 0.5% NaCl.

YEB medium: 0.5% beef extract, 0.5% saccharose, 0.1% yeast extract, 0.5% Peptone.

NZY medium: 0.5% bacto yeast extract, 0.5% NaCl, 0.2% $MgSO_4.7H_2O$, 1% caseamino acids, pH 7.5.

Super broth: 3.2% tryptone, 2% yeast extract, 0.5% NaCl, 5 ml/L 1N NaOH.

YPD medium: 2% peptone/tryptone, 1% yeast extract, (for solid media 2% agar) adjust the pH to 5.8, autoclave. On use add 40% dextrose.

MS medium: 4.3 g/L Murashige/Shoog basal salts (micro and macro elements included), Vitamins solution 1 ml/L, Myo-inositol 100 mg/L, 30 g/L glucose, 8 g/L agar, adjust the pH to 5.8-6.0, autoclave.

2.1.6 Buffers and solutions

Alkaline phosphate buffer: 100mM Tris pH 9.5, 100 mM NaCl and 50 mM $MgCl_2$

30% Acrylamide: 29.2 %(w/v)acrylamide, 0.8%(w/v)N-N´-methylene bisacrylamide in de-ionazed water.

5X MOPS: 0.2M MOPS pH 7, 50mM NaOAC, 50mM EDTA. Everything prepared in DEPC water. Autoclave.

Denhardts solution (100X): 2%(w/v) BSA, 2%(w/v) Ficoll, 2%(w/v) PVPP360.

DEPCwater 0.1% (w/v): DEPC in deionazed water shaking over night at 25°C. Aautoclave.

DNA extraction buffer: 100mM TrisHCl pH 8.5, 100mM NaCl, 50mM EDTA pH 8, 2%SDS and 0.1mg/ml fresh added proteinase K.

1X TAE buffer: 40mM Tris, 1mM EDTA, 1.162 ml/L acetic acid.

2.1.7 Protein buffers

blotting buffer: 25mM Tris-HCl pH 8.3, 192 mM glycine, 20% methanol.

extraction buffers:

For denaturing conditions isolation from bacteria:

Buffer B: 100mM NaH_2PO_4, 10 mM Tris-Cl, 8 M Urea. Adjust pH to 8.0

Buffer C (wash buffer): 100mM NaH_2PO_4, 10 mM Tris-Cl, 8 M Urea. Adjust pH to 6.3

Buffer D (elution buffer): 100mM NaH_2PO_4, 10 mM Tris-Cl, 8 M Urea, 100mM EDTA. Adjust pH to 6.3.

For native conditions from bacteria:

Buffer 1 (lysis buffer): 50mM NaH_2PO_4, 300mM NaCl, 10mM imidazole. Adjust pH to 8.0 (lysozime 200 µg/ml, DNAse 100 µg/ml, RNAse 100 µg/ml, 1mM PMSF).

Buffer 2 (wash buffer): 50mM NaH_2PO_4, 300mM NaCl, 20mM imidazole. Adjust pH to 8.0 (1mM PMSF).

Buffer 3 (elution buffer): 50mM NaH_2PO_4, 300mM NaCl, 250mM imidazole. Adjust pH to 8.0.

LaemmlI buffer(4X): 0.25M Tris-HCl pH 8.2, 0.4%(w/s) SDS, 767 mM glycine.

Loading buffer for protein gels (2X): 160mM Tris pH 6.8, 5% SDS, 20% glycerol, 0.004% Bromophenol Blue, 20% ß-mercaptoethanol (The buffer is prepared without adding the ß-mercaptoethanol, which is added just before the sample).

SDS-precipitation buffer: 3.59 ml of K_2HPO_4, 1.42 ml of KH_2PO_4 and water to 50 ml.

10X STE buffer: 100mM TRIS pH 8.0, 1.5 M NaCl, 10mM EDTA.

PBS buffer: 8 g/L NaCl, 0.2 g/L KCl, 1.44 g/L Na_2HPO_4, 0.24 g/L KH_2PO_4 use HCl to adjust the pH to 7.4.

GST elution buffer: 15mM reduced glutathione in 50mM TRIS-HCl, pH 8.0, 0.2mM PMSF, 1mM benzamidine, 5mM caproic acid.

2.1.18 RNA buffers

Buffer I: 0.1M NaCl, 0.05M TrisHCl pH 9, 0.01M EDTA, 2% SDS and 0.2mg/ml fresh added proteinase K.

Buffer II: 0.4M NaCl, 0.01 Tris-HCl pH 7.5, 0.2% SDS.

Buffer III: 0.1M NaCl, 0.02M TrisHCl pH 7.5, 0.01% SDS.

Buffer IV: 0.01M TrisHCl.

For extraction of total RNA: 0.2M Tris-HCl pH 7.5, 0.1M LiCl, 5mM EDTA, 1% SDS. Everything is prepared in DEPC water.

20X SSC: 3M NaCl, 300mM sodium citrate

20X SSPE: 200mM disodium hydrogen phosphate, 20mM sodium dihydrogen phosphate, 3.6M NaCl, 20mM EDTA pH 8.

PSE buffer: 0.5 M sodium phosphate pH 7.2, 7% SDS, 1% BSA, 1 mM EDTA (pH 8.0) and salmon sperm DNA 50µg/ml.

STE buffer (1X): 20mM Tris-HCl pH 7.5, 10mM EDTA pH 8, 100mM NaCl.

TAE buffer: 400mM Tris-HCl, 200mM NaOAC, 18mM EDTA pH 7.8 with glacial acetic acid.

TNT buffer: 100mM Tris-HCl, pH 7.5, 150 mM NaCl and 0.3% (v/v)Triton X-100.

2.1.9 Yeast buffers

LacZ plates: 100mM phosphate buffer pH 7, 10mM KCl, 2% Agar (autoclave), 1mM MgSO4 sterile, 1mM DTT sterile, 0.3-1 mg/ml X-Gal.

Z-buffer: $Na_2HPO_4.2H_2O$ 11.1 g/L, $NaH_2PO_4.H_2O$ 5.5 g/L, KCl 0.75 g/L, $MgSO_4.7H_2O$ 0.25 g/L, Adjust to pH 7.

50%PEG: PEG 4000 500 g/L, (Sigma #P-3640), Autoclave.

10X TE: 1M Tris-HCl pH 7.5, 500mM EDTA pH 7.5, Autoclave.

10X Lithium Acetate : Lithium Acetate 102 g/L (Sigma#L-6883), Adjust pH with Acetic Acid. Autoclave.

PEG/LiAc: 50%PEG 800 ml/L, 10X TE 200 ml/L,10X LiAc 200 ml/L, make fresh just before use

2.1.10 Protoplasts buffers

K3 0.4M: Murashige and Skoog medium (micro and macro elements included),Vitamins solution 1 ml/L, Myo-inositol 100 mg/L, Xylose 250 mg/L, Sucrose (0.4M) 136.92 g/L, NAA 1 mg/L, Kinetin 0.2 mg/L. The mOs should show values between 550-590. Use KOH to achieve a pH 5.7. Sterile filtration.

K3 0,4M–cellulase–macerozyme: for 100 ml of K3 0.4M, Cellulase (SIGMA) 0.1%, Macerozyme (SERVA) 0.4%. Dissolve by stirring for 4 hrs, at 28°C in darkness and filtrate sterile.

Vitamin Solution: Myo-inositol 0.5 g/L, Nicotinic acid 0.5 g/L, Pyrodoxine HCl 0.5 g/L, Thiamine HCl 0.5 g/L**.** Sterile filtration.

W5: NaCl 0.154M, $CaCl_2.2H_2O$ 0.125M, KCl 0.005M, Glucose 0.005M.

Use KOH to achieve a pH 5.7, sterile filtration.

MaMg solution: MES 0.1%, Mannitol 0.5M, $MgCl_2.6H_2O$ 0.015M.Use KOH to achieve pH 5.7, autoclave.

PEG-solution: PEG 400 (Merck) 40%, $MgCl_2.6H_2O$ 0.015M, Mannitol 0.4M, Hepes 0.1M.

Use KOH to achieve pH 7.0, filter sterilized.

Extraction buffer: NaH_2PO_4 50mM, EDTA 10mM, *b*-mercaptetanol 10mM, 0.1%(v/v), Triton X-100, pH 7.

Substrate solution: 10mM 4-methylumbelliferyl-d-glucoronide(MUG), 50mM NaH_2PO_4, 10mM EDTA, 10mM *b*-mercaptetanol, 0.1%(v/v) Triton X-100, pH 7.

LUC extraction buffer: 100mM potassium phosphate pH 7.5, 1mM DTT.

LUC assay buffer: 20mM Tricine, 2.76mM $MgSO_4$, 1.07mM, $(MgCO_3)_4Mg(OH)_2.5H_2O$, 0.1mM EDTA, 33.3mM DTT, 270µM CoA, 470µM luciferin, 530µM ATP, pH 7.8.

2.1.11 Buffers for plant GUS analysis

Staining solution: 1-3mM Xgluc, 5mM *b*-mercaptoethanol, 0.1% TRITON X-100.

Fix solution: 50mM phosphate buffer pH 7.0, 1% Glutaraldehyde.

Destaining solution: 80% ethanol.

2.1.12 λ cDNA library screening buffers

NZCYM buffer: Rich media to grow bacteria which have to be infected with lambda phages.
NZ-ammine15 g/L, NaCl 5 g/L, Casammino acids 2 g/L, Yeast extract 5 g/L, 1M Tris pH 7.4 25 ml/L, 1 M MgSO4 5 ml/L, agar 15 g/L for solid media. Add 0.4% Maltose for infections.
SM buffer: to elute, dilute and resospend lambda.
NaCl 5.8 g/L, MgSO4 7H2O 2 g/L, 1M Tris pH 7.5 50 ml/L, 2% gelatine 5ml/L. Autoclave.
Mg-Top-Agar: MgSO4 2.5 g/L, Agarose 7 g/L, Autoclave.

2.2 Methods

2.2.1 Plasmid DNA extraction

The method used is a modification of the alkaline lysis method of Birnboim and Doly (1979) taken from Sambrook (1982). For sequencing quality DNA the QIAGEN miniprep extraction Kit was used.

2.2.2 Plant genomic DNA isolation

The method used was the one developed by Sharp *et al.* (1988).

Material ground in liquid nitrogen is resuspended in DNA extraction buffer at 4 ml/g and incubated at 37°C for 30 min. The mixture is then transferred to 40 ml centrifuge tubes, and sequentially phenol extracted and centrifuged. After RNAase treatment the DNA is precipitated and quantified.

2.2.3 Isolation of total RNA and messenger RNA from plant tissues

For the isolation of total RNA the liquid nitrogen ground material is resuspended in the extraction buffer and after phenolisation is precipitated with LiCl. All buffers are prepared in DEPC-treated water.
For the messenger RNA extraction (Bartels *et al.*, 1983), the material after grinding with liquid nitrogen is resuspended in Buffer I, and after an incubation at 37°C for 30 min, is extracted with phenol/chloroform. The messenger RNA is isolated by binding to oligo d-T cellulose, which is successively washed with the buffers II and III. The RNA is eluted from de cellulose in buffer IV.

2.2.4 Genomic southern analysis

DNA was digested with different restriction enzymes and electrophoretically separated on 0.8 % agarose gels in TAE buffer. The gels were treated with a 0.125 N HCl solution for depurination, and denatured for 30 min in 0.5M NaOH and 1.5M NaCl. The neutralisation is done in 0.5M Tris-HCl pH 7.5, 1.5M NaCl for 30 min. The DNA is then transferred and bound to a Hybond N membrane (Amersham) following the standard capillary transfer procedure (Sambrook *et al.*, 1989). The filters

were UV-cross-linked and prehybridised and hybridised in 5X SSC, 0.5% SDS, 5X Denhardt´s and 50 µg/ml denatured salmon sperm DNA at 65°C.
Following hybridisation, the filters were washed twice in 2X SSC, 0.1% SDS during 10 min and twice in 1X SSC 0.1% SDS during 10 min. They were packed in plastic bags, sealed and exposed to autoradiography during the required period of time.

2.2.5 Northern blot analysis

All the instruments for RNA handling were used just for that porpoise in order to avoid RNAases contamination. After the isolation of the RNA, this was electrophoresed in a denaturing agarose gel (1.2% w/v agarose 1X MOPs buffer, 2.2M formaldehyde) using 1X MOPS as running buffer. 50 µg total RNA were used and 2 or 4 µg poly-A RNA depending on the expression level of the transcript to detect. The RNA is denatured prior to electrophoresis by heating at 65°C during 10 min, in 50% (v/v) deionised formamide, 2.2M formaldehyde, 0.5X MOPS buffer and 1X RNA loading buffer. After electrophoresis, the samples were blotted onto Amershan nylon membranes. The fixation to the filter was done by means of a UV cross-linker (120,000 µJoules cm^{-2} for 30s). Pre-hybridisation and hybridisation of the filters was done in 50% deionised formamide, 5X SSPE, 0.5% (w/v) SDS, 5X Denhardt´s solution and 50µg/ml salmon sperm DNA. Wash steps were identical to southern analysis.

2.2.6 Preparation of radioactive labelled probes

Random oligonucleotide-primed synthesis

Probes were prepared from template isolated from plasmids and purified from agarose gels. The DNA was denatured by heating and labelled using the Kleenow polymerase (Roche).
X µl of DNA solution + Y µl of H_2O were denatured for 10 min at 95°C then chilled in ice. Then the rest of the components have been added and the reaction took place at 37°C for 30min, followed by 2/3 hours at room temperature. The product of the labelling reaction is usually purified using a quick spin column (Quiagen).

X µl (20-100ng)	Dna
5 µl	10X oligo buffer
5 µl	α^{32}P-dCTP
1 µl	Kleenow pol. (2U/µl)
Y µl	H_2O
50 µl	Total volume

2.2.7 PCR amplification

Genomic and plasmidic PCR (Munemitzu *et al.*, 1990 modified)

The amount of DNA template PCR used was 50 ng of pure DNA for both *Nicotiana tabacum* SRI, and *Hordeum vulgare*. The reaction was done in 50 µl of final volume with the next amounts of components:

20 pm for each primer
50 µM dNTPs (dATP, dGTP, dCTP, dTTP)
Taq buffer to 1X with a final concentration of $MgCl_2$ of 2 mM
1 U Taq polymerase (Gibco)

The reaction was carried out in a Biometra Thermal Reactor as follows:

2 m n 30 s at 85 °C	initial denaturatio
40 s at 94°C	denaturation
1 m n at 60°C	annealing
1 m n at 72°C	extention

The last three steps were cycled 35 times, followed by an extension of 5 min to ensure the completion of the reactions.

Plasmid PCR amplification

The amplification was done either from 10 ng plasmid or from single colonies. The reaction was done in 20µl of final volume and under the same conditions as before. The program contained a longer denaturing initial cycle (3 min at 94°C) and the number of cycles was decreased to 25.

2.2.8 Cloning methods

Preparation and transformation of *E.coli* competent cells

Preparation of E.coli electro competent cells

A single colony is used to inoculate a 10 ml LB culture that is grown over night. The next day a dilution 1/100 is done and the medium is kept growing till an OD_{600} of 0.6 is reached. The culture is pelleted during 10 min at 2000X g and gently resuspended in ice-cold sterile water. The process is repeated twice. After the washing, the cells are gently resuspended in a 1/100 volume of the initial culture in 10% sterile glycerol, again pelleted as before and resuspended in 1/1000 volume of glycerol. The cells can be already aliquoted and stored. The aliquots should be immediately frozen in liquid nitrogen.

Transformation of E.coli electro competent cells

20 to 50 ng of salt-free ligation DNA is mixed with the competent cells, and so pipetted to the electroporation cuvette. A pulse of 9 kv/cm; capacitance 25 µF; resistance 400-600 Ù. Immediately after the pulse, the cells are resuspended in 1 ml of LB and incubated at 37°C before plating.

Preparation of $CaCl_2$ competent cells

A single colony is used to inoculate a 10 ml LB culture that is grown over night. The next day 100 ml of medium are inoculated with 1 ml of the o.n culture and the medium is kept growing till an OD_{600} of 0.6 is reached. The medium is pelleted at 2000X g in a cold centrifuge and the pellet is resuspended in 10 ml of ice-cold 0.1M $CaCl_2$ and kept in ice during 20 min. The cells are pelleted again under the same conditions and resuspended in 1ml of 15% glycerol 0.1M $CaCl_2$. 100 µl aliquots are prepared and immediately frozen in liquid nitrogen.

Transformation of $CaCl_2$ competent cells

The cells are thawed on ice during 20min. Meanwhile the ligation is diluted twice with ice-cold 0.1M $CaCl_2$. Afterwards the ligation is carefully mixed with the cells and kept in ice during 1h. After this period the cells are heat-shocked at 42°C during 2 min and resuspended in 300-500 µl of LB before plating.

Ligation of double stranded DNA fragments

For this purpose the enzyme used was the ligase purchased by GIBCO-BRL®. One unit of enzyme was used per reaction. The vector was dephosphorilated and 10 ng were used for the reaction in the presence of different ratios of insert DNA.

2.2.9 *In vitro* transcription and translation

The open reading frame of the gene of interest was cloned in pBluescript that allows in vitro-transcription via the T3- or T7 promoter.

Plasmids are linearised 3' of the insert by restriction digest, preferably with an enzyme generating 5' protruding ends. About 1 µg of the linearised plasmid is used as template for reverse transcription from the promoter at the 5'-end of the insert.

The components of the reactions are: 5 µl of DNA template (200ng/µl), 5 µl of 10X reaction buffer, 7.5 µl of 3.3mM (CTP/ATP/UTP), 5 µl m^7 G(5')ppp(5')G, 1.5 µl Rnase inhibitor (40U/µl), 1.5 µl T3/T7 RNA-Polymerase (Roche) (5 U/µl) in a total volume of 50 µl.

The samples were incubated at 37°C for 90 min followed by a Dnase digestion, a phenol/chloroform extraction and an ethanol precipitation. The pellet has been taken up in 10 µl of H_2O and 1 µl has been used for quantification.

For in vitro translation both the wheat germ extract system (Promega) and the reticulocyte lysate system (Amersham) were used following the protocols of the suppliers.

In both cases the *in vitro* transcribed RNA was used at a dilution of 100 times, ^{35}S-Methionine was the label amino acid of choice and Rnase inhibitor (40U/µl) was included. After the two different extracts were added, the reactions were incubated at 25°C for 2h in the case of the weat germ and at 30°C for 1h when the reticulocytes were used. Subsequently the reaction were checked on SDS-PAGE and exposed for autoradiografy.

Detection of ^{35}S-labelled proteins in polyacrylamide gels

After the proteins were electrophoretically separated on an SDS minigel, the gel was incubated on a shaker for 1 h in 40% Methanol/10% Acetic acid, 2 times, for 30 min each time, in DMSO, 1 h in Rotifluorezint D and 1 h in H_2O. After the gel was vacuum dried, signals were detected using autoradiograph film.

2.2.10 λ NM1149 cDNA library screen

A cDNA library from young inflorescences of hooded barley *K-atlas* cloned in λ NM1149 was screened according to Sambrook (1989). The *E. coli* strain POP13 has been used for infection.
Several dilutions were used to titrate the library and a suitable volume to give rise to approximately 1,000,000 pfu was used for the infection. The screen proceeded for three rounds, after each round positive colonies, with differences in signal intensities, were selected. The probes were generated using previous cDNA clones, the template was purified from a gel and labeled using random priming.
The positive single λ plagues were prepared with the commercial λ mini-kit (Quiagen), the DNA was digested with *EcoRI* and the insert was checked by southern analysis.
After the confirmation of the hybridization signal, the inserts were cloned into pBluKS and subsequently sequenced.

2.2.11 Yeast two hybrid system

Yeast transformation

A single colony of YRG-2 was picked from solid YPD plates and 20 ml of liquid YPD were inoculated. The inocule was grown at 30°C for 1-2 h, then the cells were centrifuged and washed before with 2 ml of ice-cold water and then with 2 ml of 1X Li/TE. Finally the cells pellet was resospended in 200 µl of 1X Li/TE. Around 5 µg of plasmid DNA and 50 µg of carrier salmon sperm DNA were mixed with 50 µl of the yeast suspension and 500 µl of PEG/Li/TE and subsequently incubated for 30 min at 30°C, for 15 min at 42°C and then chilled down in ice. The cells were centrifuged for 1 min at maximum speed, the supernatant removed, 200 µl of 1X TE

added and everything plated on the right solid selective media. Then the plates were incubated for 2-3 days at 30°C.

β-galactosidase activity assay on filters

Positive transformed colonies were transferred on Hybond-N nylon membranes (Amersham). The membranes were placed on solid selective media and grown for 2 days, then placed side up in liquid nitrogen for 5-10 s and transferred on LacZ-plates and incubated at 30°C. The time of incubation is variable.

2.2.12 Protein purification from *E. coli*

Combination between batch and column preparation methods for BBR::*GST* purification from *E. coli*.

The vector system used to overexpress BBR in *E.coli* is pGex-5X (Pharmacia; 2.1.4).
From a positive clone overexpressing, under induction, BBR*::GST* fusion protein, cells were grown in 10 ml LB medium o.n. and then used to inocule 100 ml of fresh medium. They were grown for 1 h, then induced with 0.5mM of IPTG and grown for other 4 h. At this point the cells were collected washed with STE buffer and the pellet frozen at –20°C.
The frozen pellet was resospended in 1X STE buffer with 100 µg/ml Lysozym, mixed, incubated in ice for 20 min. Subsequently 12 µl of 0.5M DTT and 180 µl of 10% sarkosyl in 1X STE were added mixed and placed in ice. Afterwards the preparation has been sonicate for 1 min, centrifuged at 10,000 rpm for 5 min., then the supernatant transferred in a new tube and 700 µl of glutathione sepharose beads (50% slurry) added. This mix was shaken for 30 min at 4°C to promote binding of the proteins to the beads. The matrix was placed on a column with a normal polyethylene filter, let sediment and washed two times with cold PBS buffer. Finally the elution buffer was applied and left o.n. to equilibrate with the matrix. The next day two successive elutions steps were performed.
An aliquot of the purified protein was checked on SDS-PAGE gel, the rest dialysed against 5 L of 2X G1 binding EMSA buffer.

Western blot analysis (Sambrook *et al.* 1989)

Protein were run overnight on a discontinuous SDS-PAGE gel, electro-blotted in blotting buffer to supported nitrocellulose membrane and blocked overnight at 4°C in 5% milk TTBS. The first antibody was diluted in the blocking solution and incubated together with the filter at room temperature during 2 h. After three washes of 10 min in TTBS the secondary antibody diluted 5000 fold in blocking solution, is added to the filter and incubated as before during 45 min. After three washes of ten min in TTBS, the detection is done by means of the ECL system from Amersham, followed by autoradiography.

2.2.13 EMSA

Once the protein is purified and dialysed what is needed for the gel retardation assay is to produce the probe DNA, set up the binding reaction and electrophoretically separate the complexes.
A fill in reaction and a terminal labelling procedure were used to generate probes respectively from 5'-protuding digested DNA and from oligonucleotides.
Fill in reaction: 20 µCi of the desired [α^{32}P] dNTP were used together with 1µl of 5mM mix of the other 3 dNTPs, 1U of Klenow enzyme, around 100 ng of template DNA and 2 µl of 10X Klenow buffer. The reactions were placed at 30°C for 15 min.
Teminal labelling: 150 µCi [γ^{32}P]ATP were used together with 3µl of 10X T4 polynucleotide kinase buffer, 1 to 50 pmol dephosphorylated DNA 5' ends, 20 U of T4 polynucleotide kinase and 50µg/ml BSA. The reactions were incubated at 37°C for 60 min.
The binding conditions can modify efficiency and specificity of binding. For this reason several buffers were tested. The G1 buffer was chosen for BBR according to the first shift revealed with the *in vitro* transcribed and translated proteins. In the table below are summarized the buffers tested and their composition.
Different amount of proteins, from 10 to 50 ng, were tested for binding to the labelled DNA.
The reaction was conducted at 4°C for 20 min.
Proteins, DNA and complexes were separated on a 4% native polyacrylamide gel, using 0.2X TAE as buffer system. After a pre run the sample were loaded and run for 6 h, then the gel was transferred on 3MM paper, vacuum dried and used to expose a film for autoradiografy.

	EDTA	TRIS/HCl	HEPES	KCl/$MgCl_2$/Nacl	glycerol	*b*-mercapto/ DTT	PMSF	Poly(dI-dC)
ORCA	0.1	-	25mM pH 7.2	100 mM KCl	10%	-	-	10ng/µl
G1	0.5	10mM pH 7.5	-	1mM $MgCl_2$/50 mM NaCl	4%	0.5mM DTT	-	50ng/µl
G2	1	-	10mM pH 7.9	50 mM KaCl	10%	1 mM DTT	-	50ng/µl
WB1	1	10mM pH 7.5	7.5mM pH 7.5	-	4%	10mM β-mer.	-	50ng/µl
WB2	0.07	4mM pH 7.5	-	60 mM KCl	7%	3.5mM β-mer	0.05 mM	50ng/µl
GT1	-	20mM pH 7.9	-	50mM NaCl	10%	0.1mM DTT	-	-
GT2	1	20mM pH 7.9	7.2mM pH 7.3	2mM $MgCl_2$/50mM NaCl	10%	1mM DTT	-	-
AM	1.13	-		-	-	0.7 mM DTT	-	-

2.2.14 Transient expression analysis in *N. tabacum* SRI protoplasts

Protoplasts isolation

Tobacco plants were grown under sterile condition until they reached about 7 cm (the leaves show a diameter of 4 cm at this stage). Leaves were cut and transferred to flasks containing 50 ml K3 0.4M cellulase-macerozyme. The cellulase/macerozyma digestion was performed for 20-22 h in the dark at 28°C. After overnight incubation, the flasks were shaken slowly for 15 min at room temperature. The protoplast/debris solution was transferred via sterile pipettes on 100 µm mesh screen and the debris was filtered out. The flow-through was transferred to 50 ml Falcon tubes. These tubes were centrifuged at 250 rpm for 5 min. The protoplast remained at the top of the solution and the debris at the bottom forming the pellet. A capillary pump was used to get rid of the debris and the interphase. The protoplast band was resuspended very carefully in W5 to a volume of 50 ml and the

Falcon tubes were centrifuged at 500 rpm for 5 min. The pellet of the protoplasts was resuspended again in W5. Centrifugation and resuspension in fresh medium was repeated 2 times. The protoplast titre was determined with a Fuchs-Rosenthal chamber. Protoplasts were pelleted at 500 rpm for 5 min and resuspended to 1,000,000 protoplasts per ml in MaMg solution.

Protoplast transformation

The protoplast transformation was performed in 12 ml Falcon tube as following:
50 µg carrier plasmid and 5 µg each construct were added to 330,000 protoplast (330 µl), after an incubation step of 10 min at room temperature, 660 µl PEG solution (40%) was added and a second incubation step of 30 min at room temperature was performed. Finally 5 ml of K3 0.4M (250 mg/L Cefatoxime) for transient incubation was introduced.
The mix was incubated for two days at 28°C in the dark.
After 48 hrs, 6ml of W5 was added to ensure that the protoplasts were pelleted down. The samples were centrifuged for 5 min at 5,000 rpm. The supernatant was removed, except the last 1ml, and then the suspension was transferred to a 1,5 ml Eppendorf tube. The tube was microcentrifuged at maximum speed for 2 min. The supernatant was removed and the pellet was frozen in liquid nitrogen and was stored at -80°C.

GUS assay on protoplast cells

To determine reporter enzyme activity, protoplast pellets were througly vortexed for 30 s with 800 µl of LUC extraction buffer. If only GUS activity was to be determined, LUC extraction buffer was replaced by GUS extraction buffer. The resulting suspension was centrifuged for 10 min at 4°C 13,000 rpm. The protein content was determined using the Bradford assay (Bradford, 1976) and BSA as a standard. Luiciferase (LUC) activity was determined in 10 µl of the protoplast extract by addition of 100 µl LUC assay buffer and subsequent determination of the photons emitted over 10 s. In order to determine the GUS activity, 72 µl of supernatant was mixed with 8 µl of substrate solution 10mM 4-methylumbelliferyl-d-glucoronide (MUG) and incubated for 30, 60 and 120 min at 37°C. The reaction was stopped by adding 1 ml of 0.2M Na_2CO_3. The enzyme activity was measured by fluorimetric assay using excitation and emission wavelenghts of 365 and 455 nm, respectively. Specific GUS activity is given in pmol 4-MU per µg protein per minute. Standardised specific GUS activity was calculated by division of each amount of the GUS activity by the luciferase reading in each extract (pmol 4-MU /min/lu/10 s).

2.2.15 *Agrobacterium* mediated transformation of *N. tabacum* SRI leaf disks

Agrobacterium transformation

The binary vector was introduced into the *Agrobacterium* desired strain using electroporation (Bio-Rad, GenePulser) in 0.1 cm cuvettes having a field strength of 1.25 KV a capacitance of 25 μF and a resistor of 600 ohm in parallel. After a recovery period of 2h, the bacteria were plated on solid YEB medium and selected according to the vector system used. A positive colony was grown in 5ml YEB medium with selective antibiotics for two days, 4 ml of this preinocule were inoculated into other 80 ml of YEB selective medium.
The *Agrobacterium* was grown overnight until an OD_{600} of 0.1 then the suspension was used to infect tobacco SRI leaf disks.

N. tabacum leaf disks transformation

Adult leaves were cut into small pieces ($1cm^2$) in order to have around 70 fragments each construct to transform. The fragments were placed on solid MS medium with BAP 1 mg/L, MES 0.5 g/L, NAA 0.1 mg/L (MSI) and incubated for two days in growth chamber. Subsequently they were incubated for 20 min with the *Agrobacterium* suspension to generate the infection and replaced on the MSI solid media. After two days the fragment were transferred on MSI medium containing Claforane 500mg/ml to stop the bacteria's growth. From this point on the fragment and the calli on them were transferred weakly on fresh plates, containing the plant selective antibiotic, until the first green shoots were produced.

2.2.16 Hystochemical GUS analysis

Tissue material was infiltrated and incubated with 100mM sodium phosphate buffer (pH 7.0) containing 1mM X-Gluc, 0.1 Triton X 100 and 8mM *b*-mercaptoethanol at 37°C overnight. Then the tissue was fixed with 1% Glutaraldehyde in 50mM posphate buffer (pH 7.0), distained with ethanol 100% and stored in 50% glycerol at 4°C.

3 Results

3.1 Analysis of the *Bkn3* intron IV/305 enhancer using GUS reporter gene activity in tobacco transgenic plants

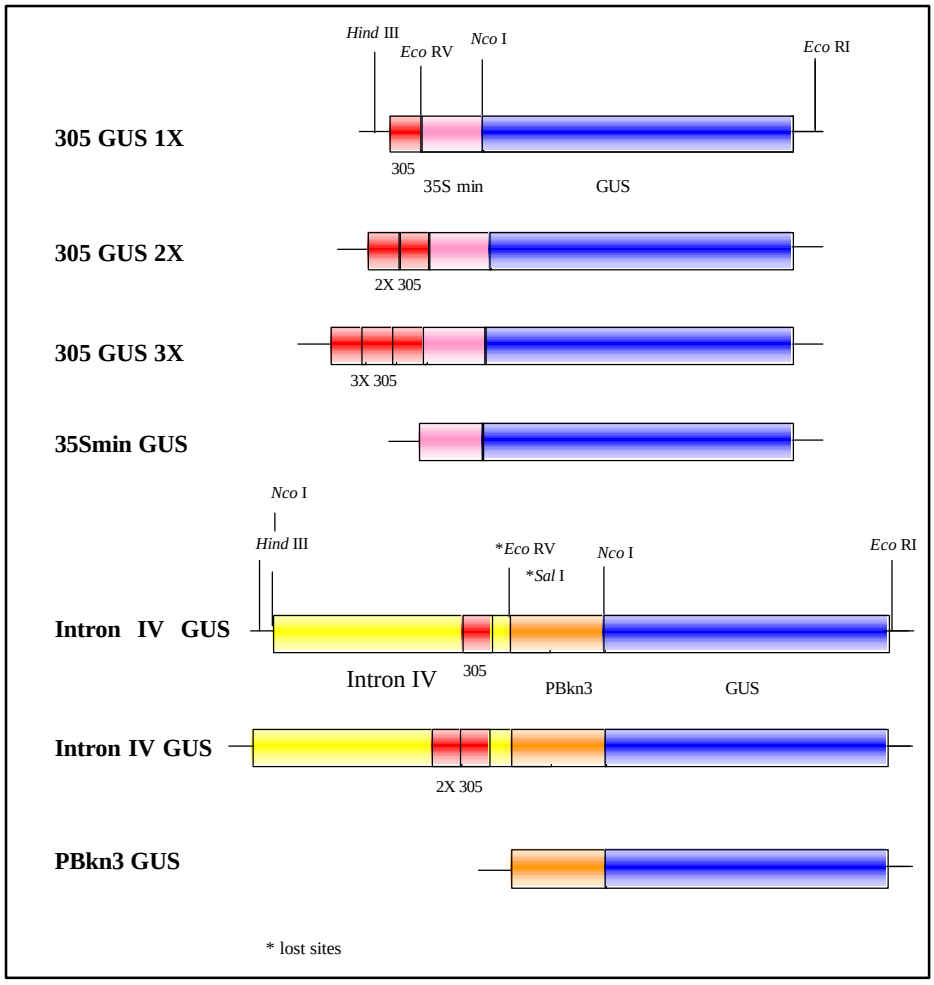
The duplication of 305bp in the large intron IV of *Bkn3* is the cause, at the molecular level, of the *Hooded* (*K*) phenotype. *Hooded* is a dominant, gain of function mutant. The direct duplication influences the expression of *Bkn3* ectopically expressed in the meristematic cushion from which an extra flower develops. This has suggested that the 305bp could mediate some regulatory processes influencing the expression of *Bkn3* itself. If so, the intron IV and the 305bp fragment would act as an enhancer like sequence that, through interaction with protein partners, regulates the expression of the gene. To test this hypothesis, several promoter GUS constructs were produced based on the 305bp alone, as well as in duplicate and in triplicate and in the context of the endogenous intron. The constructs were assembled in PUC vectors, then moved into the pBIN19 binary vector (Frish *et al.*, 1995) suitable to mediate integration in the tobacco genome through *Agrobacterium tumefaciens* mediated transformation. The plants were analysed for GUS activity.

3.1.1 The intron IV/305bp GUS reporter vectors system

The constructs can be divided in two different groups. In the first set, the 305bp sequence in one, two or three copies was cloned upstream the 35S minimal promoter from the cauliflower mosaic virus (CaMV). The minimal promoter was defined, with respect to the transcription starting point, as the –46 / +9 portion of the full size promoter (-941 to +9). It contains the TATA box region and several studies indicate for this promoter a very low basal level of expression (Fang *et al.*, 1989). In the second set, the whole intron IV (2.kb) with one or two copies of the 305bp were cloned in front of the *Bkn3* "minimal promoter" (P*Bkn3* 0.655), which is a region of 655bp upstream of the ATG translational start codon in the *Bkn3* gene. The part of the intron IV used extended in the intron until the *Eco*RV site located 3' downstream to the 305bp sequence.

The 305bp sequences were isolated, using the polymerase chain reaction, from constructs 39, 40 and 41, where they have been already cloned in one two or three copies respectively by Kai J. Muller. The primers used (305H, 305EV) were designed on the sequences flanking the 305bp sequence and *Hind*III and *Eco*RV restriction sites were artificially inserted. The inserts were cloned in the *Hind*III and *Eco*RV sites of the K373 reporter vector (2.1.4) upstream the 35S minimal promoter and the GUS gene.

For the second set of constructs the intron sequence was isolated from *Bkn3* genomic sequences as an *Nco*I-*Eco*RV fragment in wild type and *K*, to obtain one or two copies respectively. Using the same restriction enzymes, the fragment was also cloned in the K373 reporter vector. Subsequently, the P*Bkn3* 0.655 was exchanged with the 35S minimal promoter of K373. The promoter was isolated using S*alI* and N*coI* restriction sites and cloned between the P*Bkn3*and the *GUS* gene using on one side the *Nco* I site and on the other a blunt end generated by fill in reaction on the digested *Sal*I site.



Schematic representation of the intron IV-GUS constructs.

**Eco* RV
**Sal* I

3.1.2 *N. tabacum* transformations and GUS analisys of T_0-T_1-T_2 plants

The expression cassette of all the constructs was cloned from the PUC derivative vector to pBIN19 binary vector (Frisch *et al*, 1995). All the inserts were gel cleaned after a *Hind*III-*Eco*RI digestion and ligated into the binary vector using the same restriction sites. Following transformation into *E. coli*, the plasmids were selected on kanamicyn, propagated, controlled and used to transform *Agrobacterium tumefaciens* strain LBA 4404 utilising electroporation. The *Agrobacterium* strain LBA 4404 carries rifampicillin and streptomycin resistance and the transformed clones are also kanamycin resistant due to the pBIN19 selectable marker. The *Agrobacterium* transformed clones, carrying the different constructs, were selected on the three antibiotics, grown until optimal density and used to infect leaf disks of *Nicotiana tabacum* SRI wild type as described in materials and methods (2.2.15).

Ten resistant calli for each construct were selected on MS medium with kanamycin at 100mg/ml. The plants were grown until flowering, afterwards the following tissues of the primary transformants were screened for GUS activity: Vegetative apices of main and lateral shoots; buds of lateral shoots; leaves close to apices; apices with early inflorescences; later inflorescences and flowers. The plants carrying the same construct were divided into three classes of expression according to the time needed to develop GUS staining, as well as based on a mutual comparison of the GUS activity within independent transformants. GUS activity was evaluated as strong, medium and weak; strong GUS activity resulted in blue staining of plant material within one hour of incubation with the vacuum-infiltrated GUS substrate. Medium GUS activity required staining for 3 to 6 hours and weak GUS activity was attributed to transformants that had to be incubated over night. Differences in the expression levels among plants carrying the same construct can be attributed to positional effects due to different integration sites in the tobacco genome. One plant for each expression class, and for each different construct, was selected and the T_1 and T_2 generations were tested for GUS activity. The analysis of the progenies validated the T_0 analysis, with the exception of three lines where GUS staining was not any longer detectable in all the analysed members of the progeny.

Strong *bkn3* promoter activity was detected in the shoot apex, while only medium 35S minimal promoter activity was detected in the roots of T_1-plants. The expression domain of the shoot apex was characteristic for 305-35S minimal and intron IV-*Bkn3* promoters. Since no differences in their expression domain became obvious they will just be indicated 305-GUS constructs. During inflorescence and flower development, the 305-GUS constructs activated reporter gene expression in young inflorescence apices, while the 35S minimal and the *Bkn3* promoter showed no activity.

Two expression domains during reproductive development were found exclusive for 305-GUS constructs: The branching points within the tobacco inflorescence and the base of the flower. During vegetative development the first expression domain corresponded to the position where flowers or lateral inflorescence shoots branch off. GUS staining at the base of the flower was located at the base of petals and carpels and stained the whole base of the flower (Fig. 3.1.1).

Expression in the main shoot apex and in the emerging lateral vegetative buds was common for intron IV-GUS and for the isolated B*kn3* promoter, thus reflecting overlapping expression domains. Unlike the promoter of *Bkn3*, the intron IV and the 305bp element responded with enhanced expression in lateral shoot apices. In the tobacco plants these expression domains are the sites of organ additions or neo-meristem formation.

	STRONG	**MEDIUM**	**WEAK**
305 GUS 1X	2,**6**,7	1,8,**9**,10	3,4,**5**
305 GUS 2X	**3**,5,6,7	**2**,8,10	**1**,4,9
305 GUS 3X	**2**,6,8,9	**3**,7,10	**1**,4,5
Intron IV GUS 1X	**1**,5,9	**6**,7,8,10	**2**,3,4
Intron IV GUS 2X	**6**,8,9,10	**2**,4,5,7	1,**3**
P*Bkn3* GUS	-	-	1-**10**

Tab. 3.1 T_0 transgenic tobacco plants harbouring the intron IV GUS constructs. One plant for each construct and for each expression class was selected (bold).

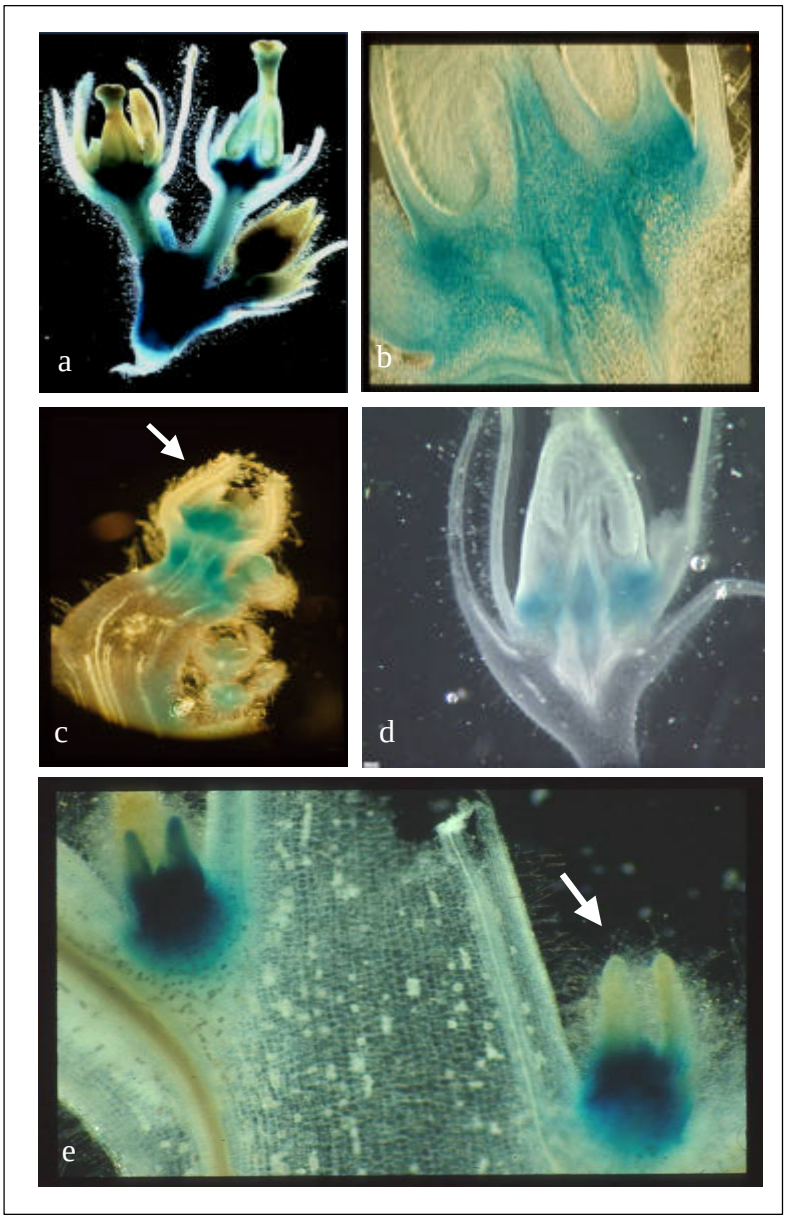


Fig. 3.1.1 GUS expression detected in transgenic tobacco plants harbouring the intron IV-GUS constructs. Enzyme activity is localised in vegetative and reproductive active meristems, a-d) at the base of the flower and e) at the emerging of vegetative side branches. Expression is absent from organ primordia as indicated by the arrows.

3.2 cDNA-Library screen: additional sequence information and homologies of BEIL, BBR, BGRF and BAPL

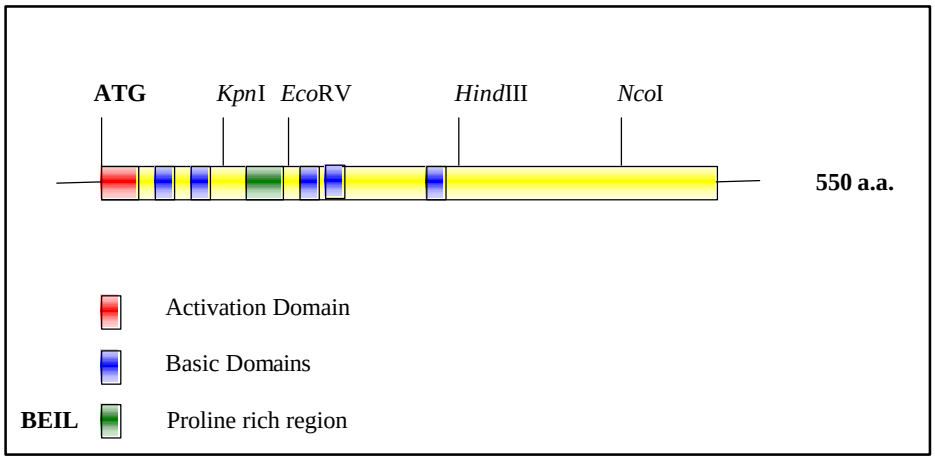
The 305bp sequence has been used as target element in a one-hybrid screen (K, Muller, unpublished) to identify binding factors participating to the regulation of *Bkn3*. Four coding sequences were the results of the screening and the proteins were named BEIL, BBR, BGRF and BAPL. Two different clones were identified for BEIL, four for BBR and one for BGRF and BAPL.
A cDNA library, generated from the *K*-Atlas barley mutant, was screened in order to acquire further information on the predicted full open reading frames and the 3' and 5' untranslated regions (UTR). The library has been previously produced in λNM1149 (Frischauf *et al.*, 1983) using an RNA template from inflorescences developing the meristematic cushion, typical of the *K* mutant. λNM1149 is a λ phage insertion vector suitable for cloning small (+/- 8kb) cDNA fragments in the unique *Eco*RI site of the immunity region. The screening for the four genes was carried on in parallel using probes derived from the four "one-hybrid sequences". The four inserts were isolated from the pAD-Gal4 vectors and used as template for random priming labelling, as described in materials and methods (2.2.6). The library titre was 0.6 pfu/ml. To obtain around 1,000,000 pfu, 20µl were used to infect *E.coli* strain POP13 and plated, for the first round, on 50 plates for each gene. The screening proceeded through an additional three rounds. Twenty single plagues of different hybridisation intensity were isolated for each gene. The phages were grown in liquid medium and the barley DNA sequences were extracted from them. Southern blot and restriction digestion investigations were conducted on the clones to avoid duplicates and false positives. The longest clones were selected, subcloned undirectionally in pBlu KS using the *EcoRI* site and sequenced.

Clones	λ-cDNA	One-hy-cDNA
BEIL	1.5 Kb	1.5 Kb
BBR	1.63 Kb	1.25 Kb
BGRF	1.6 Kb	1.041 Kb
BAPL	1.5 Kb	1.6 Kb

Tab. 3.2.1 Comparison of the longest cDNA clones isolated for BEIL, BBR, BGRF, and BAPL from their respective library screen.

3.2.1 Sequences and homologies

BEIL: The start and the stop of the translation still remain undefined. The cDNA was of the same length and mostly overlapping with the one-hybrid sequence, but shifted 132bp at the 3' end.
The first methionine found in the one-hybrid cDNA is likely to be the ATG start site for homology reasons, but nothing is known on the 5' UTR. Even if with the screening the predicted ORF was extended 44 a.a., no stop codon was identified leaving the a.a. sequence incomplete at the C-terminus; the partial ORF is 550 a.a. long.



Schematic representation of the BEIL ORF.

A database search was performed using the BLAST (Altschul *et al.*, 1997) for protein sequences. The three highest homologous proteins found were ethylene insensitive like 3 (EIL3) from *Arabidopsis thaliana*, showing 63% identity and two other EIL3 like proteins from *Cucumis melo* and *Lycopersicon esculentum*, showing 58% and 57% identity respectively (Fig.3.2.1). EIL3 acts in the ethylene pathway and is supposed to work with proteins of the EREBP subfamily to regulate gene trascription.

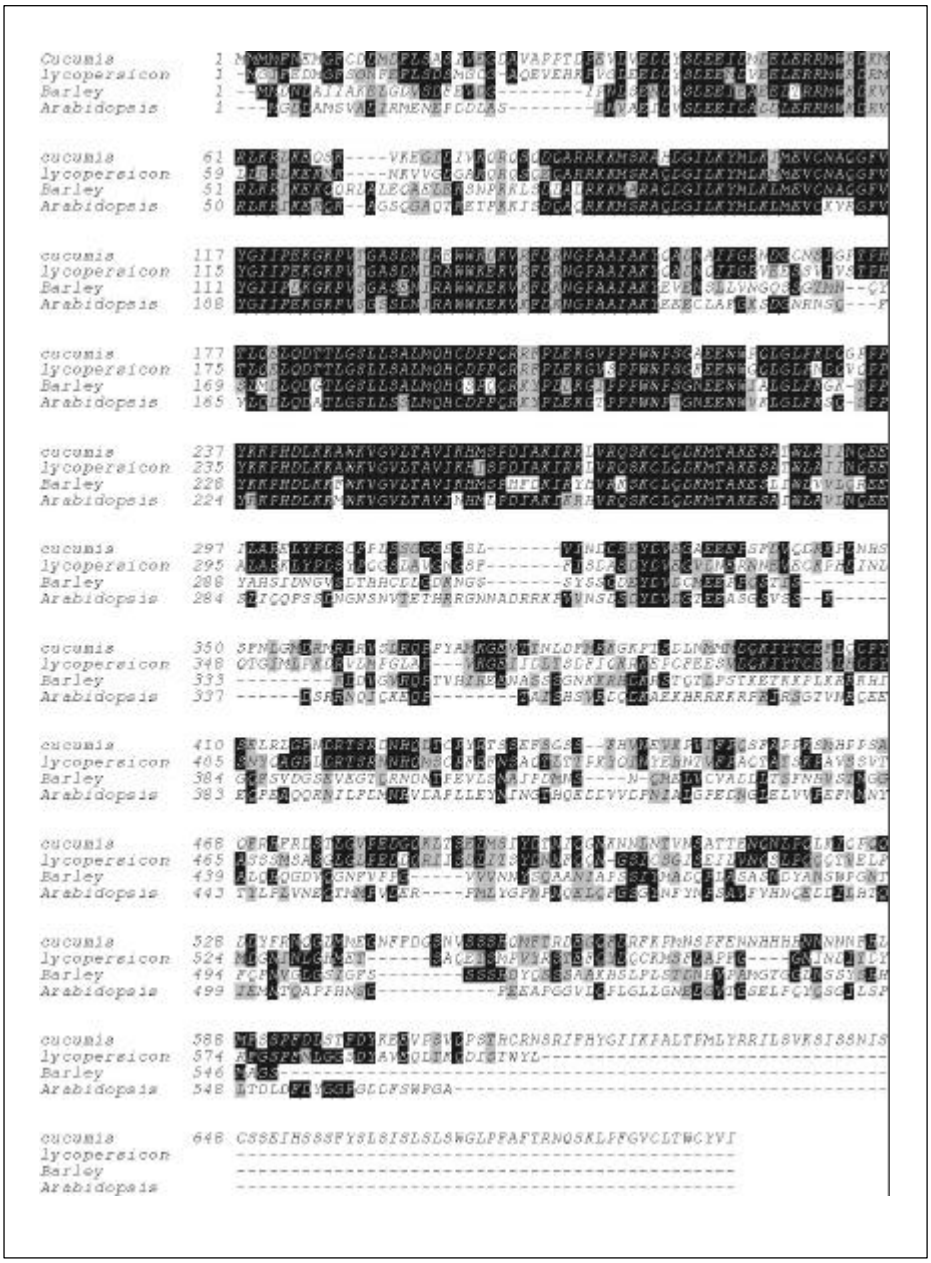
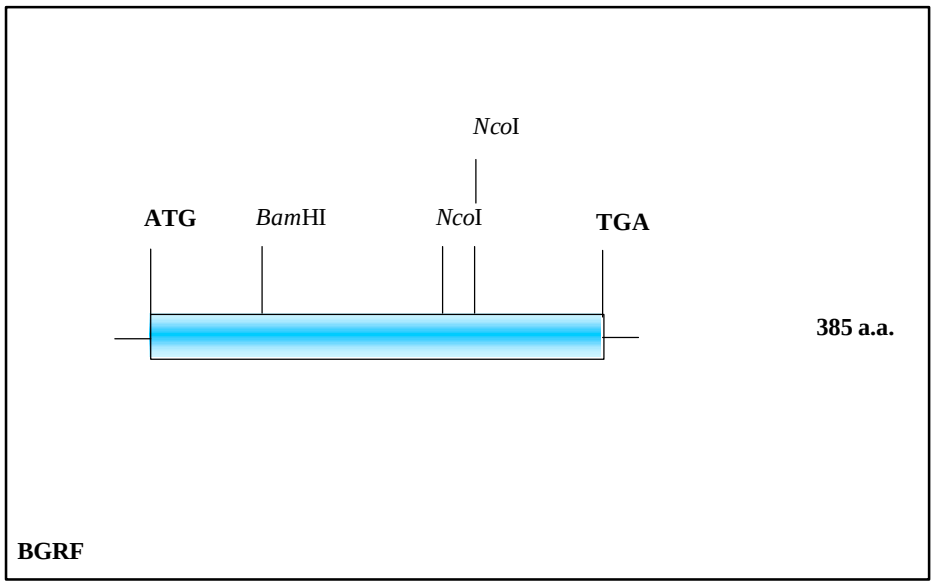


Fig.3.2.1 Alignment of the amino acid sequences of BEIL from Barley and the highest homologous sequences identified through a BLAST search (Altschul *et al*., 1997); one sequence is from *Arabibopsis*, one from *Lycopersicon esculentum* and one from *Cucumis melo*. The sequences were aligned using ClustalW and Boxshade programmes at the gene regulation web site www.gene-regulation.de.

BEIL, as the homologue of *Arabidopsis*, has five small clusters of basic amino acids dispersed throughout the sequence at the following positions: 56-69 (basic domain I, BD I); 91-97 (BD II); 238-248 (BD III); 265-274 (BD IV); and 378-384 (BD V). α-helices are predicted in BD II and BD III, whereas amino acids 199-240 comprise a proline rich domain (Chao Q. *et al.* 1997). The amino terminal half of the BEIL peptide is more conserved than the carboxy-terminal regions.

BGRF: The cDNA of 1.6 Kb, which completely contained the 1.050 Kb sequence of the one-hybrid cDNA, revealed a new 90 bp on the 5' and new 420bp fragment on the 3' of the sequence. With this new sequence information it was possible to define the starting methionine and the stop codon that delimit this protein to an ORF of 385 amino acids. The clone also contains 78bp of 5' UTR and 300bp of 3'UTR.



Schematic representation of the BGRF ORF.

A search for protein sequence homology was performed using BLAST (Altschul *et al.*, 1997). The three highest homologous proteins found were two predicted proteins from the full genome sequence of *Arabidopsis thaliana* showing 63% and 50% homology, and one "growth regulating factor" from *Oryza sativa* (Fig. 3.2.2). This last gene is a novel putative gibberellin-induced gene from rice involved in the regulation of the stem growth (unpublished).

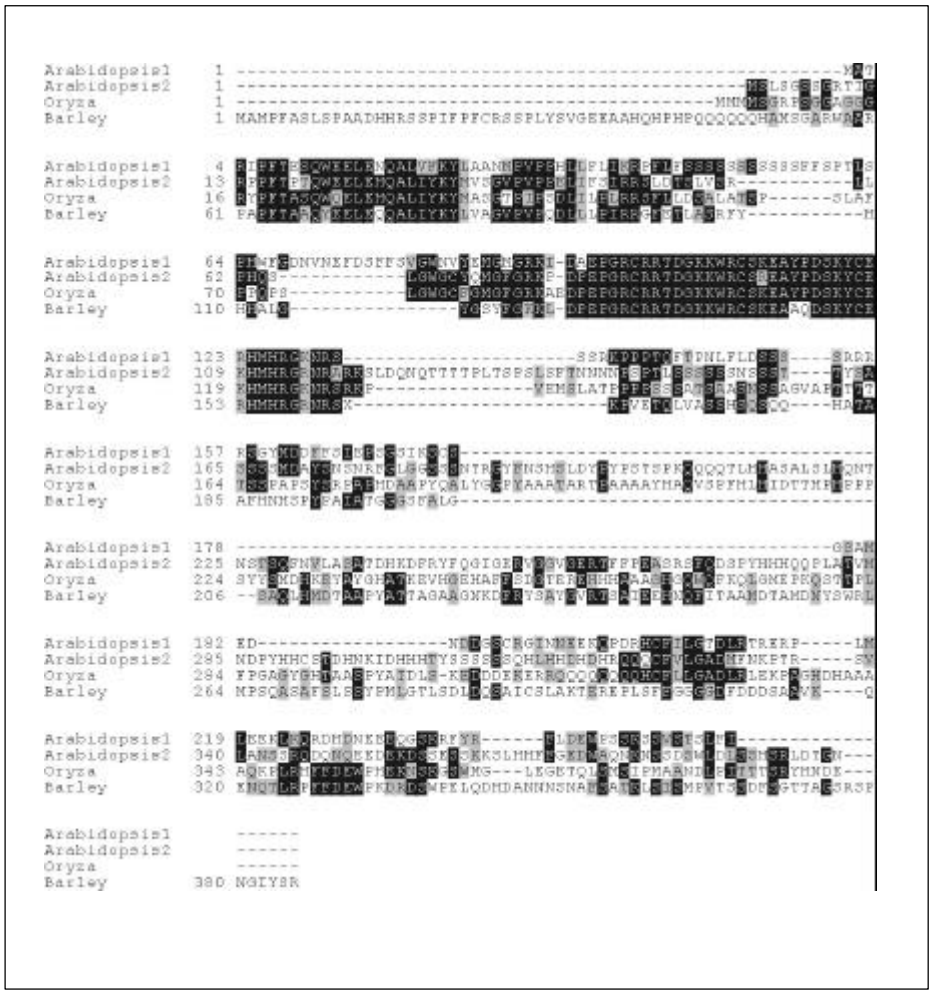
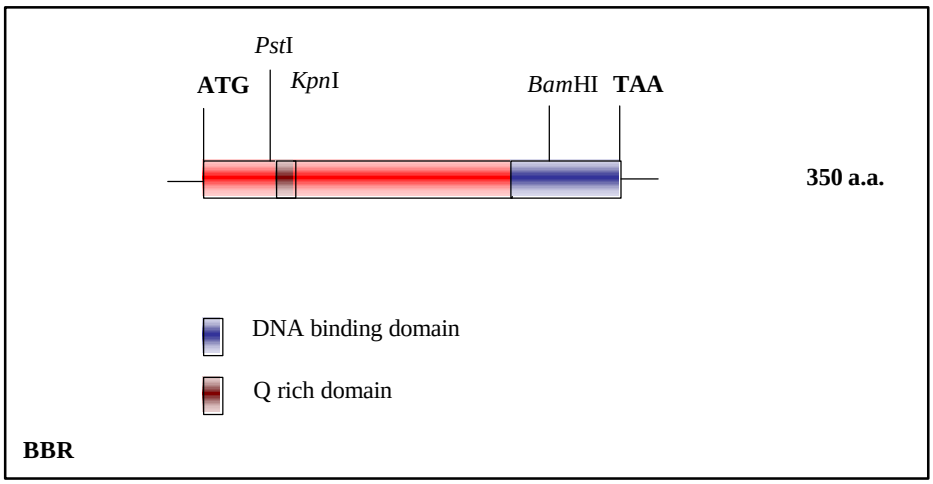


Fig. 3.2.2 Boxshade alignment of the BGRF protein with two homologous putative proteins from *Arabidopsis* and one from rice.

BBR: Also for this gene, the cDNA clone was longer than the previous clones and the new sequence revealed 183bp on the 5' end, of which 75bp were encoding amino acids, while 108bp were part of the 5' UTR. The first methionine is surely defined: being the first one presents after a stop codon in frame at the 5' UTR, and for analogy to the homologous sequences found in the database (Fig. 3.2.3). The full ORF consists of 350 a.a. and the 3' UTR is of 460bp. The BBR sequence has been cloned independently four times using the yeast one-hybrid system, the longest clone encodes for 325 a.a. of the protein: 25 a.a. are missing at the N-terminus, compared to the full

length protein sequence. The shortest encodes for the last 128 amino acids at the C-terminus, while the other two sequences code for two proteins of 161 and 244 amino acids, respectively, upstream from the C-terminus of the protein.



Schematic representation of the BBR ORF.

BBR has a homology of 60% with a putative nucleic acids binding protein from *Oryza sativa* (unpublished data) and of 74% and 49%, respectively, with two putative proteins in *Arabidopsis thaliana*. As shown in figure 3.2.3, the homology is particularly high for the last 120 amino acids at the C-terminus of the protein. This, together with the fact that this region of the protein is functional in the yeast one-hybrid -the shortest functional clone consists of the last 128 a.a.- clearly suggests that this is the DNA binding domain or that this region contains the DNA binding domain. Nothing is known from the literature on such a kind of domain. The amino acid composition suggests DNA binding properties as it is relatively high in the basic amino acids K, R and H. In yeast, BBR is able to activate the expression of a reporter gene proportionally to the repetition of the target sequence, from 1 to 3 times. Characteristic feature of this protein is the long stretch of QH residues from position 63 to position 123. Glutamine rich domains are reported to be responsible for protein-protein interactions of GAGA factors (Wilkins *et al.*, 1999), so essential for homo and hetero associations and formation of multicomponent complexes. For example, the products of the Groucho related gene (grg) family can dimerise through their amino-terminal Q domain, and glutamine domains have been shown to mediate stable dimerisation *in vitro* (Pinto & Lobe, 1996).

In the *Drosophila* GAGA factor the Q domain alone binds single stranded DNA and appears to form stable tetramers in solution (Wilkins and Lis, 1999).

BAPL: This was the longest cDNA clone from the whole screening: 3.4 Kb. Unfortunately sequence and BLAST analysis revealed that it was a chimeric cDNA, comprising two different clones joined by an *Eco*RI site. Of the 3.4 Kb, 1.5 Kb belonged to BAPL and the rest (2 Kb) to the *Hordeum vulgare* tonoplast intrinsic protein 2 (TIP2). The latter has not been characterised further.

The BAPL cDNA sequence exceeded the previous clone by 12bp at the 5' end. The starting ATG has the same position as in the homologous clone from *Oriza sativa*. The full ORF consists of 360

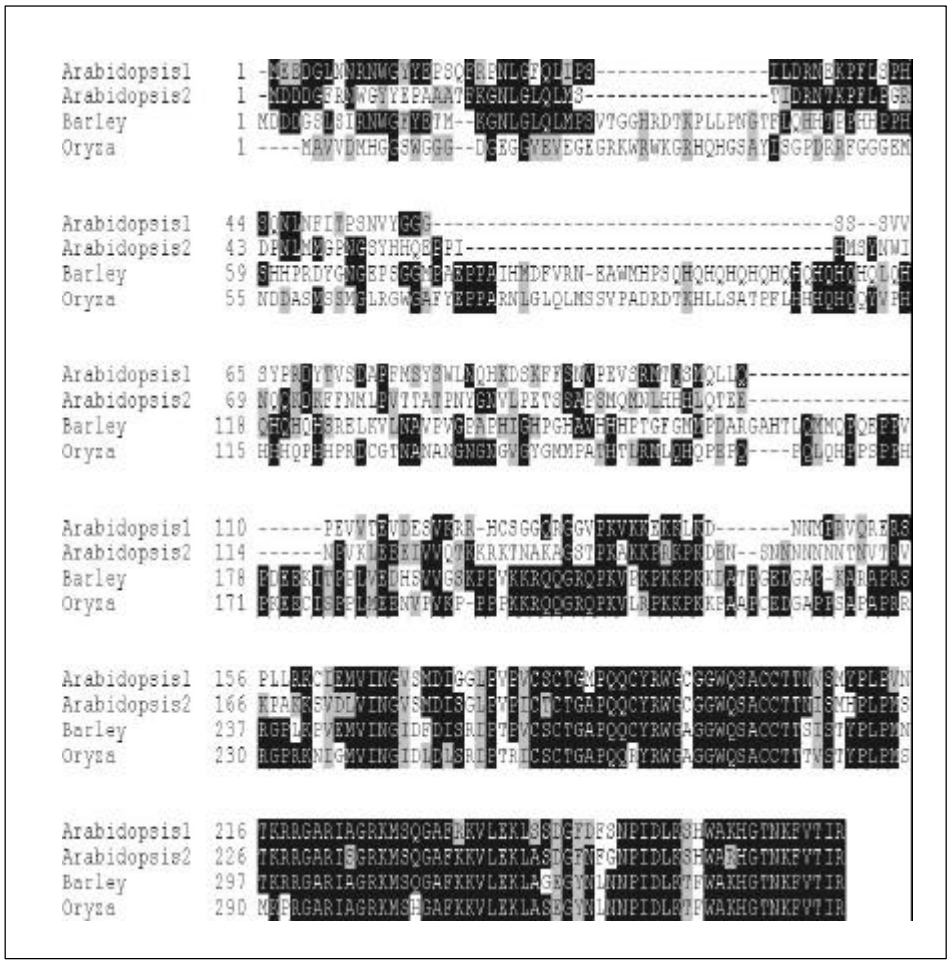
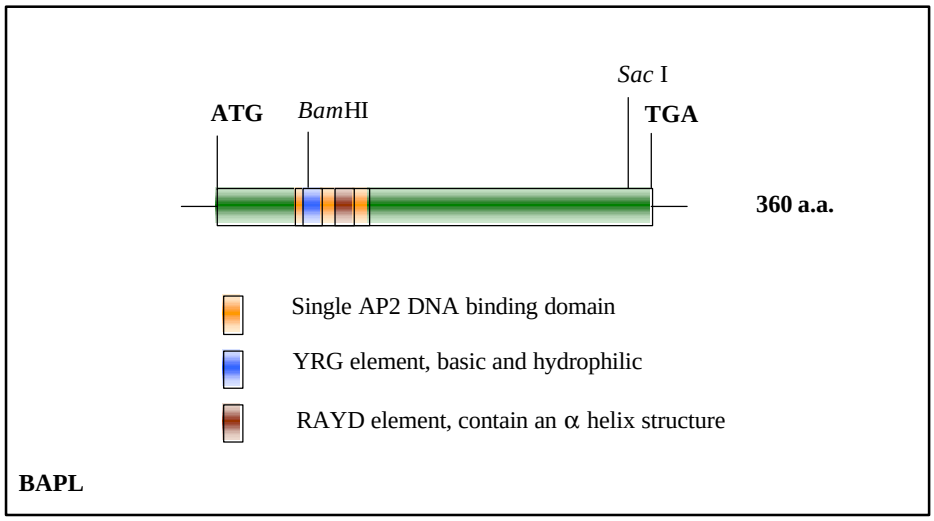


Fig. 3.2.3 Amino acid sequence alignment of BBR and the homologous sequences identified through a BLAST search (Altshul *et al.*, 1997). The region defined by the last 120 a.a. at the C-terminal end, which shows the highest level of homology, contains the DNA binding domain.

amino acids and database homology investigations show that has an ethylene responsive element binding protein domain (EREBP domain) that is a clearly defined DNA binding domain.



Schematic representation of the BAPL ORF.

BAPL shows homology to an ethylene responsive element binding protein from rice and two other proteins containing an APETALA2 (AP2) domain, one from *Prunus armeniaca* and one from *Arabidopsis thaliana* (Fig. 3.2.4). AP2 (APETALA2) and EREBPs (ethylene-responsive element binding proteins) are the prototypic members of a family of transcription factors unique to plants, whose distinguishing characteristic is that they contain the so called AP2 DNA binding domain. *AP2/EREBP* genes form a large multigene family, and they play a variety of roles throughout the plant life cycle: from being regulators of several developmental processes, like floral organ identity, to form part of the mechanism used by plants to respond to various types of biotic and environmental stresses. The peculiar characteristic of proteins of the AP2/EREBP family is that they contain either two or one DNA binding domains respectively. The AP2 domain was first recognized as a repeated motif within the *Arabidopsis thaliana* AP2 protein (Jofuku *et al.*, 1994). Shortly afterward, four DNA binding proteins from tobacco were identified that interact with a sequence that is essential for the responsiveness to the plant hormone ethylene and were designed as **e**thylene **r**esponsive **e**lement **b**inding **p**roteins (EREBPs).

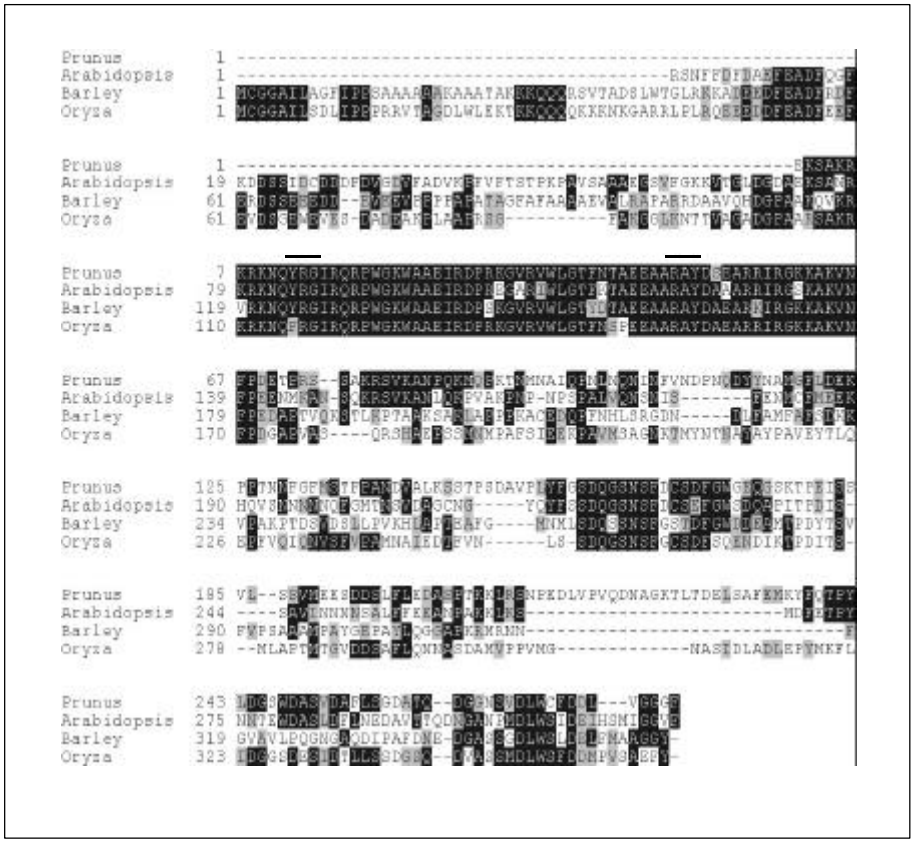


Fig. 3.2.4 Protein sequence alignment of BAPL together with proteins belonging to the AP2/EREBP family of transcription factors from different plant species: *Prunus armeniaca*, *Arabidopsis thaliana* and *Oriza sativa*. The region of highest homology defines the AP2 DNA binding domain. The overhead-lines indicate the YRG and RAYD elements typical of this domain.

This multigene family can be divided into two subfamilies based on whether the proteins contain one or two AP2 domains, denoting the AP2 or the EREBP subfamily, respectively. Two mostly conserved amino acid segments found within each AP2 domain, are referred to as the YRG element and the RAYD element. The amino terminal part of the AP2 domain (the YRG element) is basic and hydrophilic. The carboxyl RAYD element contains a central region that in almost all the AP2 domains is predicted to adopt the configuration of an λ-helix of amphipatic character. As all the other members of both subfamilies, BAPL has five amino acid residues absolutely conserved. These are: G at position 126, G at 133, Y at 163, G at 172 and N at 178. The conservation of these residues, obviously suggests the importance for the structure/function of the AP2 domain, of which very little is known.

3.3 Two hybrid assay: self and hetero interactions of BEIL, BGRF, BBR and BAPL

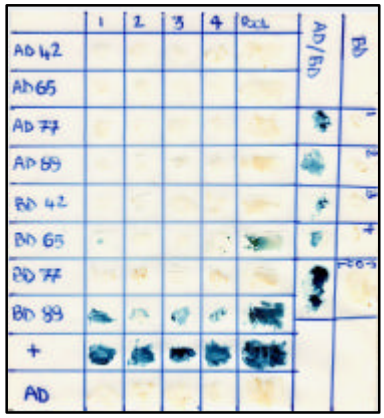
The four fragments, which encode for the four different proteins, were isolated by digestion with *Eco*RI and *Pst*I restriction enzymes from the pAD-GAL4 vector used in the one-hybrid screen and subcloned into pBD-GAL4 using the same restriction sites. Since these two different sites have the same frame in both vectors, all four constructs resulted in frame with the *Gal4* BD.

Transformation of *Saccharomyces cerevisiae* strain YRG-2 was carried out as described in materials and methods (2.2.11). The YRG-2 strain contains a dual selection system with *lacZ* and *HIS3* reporter gene constructs, which can be activated only in presence of a specific protein-protein interaction. LEU2 and TRP1 are the transformation markers. Before testing the eight constructs, several control transformations were performed. All constructs were transformed alone, as well as the empty pAD and pBD vectors and two positive controls: the pGAL4 plasmid, which contains the entire coding sequence of the wild type GAL4 protein (AD/BD) and p53 plasmid, which contains a.a. 72-390 of the murine p53 cloned in the pBD-GAL4 vector (+). Transgenic yeasts colonies were selected for their nutritional requirement on solid media lacking tryptophan (TRP) for the pBD derivatives, while leucine (LEU) deficient plates were used for the pAD derivatives and the pGAL4 construct. Four independent colonies and a pool of the remaining were tested for galactosidase gene activity on Hybond-N membranes (Amersham). Results obtained, included those of control experiments, are shown in figure 3.3.1 (a/b).

All constructs proved to be suitable for testing interactions except BD-BGRF and BD-BAPL which activated the reporter system per se. This may occur if the bait is a transcriptional activator or contains a region of amino acids which are highly acidic. The 6 remaining constructs were tested in all possible combinations pair wise among each other and with the empty pBD-GAL4 and pAD-GAL4. Yeast strain YRG-2 was transformed, transgenic colonies selected on solid media lacking Leu and Trp and bulks of these positives colonies were tested for galactosidase gene activity. The filters and the results of control experiments are summarized in figure 3.3.1 (c/d).

Interactions between proteins were not evident except for BEIL homodimerisation (AD-BEIL/BD-BEIL). This interaction can be considered spurious because the transformation of BD-BEIL sing the empty pAD-GAL4, also resulted in β-gal activity. BEIL was in fact capable to activate the reporter in the one-hybrid system without fusion with the Gal4AD in the chimeric protein. This demonstrated that BEIL contains an activation domain functional in yeast

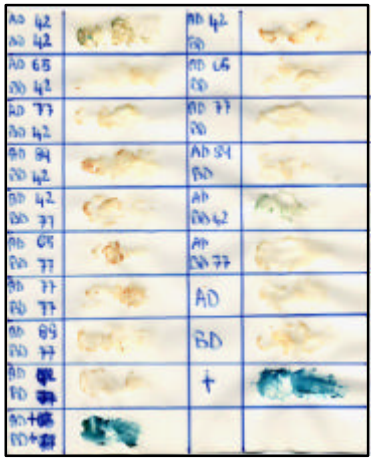
a)



b)

Yeast strain	Vector	-Leu	-Trp	-Ura(*)	β-gal assay (*)
YRG-2				+	-
YRG-2	ADBEIL	+		+	-
YRG-2	ADBGRF	+		+	-
YRG-2	ADBBR	+		+	-
YRG-2	ADBAPL	+		+	-
YRG-2	BDBEIL		+	+	-
YRG-2	BDBGRF		+	+	+
YRG-2	BDBBR		+	+	-
YRG-2	BDBAPL		+	+	+
YRG-2	AD	+		+	-
YRG-2	BD		+	+	-
YRG-2	GAL4ADBD	+		+	+
YRG-2	p53 +		+	+	+

c)



d)

Vector combination	β-gal assay	Vector combination	β-gal assay
ADBEIL BDBEIL	+	ADBEIL BD	-
ADBGRF BDBEIL	-	ADBGRF BD	-
ADBBR BDBEIL	-	ADBBR BD	-
ADBAPL BDBEIL	-	ADBAPL BD	-
ADBEIL BDBBR	-	AD BDBEIL	+
ADBBR BDBBR	-	AD BDBBR	-
ADBAPL BDBBR	-	AD	-
GAL4	+	BD	-
P53	+		

Fig. 3.3.1 Interactions of BEIL, BGRF, BBR and BAPL by means of the two-hybrid assay using the galactosidase reporter gene. a-b) Galactosidase activity in control transformations, all constructs were tested alone as well as with the pAD-GAL4 and the pBD-GAL4 empty vectors. Positive controls included were the GAL4 protein (AD/BD) and the AD of the murine p53 cloned in pBD-GAL4 (+). c-d) Galactosidase activity in tranformation using constructs tested in all possible combinations and against the empty pAD-GAL4 and the pBD-GAL4 vectors. The experiment revealed neither homo- nor hetero-interactions.

The two hybrid experiment can be considered uninformative, based on the results, but also because the sequences used where derived from the one-hybrid cDNA. As revealed from the cDNA screen conducted later, these sequences did not encode the full ORFs: BEIL lacked at least 44 a.a. at the C-terminal end; BGRF, 4 a.a. at the N-terminal and 41 a.a. at the C-terminal end; BAPL, 12 a.a. at the N-terminal end and BBR, 25 a.a. at the N-terminal part of the protein. The experiment is in the process to be repeated using the full-length cDNAs and deletions for the self-interacting clones.

The information already obtained, in any case, were useful to understand that it may be difficult to develop experiments directed to the demonstration that a protein complex exists, mediating the intron IV-type of homeobox regulation.

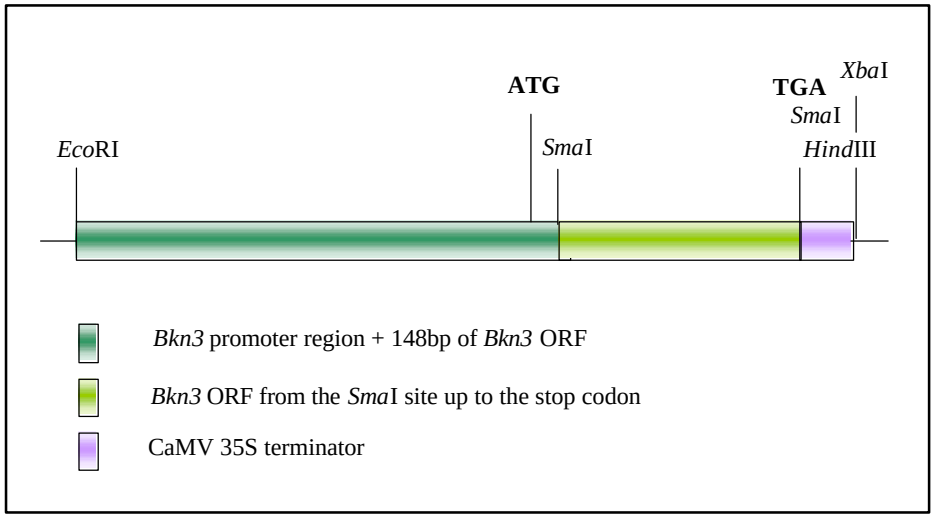
3.4 *Bkn3* promoter-*Bkn3* gene construct in transgenic *N. tabacum*

Transgenic tobacco was frequently used to over express homeobox genes to understand their morphogenetic capacities and their role and function in plant leaf development. Transgenic tobacco plants expressing the barley homeobox genes *BKN1* and *BKN3*, under the control of the full cauliflower mosaic virus (CaMV) 35S promoter (-343 to +8), have been morphologically extensively analysed (Lin *et al.*, 2001). The over expression of *BKN3* showed striking effects compared to wild type which can be summarised as follows: the overall size of the transgenic lines was dramatically reduced, rarely reaching one-fifth (25-40 cm) of the wild type; epiphyllous appendages were found on the adaxial leaf surface, often changing the leaf shape from a bilateral to a radial organisation. Epiphyllous structures were grouped into two categories: epiphyllous, vegetative shoots growing on lower levels of heart-shaped leaves, and inflorescences or isolated flowers on the upper levels of lanceolate leaves, corresponding to the juvenile and mature or adult stage in the wild-type plant, respectively. Every terminal flower of the main inflorescence showed an unusual phenotype having petals fused only at the lower part and displaying a tubular shape at the tip. In addition outgrows were present on the petal margin. These extreme tobacco phenotypes reproduce, in a heterologous system the effect of the *Hooded* mutation in barley, showing a clear phenomenon of epiphilly as in the case of the 305bp element direct duplication in the large intron IV of barley *Bkn3* (Mueller et. al. 1995). Knowing how *Bkn3* can affect development under the control of its own promoter in tobacco helps to further understand the role of this promoter.

3.4.1 Generation of the P*bkn3*::*Bkn3* construct

The promoter region and part of the first exon coding sequence was isolated from the *Bkn3* genomic clone, and the remaining coding sequence was obtained from the *Bkn3* cDNA clone. The 35S poly A signal from the pRT 101 vector was used as a terminator of transcription and pBIN19 was the binary vector chosen for plant transformation. In the first step the 35S terminator was to isolated from pRT101 and subcloned it into pBIN19: the fragment was generated by a *Hind*III-*Xba*I digestion and ligated into the *Hind*III-*Xba*I sites in the polylinker of pBin19. A fragment from the genomic clone of Bkn3 was produced by an *Eco*RI-*Sma*I digestion and subcloned as an *Eco*RI-*Sma*I fragment upstream the 35S terminator, previously cloned into pBIN19. This sequence contains a fragment of 4Kb from the genomic region preceding the ATG start codon, and 148 bp of the first exon of *Bkn3*. A PCR product, using a proof reading termostable DNA polimerase (Pwo DNA polymerase, Roche) and the *Bkn3* cDNA clone as template, was generated using two primers (KS1, KS2), one on the endogenous *Sma*I site, the other just downstream of the stop codon, in this case

including an artificial *Sma*I site. The product was digested with *Sma*I restriction enzyme and ligated into the afore mentioned construct, which was opened as *Sma*I and dephosphorilated using Calf Intestinal Phosphatase (Roche). A clone containing the right orientation of the insert was selected and the complete fusion was sequenced



Schematic representation of the P*bkn3*::*BKn3* construct.

3.4.2 Tobacco transformation and morphological analysis

The p*Bkn3-Bkn*3 construct was moved from *E.coli* to *Agrobacterium tumefaciens* strain LBA 4404 using electroporation. Tobacco leaf disks were infected, and the resulting calli produced and selected on MSI medium with kanamicyn at 100mg/ml, as described in materials and methods. A total of 30 independent transgenic lines were analysed phenotipically. Twenty-six out of the 30 plants regenerated showed a clear mutant phenotype, even if with a different degree of intensity. A plant was classified as having the altered phenotype whenever it showed defined characteristics: reduced plant height, rounded leaves heart shaped with early divergence of the lateral vein from the mid-rib, and, occasionally, epiphilly (Fig. 3.4.1). The 26 transgenic lines were classified into two groups according to their phenotypes, representing two different degrees of phenotypic modification.

The two groups were compared with the 35S-*Bkn3* transgenic tobacco plants. The transgenic plants expressing the same gene under the control of the two different promoters, were phenotypically

very similar. Also, the 35S-*Bkn3* plants could be, in fact, divided into different classes but, as a rule the effect of *Bkn3* under its own promoter was milder than under the viral promoter: the reduction of plant height was never less than ½of the wild type; epiphilly was clearly present with most shoots developing from the mid-rib of the leaf, particularly in the group of plants with the more extreme phenotypes, but never so intense as the "35S plants", which, in some cases showed up to 30 epiphillous shoots close to the junction between petiol and lamina of the heart shaped leaves. Epiphyllous shoots close to the leaf tip, sporadically found in the 35S-*Bkn3* transgenics, was never observed in the p*Bkn3*-*Bkn3* plants. Additionally, epiphilly observed on lanceolate leaves as in the 35S-*Bkn3* subtending the branches of inflorescences or the appearance of isolated flowers on the same leaves was never observed. This was also the case of the classical *Bkn3* over expressing flower phenotype: tubular shaped petals just fused at the base with appendages as outgrows at the margin.

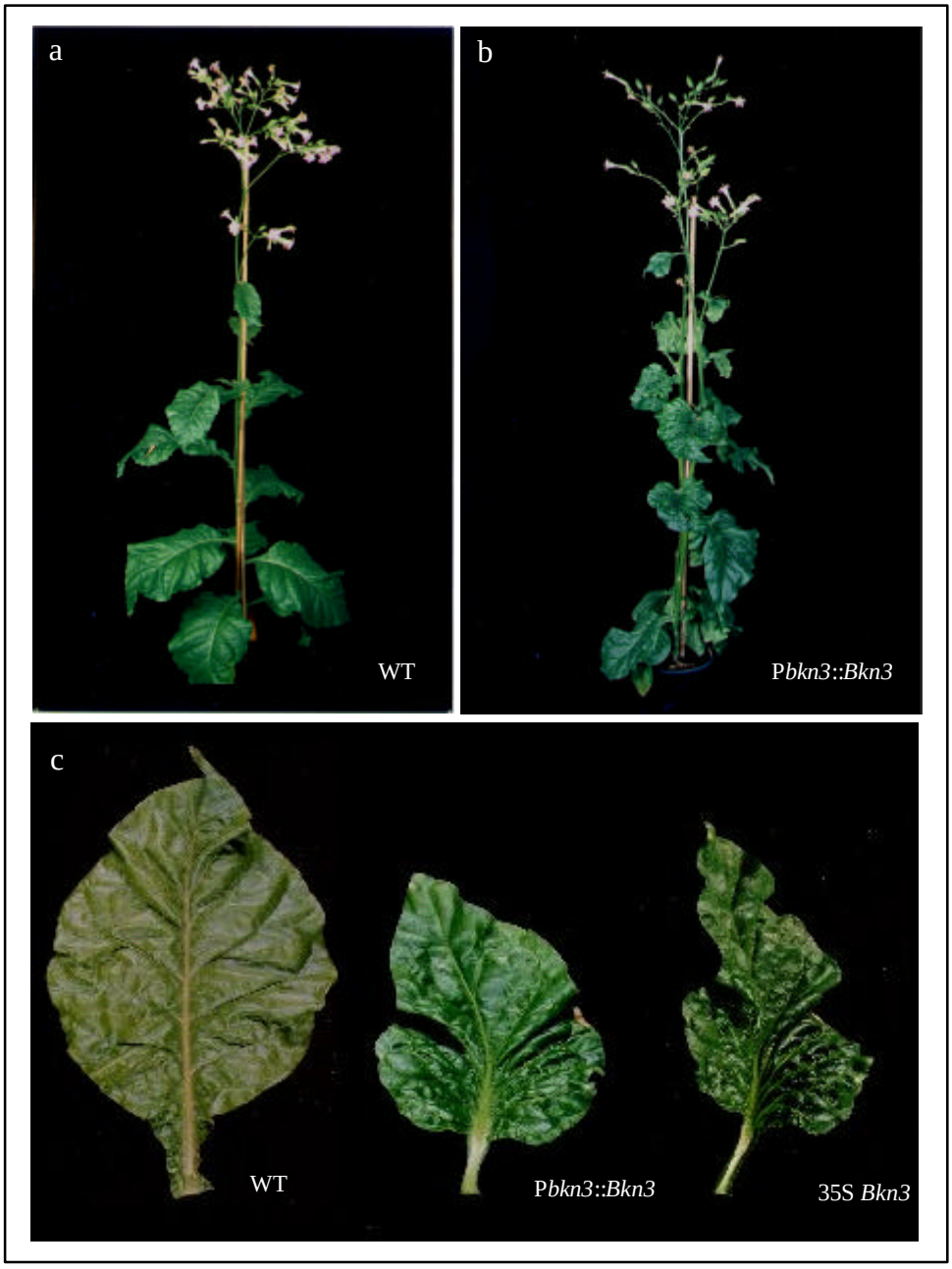


Fig. 3.4.1 Comparison of the gross plant morphology of WT SR1 *N. tabacum* (a), with a P*bkn3*::*Bkn3* transgenic line (b) and of the leaf morphology of WT with 35S BBR and p*Bkn3::Bkn3* transgenic lines (c).

3.5 Protein expression

For the four sequences isolated by the one-hybrid screen due to their ability to bind in yeast the 305bp element, the binding capacity in an *in vitro* protein-DNA interaction assay was tested. In yeast, transcription is in fact driven by an heterologous complex and the protein identified may not have a real DNA binding properties. The first step was to obtain relatively pure proteins to be used in an electro-mobility-shift-assay.

3.5.1 in vitro

The proteins were synthesised performing transcription followed by translation *in vitro*, utilising the two most commonly used systems: the wheat germ extracts (Promega) and the rabbit reticulocyte lysate system (Amersham). All sequences had to be moved from the pAD vectors to pBlu KS, which carries the T3 and T7 RNA polymerase promoters, necessary for *in vitro* transcription, at either side of the polylinker. The coding sequence of the four genes were isolated with a *EcoRI-PstI* digestion and cloned in the *Eco*RI-*Pst*I restriction sites of the multiple cloning site of pBlu KS. According to the polylinker sequence and to the direction of the insertions of the four fragments the T3 promoter was chosen for the *in vitro* transcription of the sense RNAs. These vectors were linearised at the 3'end: BEIL and BBR with *BamHI* and BGRF and BAPL with *SmaI* restriction enzyme. Around 2µg of each linearised plasmid were used as template for the T3 RNA polymerase as described in materials and methods (2.2.9). The products of the reactions, after a DnaseI treatment, were checked on a normal agarose gel to have a rough indication of the reaction efficiency and to decide the amount to be used in the translation step.

As mentioned before, the wheat germ and the rabbit reticulocyte extracts were used for each clone. Both extracts are considered highly efficient cell free systems for synthesis of proteins from exogenous RNA templates. The two types of reactions were carried out according to the kit instructions, in duplicate, including in one case hot [^{35}S], in the other cold methionine. This was done in order to check the reactions with the radioactive amino acids and to use the cold preparation directly in a band shift experiment with the whole 305bp element as DNA target sequence.

The reactions were conducted in parallel for 2h at 25°C for the wheat germ and for 1h at 30°C in the case of the reticulocytes. An aliquot of the radioactive reactions was loaded on an polyacrylamide mini gel and the gel was treated and exposed to autoradiography film in order to detect the [^{35}S] labelled proteins. In both systems the proteins of the expected molecular weight were produced, even if for BAPL more than one clear band were present, probably due to an alternative starting point and, or to degradation of the protein.

The eight preparations were tested at three different concentrations in a gel shift assay (see materials and methods and the next sessions of results for details). None gave a clear shift, except BBR, which, when translated with the reticulocytes extract, produced a small shift of the sequence, even if weak and fuzzy (Fig. 3.5.1).

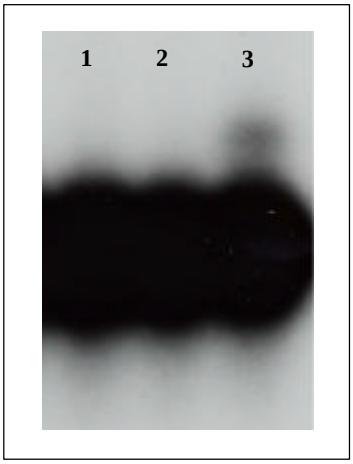


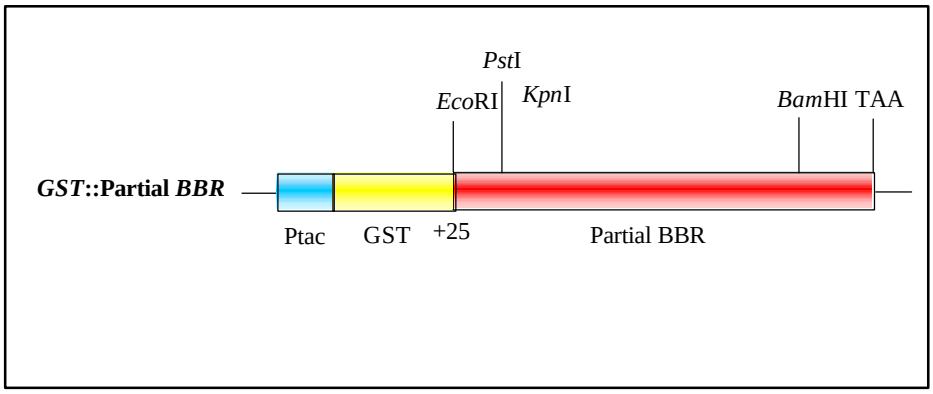
Fig. 3.5.1 Initial EMSA experiment testing the interaction of the 305bp fragment with *in vitro* translated BBR, synthesized using the reticulocyte system (Amersham). In lane 1 the 305bp free probe alone and together in lane 2, with the reticulocyte extract control and in lanes 3 with 4µl of the translated BBR protein product.

3.5.2 In *E. coli* (In collaboration with Yamei Wang)

When using the *in vitro* synthesised protein, even if the major product and component at the end of the reaction is the protein of interest, all other proteins present in the extract, necessary for the translation, are also present and they can affect protein-DNA binding; following the weak indication of binding shown by BBR, it was decided to obtain, at least for BBR, a protein of higher purity. BBR was overexpressed in *E.coli*, under an inducible promoter, fused with a sequence tag to enable protein purification utilising antibodies available against the tag, or other specific chemical properties of the tag itself. The pGex-5X was the vector (Pharmacia), which contains the glutathioneS-transferase (GST) at 5'end of the multiple cloning site to produce fusions that carry the GST tag at the N terminal part of the protein.

The glutathoine-S-transferase is an enzyme that is concerned with the detoxification of several substances such as xenobiotics. It is a small enzyme that contains 49 groups and the only two α-helices are made up of a total of 24 bases. The tac promoter with the lac operator downstream allow induction and modulation of the chimeric protein. The ribosomal binding site of GST, at the 3'of the ATG start codon, increases the translation level. Apart from the origin of replication and

the resistance marker, the *lac I*^{q} gene is also present because necessary to induce gene expression with isopropyl-β-D-thiogalactoside (IPTG).



Schematic representation of the *GST*::*BBR* contruct.

The partial BBR (Δ25 a.a. at the N-terminal end) was subcloned as a *Eco*RI fragment from the one-hybrid pAD-GAL4 vector. The construct was transferred from the *E.coli* strain DH10B, used to maintain and clone the plasmid to *E.coli* strain BL21, which is a good host for the expression of recombinant proteins. Like all the *E.coli B* strains, it lacks the major protease, encoded by the *lon* gene, which catalyses the endoproteolytic cleavage of damaged and recombinant proteins in the cell. Five transformed colonies carrying the GST-BBR fusion and one colony carrying the pGex-5X empty vector, used as a control, were grown over night; 500 µl of the overnight culture were used to inoculate 4.5 ml of fresh LB medium and grown for 2 h. The cultures were split into two sets of cultures of 2.5 ml each. The first was induced with IPTG 1mM final concentration and grown for 4 h, while the second was grown for 3 h at 37°C without the inductor. A 1.5 ml aliquot of the culture was centrifuged, loading buffer added to the pellet, denatured at 95°C for 5 minutes and 15 µl of the supernatant loaded on an acrylamide mini gel. Comparing induced and non induced lysates carrying the empty vector, as well as the fusion construct, it became apparent that after induction at least three of the five colonies had a dramatic increase, of a band of roughly 74 kD, that is the expected size of BBR when fused to the 26 kD of the GST.

The three colonies were inoculated, grown, induced and lysed as described in material and methods. At first, a batch purification using glutathione sepharose beads was carried out without success, testing aliquots of the different steps of the purification procedure: full lysate, matrix, first wash, second wash, first elute and second elute. However the protein was present but still bound to the

matrix and was not properly eluted. It was decided to combine the batch purification with the column method. A normal protocol for batch purification was followed until the binding of the glutathione sepharose beads, then the matrix was loaded on a column at 4°C until sedimentation. Two washing steps were performed and the elution buffer (containing 15 mM of reduced glutathion and three proteinase inhibitors: benzamidin, caproic acid and PMSF) was applied and left equilibrating overnight at 4°C. The next day two successive elution steps were performed (see materials and methods 2.2.12). Using this combined method the purification was successful; the only band present had the expected molecular weight of the chimerical GST-BBR protein with no degradation products visible (Fig. 3.5.2). Subsequently, the 1 ml of elute was dialysed overnight at 4°C against 5 L of the 2X G1 binding buffer used for the gel retardation assay.

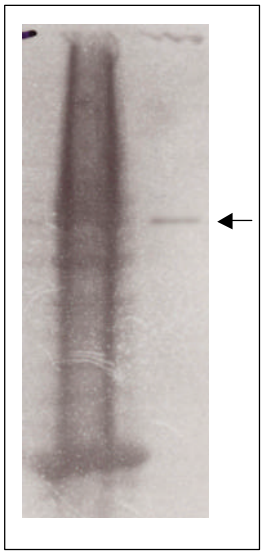


Fig. 3.5.2 Purified *BBR::GST* (74Kd) product (arrow) after a combined batch and column purification.

3.6 *In vitro* BBR-DNA interaction

The gel mobility shift assay is a simple and sensitive method for determining interactions between protein and DNA. This assay was developed by Fried and Crothers (1981) and Garner and Revsin (1981) for analysing protein-DNA interactions, it separates protein-DNA complexes from free DNA by non-denaturing polyacrylamide gel electrophoresis.

3.6.1 305bp element retardation

As a target sequence, the whole 305bp element was chosen for the first experiments. The 305bp fragment contains three unique restriction sites within its sequence: *Xho*I, *Bgl*II and *Nhe*I. The *Xho*I and *Nhe*I fragments generated two different probes, which were tested independently afterwards. This was done to avoid the putative binding site being cut by the restriction enzyme used to produce the probe, thus abolishing the binding. They were isolated from construct 305-GUS-3X (3.1.3) which contains the 305bp repetition in triplicate. The 3'-end of the DNA was labelled using the 5'-3' polymerase activity of the Klenow fragment of DNA polymerase I. The recessed 3' end, produced by the *Xho*I and *Nhe*I restriction, was filled using the corresponding 5' overhangs as a template. The 5' overhangs were composed by G, A, T, C in the case of *Nhe*I and by A, G, C, T in the case of *Xho*I. Both reactions included cold nucleotides at a concentration of 5mM each and α-^{32}P labelled nucleotide: dCTP for the *Nhe*I fragment and dGTP for the *Xho*I fragment. A wide variety of parameters can influence protein-DNA interactions and affect the results of the mobility shift assay. For example, monovalent (Na^+ and K^+) and divalent ($Mg2^+$) cations, non ionic detergents, binding temperature, binding time and protein concentration. In addition, the gel composition as well as the electrophoretic conditions can significantly alter the mobility of a given protein-DNA complex. The binding buffers were prepared according to different available literature sources, eight different binding buffers were produced (see material and methods 2.2.13) and tested at the beginning with the in vitro produced proteins. Since the shift indication given by BBR in the *in vitro* transcription and translation took place using the G1 buffer, all the other experiments used the G1 as the binding buffer of choice.

In general it is reported that low salt conditions tends to favour protein-DNA interactions by decreasing the dissociation rate of the complex. Conversely, higher salt concentration facilitates dissociation.

Addition of the non-specific competitor DNA to the binding reaction is important when one is trying to select single DNA binding proteins from a vast number of proteins such as in crude extracts. The most commonly used DNA competitors are synthetic copolymers such as poly(dI-dC). These copolymers provides a vast excess of low affinity binding sites which absorb DNA binding proteins non-specifically, thus allowing detection of specific protein-DNA complexes. In these experiments, dI-dC was always included in the binding reactions as a source of unspecific cold DNA competitor. The binding reactions took place during twenty minutes at room temperature using different amounts of proteins and probe DNA. A run of eight hours on a 4% non denaturing

polyacrylamide gel at 4°C clearly confirmed the shift found previously using BBR, with both the 305bp elements used to produce the probe; the *Xho*I and *Nhe*I fragments.
Subsequently in order to confirm the shift and to test the specificity of the binding, several controls were performed. Since the BBR used was a fusion with GST, free GST was used alone with the probe DNA. In addition, optimised increased levels of proteins as well as increasing concentrations of unlabeled cold specific DNA competitor were used. Cold competitor at high concentration was able to abolish the complex observed from the trace amount of DNA in the reaction (Fig. 3.6.1).
Since the shift was specific and completely out competed by a 200 fold excess of unlabelled DNA probe and it was consistent and reproducible, the binding properties of BBR could be used to obtain more information on the DNA binding site within the 305bp.

3.6.2 Mapping of the DNA binding site (In collaboration with Yamei Wang)

Several deletions were produced. At the beginning, three overlapping fragments of at least 20bp were produced by PCR, each pair of primers (305a, 305b; 305c, 305d; 305e, 305f) contained the artificial restriction sites *Eco*RI and *Xba*I to allow cloning into pBlu KS where the fragments were inserted in order to be kept as a source of template. The fragment containing the last 125bp (180-305) was the only one to be retarded by BBR and it was subsequently divided into four shorter fragments, with in this case 10bp overlaps. From this point on, since the DNA was shorter than 50bp, synthetic oligonucleotides were used. Each of the four fragments was made of two complementary annealed oligonucleotides (401, 402; 403, 404; 405, 406; 407, 408) and lacking the 5'-phosphate they were labelled using the T4 polynucleotide kinase and γ^{32}P ATP. Again only one fragment (238-276) showed exclusive binding, defining just 39bp of binding sequence. The fragment was shortened this time to a 26bp sequence (243-268; oligonucleotides 261, 262) and at the end the use of two pairs of oligonucleotides (171, 172; 161, 162) TCTCTGGTCTCTCTCTC and TCTCTCTCTCTCTCTC each one annealed together with his complementary sequence, undoubtedly allowed the precise identification of the binding site on the DNA.

BBR was able to bind and shift specifically and exclusively the last of the two double stranded DNA sequences, with respect to the control experiments (Fig. 3.6.1).

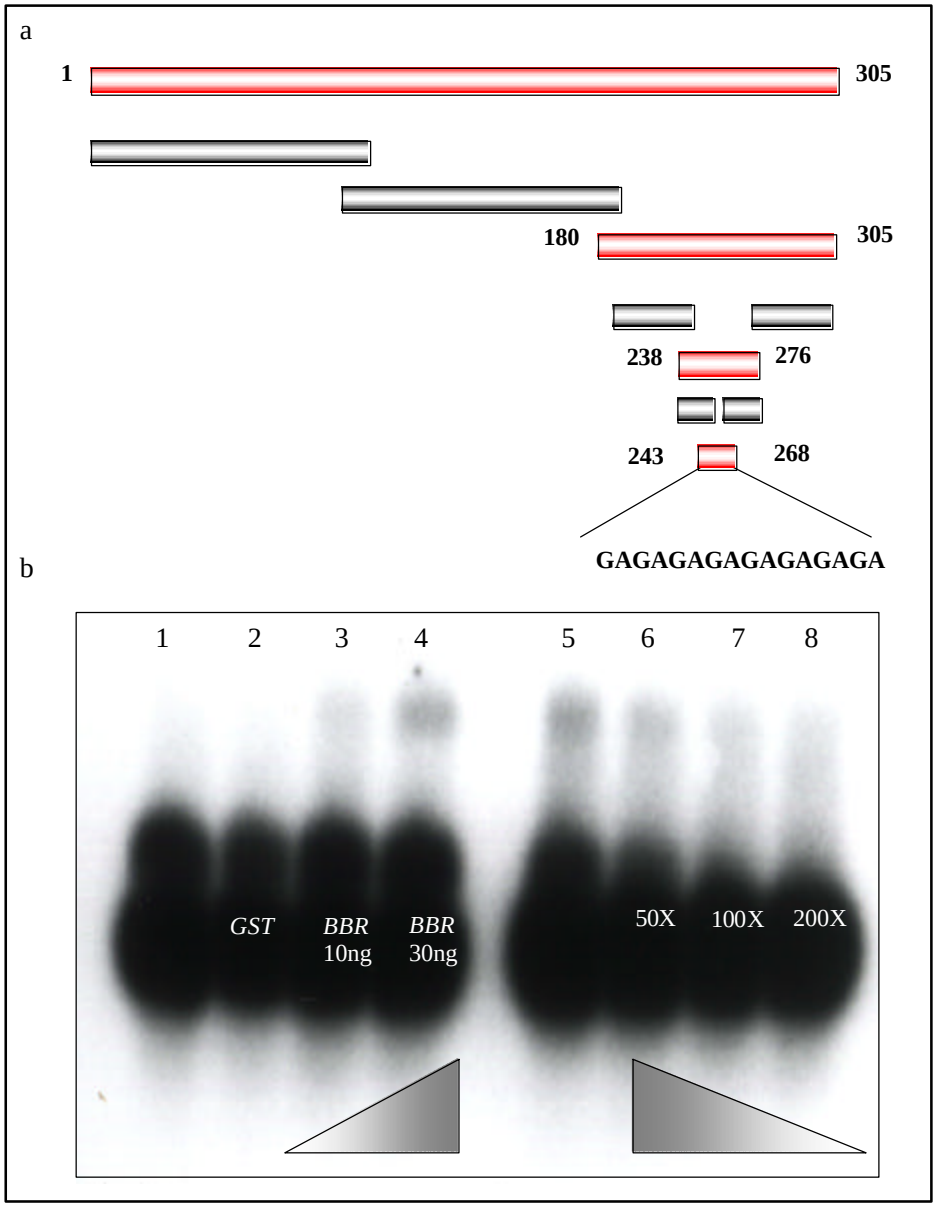


Fig. 3.6.1 a) The mapping strategy to identify the shortest target sequence able to bind BBR is summarised. The BBR DNA binding site is defined by 16bp of dsDNA composed by a $(TC)_8$ repeat together with a $(GA)_8$ repeat on the complementary strand. b) EMSA. The 305bp free probe alone (lane 1) and together with 500 ng of the purified GST control (lane 2) or in lanes 3 and 4, 10 and 30ng respectively of purified *BBR*::*GST*. In lanes 5, 6, 7 and 8 the competition experiment is presented. All the lanes contain 30ng of *BBR*::*GST* together with the 305bp probe. In addition, lanes 6, 7 and 8, respectively contain a 50, 100 and 200 molar excess of cold probe

3.7 Overexpression of different constructs of BBR in transgenic Tobacco plants

Following the same strategy adopted for *Bkn-1* and *Bkn-3*, also *BBR* was over expressed using *Nicotiana tabacum* as heterologous system. At the beginning, the partial BBR was used, -25 amino acids at C-terminal, alone and in fusion with the GAL4 yeast activation domain. This second construct was functional in activating the reporter in the yeast one hybrid system. Later, when the full sequence was available, the whole coding region was used to create transformants.

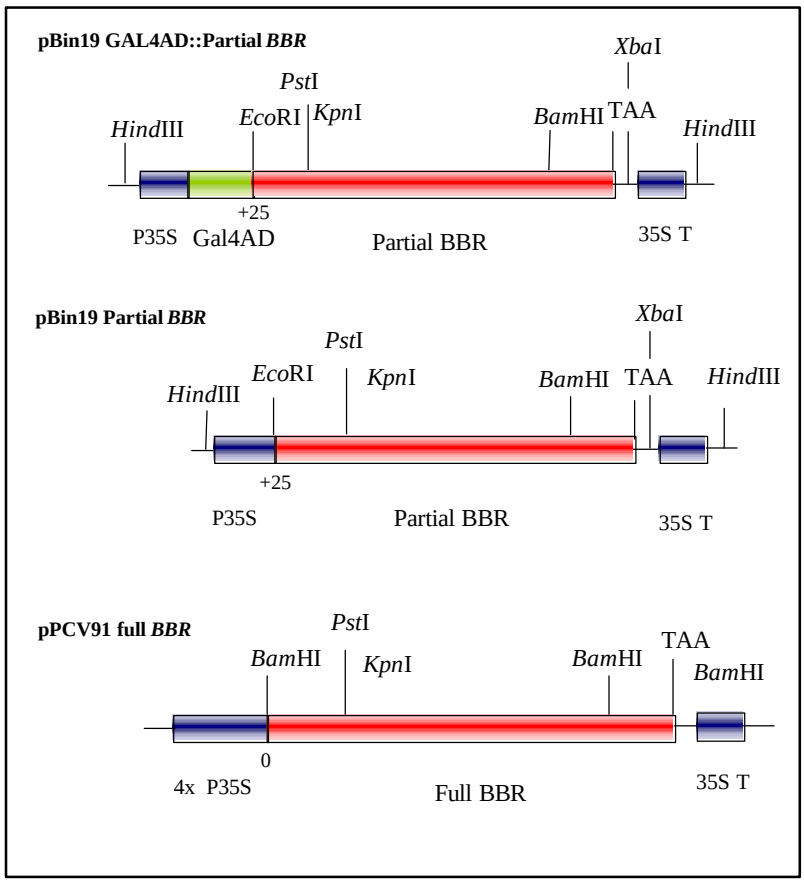
3.7.1 Constructs

The partial BBR sequence was obtained directly from the pAD constructs, with and without Gal4 AD. Since pBIN19 was the binary vector for plant transformation, the sequence was transferred in pRT101 in between the 35S promoter and the 35S polyadenylation signal (Toepfer *et al.*, 1987). Then promoter, coding sequence and terminator were moved in the pBIN19 vector.

The partial BBR was excised as an *Eco*RI/*Xba*I fragment and directly cloned in pRT101 using the same restriction sites. From pRT101 the cassette was moved to pBIN19 undirectionally using the *Hind*III sites and the right orientation was selected. For the construct containing the Gal4 AD a blunt/sticky cloning strategy was followed: the pAD-GAL4 vector containing BBR was opened with *Hind*III and filled in using the 5'→3' polimerase activity of the Klenow subunit to create a blunt end, then digested with *Xba*I and the fragment purified from the gel. The pRT101 vector was digested with *Eco*RI restriction enzyme, this site was filled in, and then the vector digested with *Xba*I. The blunt and *Xba*I sticky ends of insert and vector were ligated, subsequently the all cassette was moved in pBIN19 using *Hind*III restriction enzyme similarly to the previous construct.

When the cDNA λ screen was performed and the full sequence was available, also the full coding region of BBR was used to produce Tobacco transgenic plants. The pPCV91 was the vector of choice, which contains already a CaMV 35S promoter with four CaMV 35S enhancers upstream of it, and a CaMV 35S terminator. The sequence was directly cloned in the binary vector without additional subcloning in any other expression vector.

The insert was generated using the cDNA sequence as template for a PCR reaction with two primers that both contain an artificial *Bam*HI sites (FBBR1, FBBR2). The first step was to clone this fragment undirectionally in the unique *Bam*HI site of pBlu KS. The new construct was sequenced. From pBLU KS the insert was transferred in pPCV91 using the same *Bam*HI site, both orientations, sense and antisense, were selected from this unidirectional cloning.



Schematic representation of the different BBR contructs for plant overexpression.

The two pBIN19 derived constructs and the two pPCV91 derived constructs were transferred in two different *Agrobacterium tumefaciens* strains suitable for them: LBA 4404 and GV 3101 PMP90RK respectively. While LBA 4404 carries streptomycin and rifampicillin as strain specific markers and kanamycin resistance from pBIN19, GV 3101 PMP90RK carries rifampicillin and kanamycin and pPCV91 confers resistance to Carbenicillin.

3.7.2 Plants

Tobacco leaf disks were infected with *Agrobacterium* harbouring the two different constructs, formation of calli was induced and they were selected on MSI medium with Kanamycin 100mg/ml

as described in materials and methods (2.2.15; Murashige and Skoog, 1962). Ten different resistant calli for each kind of constructs were selected and grown to maturity. The mature plants were tested by PCR with two different pairs of primers, one gene specific and the other from the selective marker, the kanamycin resistance gene (77LU2, 77LU5; KA1, KA2).Seven on the ten plants carrying the Gal4AD-partial-BBR constructs and six on ten carrying the partial-BBR construct were positive for both combinations, hence they were chosen and their seeds harvested. Already at the T_0 generation, five on the seven transformed tobacco plants carrying the partial BBR with the Gal4 activation domain showed a clear leaf phenotype, which will be described later. None of the six selected plants carrying the BBR partial constructs showed any evident phenotype in this generation. The progeny of the chosen plants was selected again on antibiotic and grown to maturity. All the Gal4-partial-BBR transformed plants presenting the leaf phenotype confirmed the previous finding in this next generation. The progeny of the other 2 independent Gal4-partial-BBR transformants and all of the six partial-BBR confirmed absence of phenotype in this T_1 generation.

Three T_1 plants, each one coming from an independently transformed line, were selected and analysed for BBR over expression. From the Gal4-partial-BBR the three plants were chosen among the ones showing the leaf phenotype: AD-Pa BBR1, 2, 3. For what concern the partial-BBR construct they were randomly chosen: Pa-BBR1, 2, 3. Leaf material was collected and a total RNA extraction was performed as described in material and methods (2.2.3). The total RNA was used in a Northern blot hybridisation experiment. The probe was synthesised using the partial BBR sequence as template for random priming labelling and used for both kind of constructs. After a week time exposure, the film showed clearly that all the six selected plants were over expressing the transgenes. A difference in mobility due to the presence/absence of the yeast GAL4 AD was clear (Fig. 3.7.1). Subsequently, the seeds of these plants were harvested and the progeny selected again.

The T_2 GAL4-partial-BBR plants deriving from the three independent over expressing lines : AD-Pa BBR1, 2, 3 confirmed, for the third time, the leaf phenotype and several of them were kept for morphological analysis. Six for AD-Partial BBR1 (AD 1.1 to AD 1.6) five for AD-Partial BBR2 (AD 2.1 to AD 2.5) and eight for AD-Partial BBR3 (AD 3.1 to AD 3.8). None of the plant over expressing the partial-BBR construct showed a phenotype deviating from WT. They were not analysed any further.

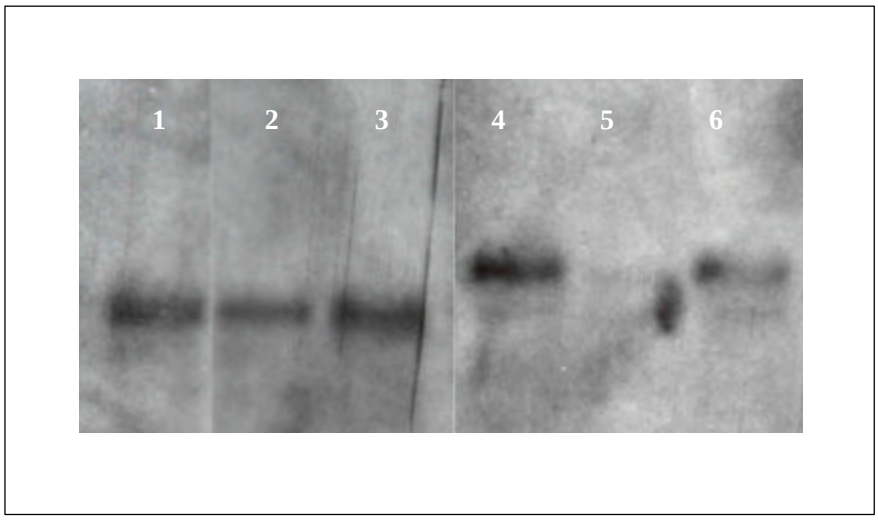


Fig. 3.7.1 RNA gel blot analysis of the three partial BBR overexpressing plant lines: Pa-BBR 1, 2, 3 lane 1, 2, 3 respectively and of the three GAL4-partial BBR plants: ADPa-BBR 1, 2, 3 in lane 4, 5 and 6. BBR cDNA was used to generate the probe for the hybridisation experiment.

3.7.3 Leaf phenotype

The phenotype showed by the GAL4-partial-BBR plants was consistent and reproducible over the three generations analysed. Basically, this transgenic tobacco plants carried narrow and elongated leaves compared to the wild type SR1 tobacco. This was clear since the very early stages of development: the plantlets presented elongated cotyledons rolled at the lateral margin and with a curved shape. The narrow and elongated characteristic of the leaves persisted during later stages of development, even if the concave shape present at the cotyledons stage was not present. Almost all the leaves of the entire plant had this feature (Fig. 3.7.2).

Characteristics correlated to the vegetative phenotype were also found in the reproductive part of the plant: sepals were elongated compared to the wild type SR1 and also the petals were longer than WT keeping the pistil and the stamen almost inside the tubular shape of the tobacco flower (Fig. 3.7.2).

Several plants were derived from three independently transformed T_1 plants: Six for AD-Pa BBR1, five for AD-Pa BBR2 and eight for AD-Pa BBR3 were analysed in comparison to two SR1 wild types to quantify phenotypic differences. The average length and width, calculated from the

19 selected AD-Pa BBR progeny, of two different leaves of the same plant, the forth and fifth developed from the base in each mature plant, were measured and compared to the average value obtained from two leaves at the same developmental stage and position of two SRI tobacco plants.

Fig. 3.7.2 Comparison between a) WT and b) ADPa-BBR plants. In transgenic tobacco plants harbouring the GAL4AD::partial BBR construct all leaves presented the characteristic elongated morphology.

Fig. 3.7.3 Comparison of a) flowers b) capsules and leaves of SRI WT tobacco plants to a GAL4AD-partial BBR fusion overexpressing plant. Leaves, sepals and petals showed a clear elongation in the BBR overexpressing lines.

The length of sepals for each plant was measured and compared with those of wild type (Fig. 3.7.3). The increase in leaf length and the decrease in leaf width were both involved in the establishment of the altered leaf phenotype, however the second parameter showed a more pronounced difference. Petals were elongated of one third in average compared to wild type. Individual measured values, averages and standard deviations are reported in appendix.

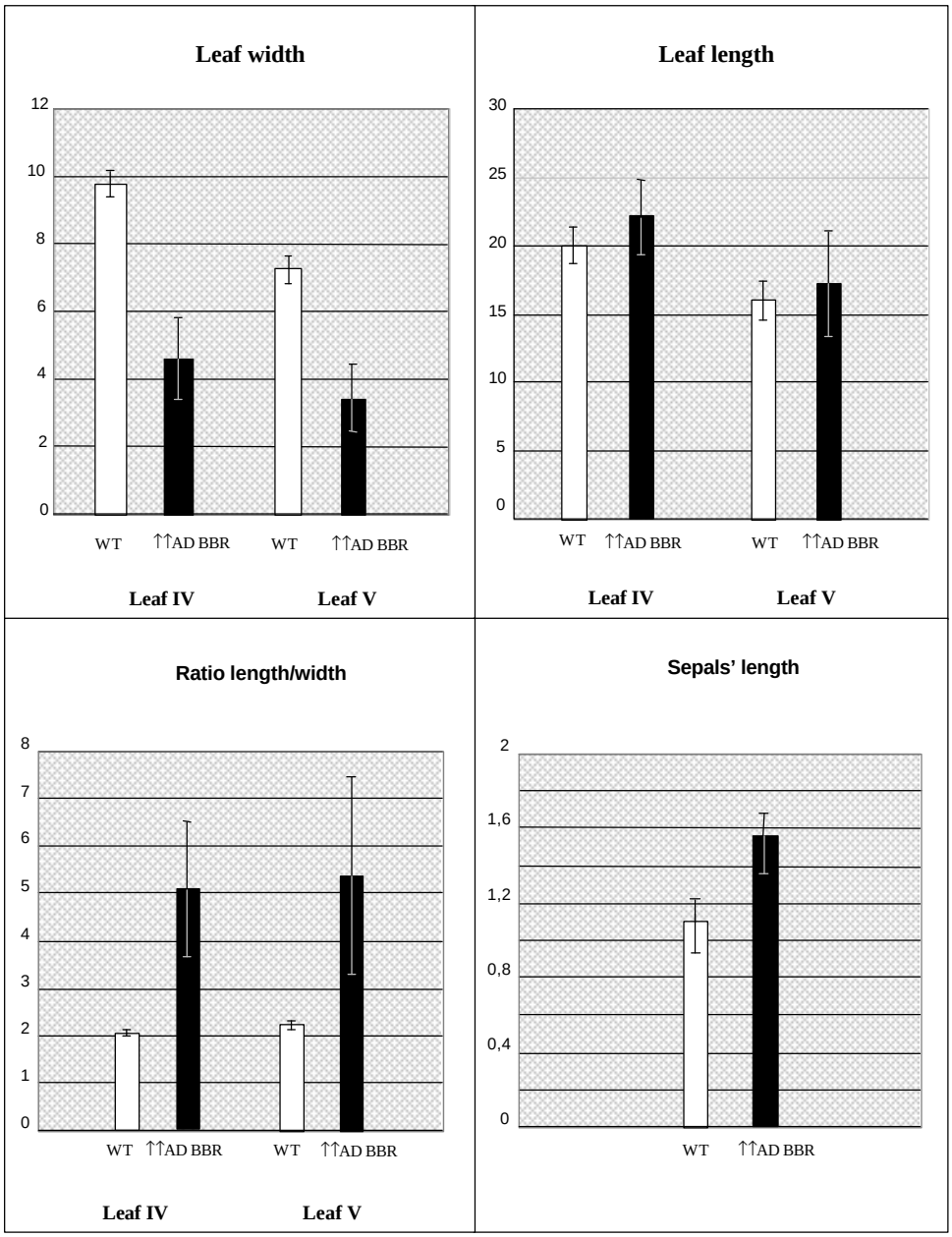


Fig. 3.7.3 Comparison of the average length (a),width (b) and ratio length-width (c) of two different leaves, the forth and fifth developed, of 19 GAL4AD::partial BBR overexpressing lines to 2 WT plants. In (d) the comparison of the sepals' length is shown. On the Y-axis values are expressed in cm.

3.7.4 Full length BBR overexpressing plants

As previously described, these constructs were assembled in the pPCV91 vector that carries hygromycin as plant selectable marker. Both sense and antisense constructs were transformed by leaf disks *Agrobacterium* mediated transformation. The efficiency was very low and few plants could be grown from the infected explants: four for the sense construct and three for the antisense one. During *in vitro* culture the majority of the calli, assumed a glass like appearance without a proper rooting of the rare shoots. Three of the four sense-BBR T_0 plants showed the modification of the phenotype, obtained for the GAL4AD::partial BBR even if less dramatic. All antisense plants had a normal WT phenotype. The three positive plants carried the construct, as revealed by PRC analysis conducted with a gene specific and a hygromycin specific pair of primers (77LU2, 77LU5; hph1, hph2). Seeds will be harvested and T_1 generation analysed.

3.8 Transient expression in tobacco protoplasts

Two different proves of BBR binding to the 305bp element have been until now achieved: in yeast and i*n vitro*. To give an additional evidence *in vivo,* and to help to understand the biological function of BBR, protoplasts from tobacco leaves were prepared and used to test how the expression of a reporter GUS (*uidA* gene) construct, carrying the 305bp element in the regulatory region, was affected by BBR, as if this gene acts as a putative transcriptional regulator.

The constructs prepared were of two kinds: effectors and reporter. Effectors constructs contained the full length BBR under the control of the full CaMV 35S promoter, in one case BBR alone and in the other BBR cloned as fusion with the VP16 activation domain. Several studies have shown that heterogeneous domains, not derived from plants, may act as functional plant activators. The herpes simplex virus protein VP16 is a powerful activator of transcription and its C-terminal activation domain has been used in transactivation studies, targeted via DNA-binding proteins (Wilde *et al.*, 1994).

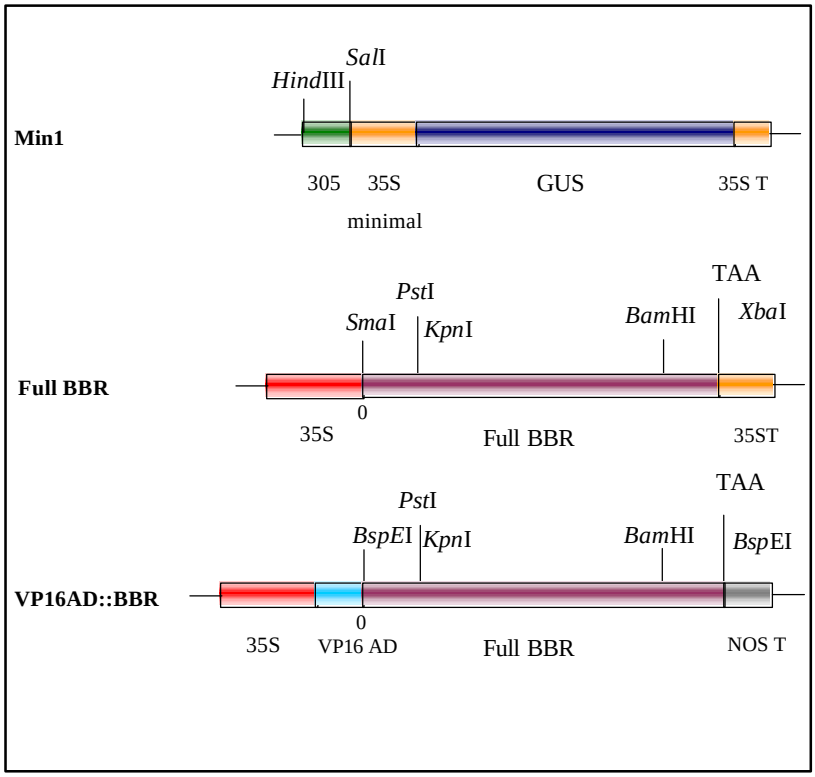
Our reporter construct was derived from K373 (Thompson, 2.1.4). It carries the CaMV 35S minimal promoter, which consists in the CaMV 35S TATA sequence (-46 to+9) in front of the *uidA* gene. The TATA sequence gives the access to generic transcription factors and RNA polymerases, but does not contain enhancer sequences. Because of this it drives very low expression level of the reporter. The 305bp element was cloned in front of the minimal promoter of K373; this construct is indicated Min1.

Constructs were designed to check the hypothesis that BBR could mediate transcription. This was possible considering the differences of the expression level of protoplasts transfected with the available reporter with and without the effector contructs.

The level of GUS activity was first normalised based on the protein content, evaluated for every sample. Later, an additional construct, which has the luciferase as reporter gene, was included in every transformation as internal control. The vector containing the luciferase is ms129-pbt8-ubi-lucm3 (Weisshaar, 2.1.4) that carries the ubiquitin promoter, which, for this porpoise, is usually preferred to the CaMV 35S full promoter that drives high levels of expression and could affect the normalisation procedure.

3.8.1 Constructs

The Min1 reporter construct was derived from K373 and the 305bp element was obtained by PCR from the construct Intron IV GUS 1X (3.1.2). Two artificial restriction sites were introduced, *Hind*III and *Sal*I (305H, 305SA1) and used to clone in the same sites of the K373 upstream to the CaMV 35S minimal promoter. In order to assemble the first of the two effector constructs, the full length BBR was generated by digestion with *Sma*I and *Xba*I of the full length BBR previously cloned in pBLU KS using the *Bam*HI site (3.7.2). This fragment was cloned, using the same restriction sites in pRT101 (2.1.4). The VP16AD-BBR full length is the second effector. In this case the vector used was pBT4 (Weisshaar B., 2.1.4). The pBT4 carries the CaMV 35S promoter, the activation domain of VP16 and the NOS terminator. In this case, a full length BBR, obtained by PCR, was cloned in frame at 3' end of the VP16AD in the unique *Bsp*EI site. The *Bsp*EI sites were artificially introduced in the primers used for the PCR reaction (BBRVP1, BBRVP2). A clone with the right orientation was selected and sequenced.



Schematic representation of the reporter and effectors constructs used in the protoplasts transient transformation experiment.

3.8.2 Protoplasts transfection

Protoplasts were isolated from *Nicotiana tabacum* SR1 leaves of plants 4/6 weeks old. The leaves were treated with macerozime (SERVA) and cellulase (SIGMA) together for 16/20h. This preparation was filtered and protoplasts separated from cellular debris. Subsequently the protoplasts were washed and transfected with 5µg of each construct and 50µg of carrier salmon sperm DNA (materials and methods 2.2.14).

In the first set of experiments the GUS activity was normalised on the total proteins.

The effector constructs were transformed alone and together with K373, which does not contain any BBR binding site. The reporters were also co-transfected with the empty pRT101 vector. These tests didn't reveal any significant modification of the reporter activity allowing proceeding in testing the effect of BBR. The constructs combinations used were: K373 alone (CaMV 35S

minimal), Min1 alone (305 + CaMV 35S minimal), BBR full length together with Min1 and VP16AD-BBR full length with Min1. After the transfection, the protoplasts were incubated for 48h, collected and lysed by freezing in liquid nitrogen. Aliquots of the supernatant were used to determine GUS activity and the protein content. GUS activity was measured at three time points: 30', 1h and 2h by fluorometric assay using the 4-methylumbelliferyl β-D glucorinide (4-MUG) as substrate. Proteins content was measured using the Bradford (1976) method.

To overcome the intrinsic variability of transient expression in protoplasts, the experiments were repeated three times and four independent transfections of the same construct combinations were performed for every experiment. The values of the GUS activity were corrected based on the total proteins amount. The four independent measurements generated average values for each time point, each expressed as folds of activation based on the K373 values calculated for the same time point. Subsequently these ratios of the normalised mediated values of each time point were mediated as well. The graphic (Fig. 3.8.1) shows the normalised mean of three experiments, consisting of four observations each, calculated for three time points and referred to K373 as activity unit. In these experiments BBR shows to activate the level the GUS reporter gene expression of, on average, 5.7 folds compared to the Min1 alone, revealing a specific binding to the 305bp target site also *in planta* and to function, at least in transient assay, as a transcriptional activator.

The same level of induction was achieved when VP16AD-BBR effector construct was used, suggesting that the full length BBR sequence contains an own activation domain functional in plants.

Also a three folds difference in the expression level between K373 and Min 1 reporter constructs, without the participation of effectors, was detected. These constructs differ just for the presence of the 305bp element upstream the CaMV 35S minimal promoter. The presence in tobacco of activator factors able to bind the 305bp element could be an explanation for this result (Fig 3.8.1). Individual GUS activity values normalised to the total protein amount of the respective sample for each time point are reported in appendix.

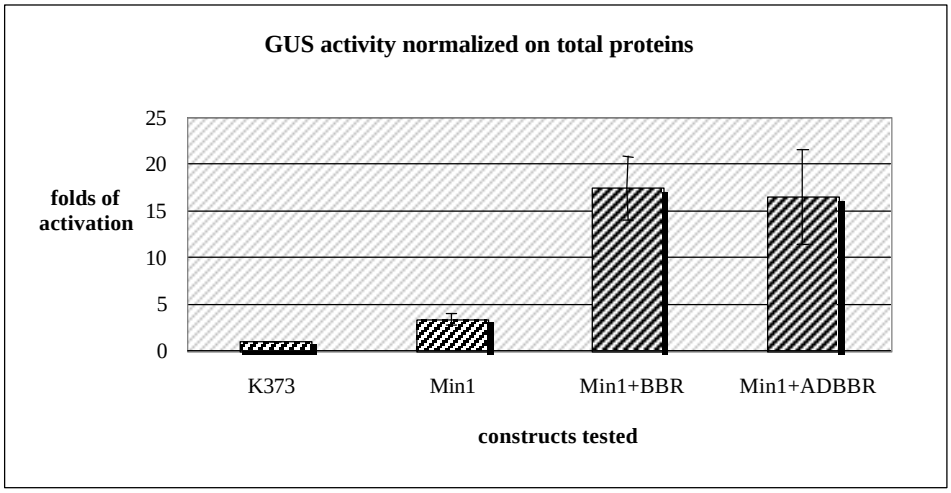


Fig. 3.8.1 Transcriptional activation of the GUS reporter gene under the control of the 305bp element and the CaMV 35S minimal promoter mediated by BBR and VP16AD::BBR. On the Y-axis the folds of activations, calculated normalising on the total proteins and referring to the activity of K373, are reported.

In the second set of experiments just the BBR full length constructs and the Min1 were used to confirm the previous data, in this case always together with ms129-pbt8-ubi-lucm3 as internal control as previously described. The use of an internal control permits to monitor the efficiency of protoplasts transformation and to normalise the reporter activity just on the population of transformed protoplasts. For this reason it is possible to compare transformation experiments coming from different protoplasts preparations and to select just the data that show a high luciferase activity. Following this strategy 10 independent transformations, showing a luciferase activity value >10000 relative light unit (RLU), were selected for each construct combination. The GUS activity was measured and mediated for each of the three time points, 30', 1h and 2h. For each time point the mediated activity values of BBR full length together Min1 were divided to the mediated activity values of Min1 alone. Subsequently these three ratio values were mediated as well.
The results confirmed the previous ones showing this time an activation of 6.7 folds on average (Fig. 3.8.2). Also in this case the measurements and their statistical analysis are included in the appendix.

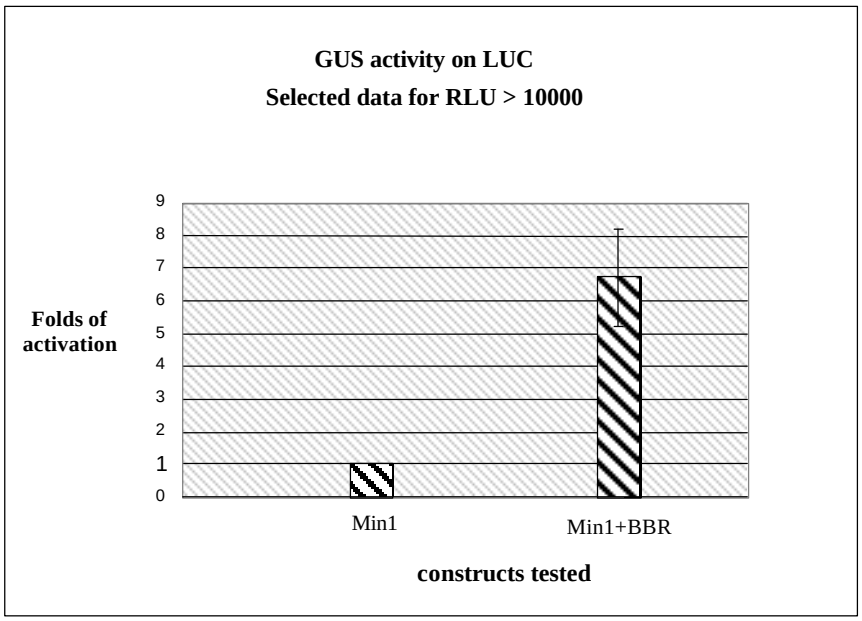


Fig. 3.8.2 The activation of the reporter system harbouring the 305bp regulatory element by means of BBR. Luciferase activity was used as internal control to normalise the reporter GUS activity.

3.9 Mapping of BBR

In order to associate genes with developmental characterised mutants, a map using Proctor and Nudinka cultivars was generated (Castiglioni *et al.*, 1998), integrated recently with a smaller Igri x Danilo linkage map. Several developmental mutants affected in the lemma awn transition zone are mapped. Also different genes have been mapped: BEIL (Wang, 2001 *Ph D thesis* University of Cologne), JUBEL1 and JUBEL2 (Muller *et al.*, 2001), *Bkn*3 (Muller *et al.*, 1995). If using gene specific sequence information is possible to detect a polymorphism between the two parental lines, in this case Proctor and Nudinka, the analysis of how the polymorphism segregates in the F3 double haploid (DH) lines allows to assign the gene to a sub-linkage group on one of the seven barley chromosomes. Recently the detection of single nucleotide polymorphisms (SNP) have been adopted especially investigating non-coding regions.

3.9.1 BBR

Investigating BBR promoter region in Proctor and Nudinka cultivars a deletion of 100bp was detected. This gave the possibility, using a pair of primers (PBBR1, PBBR2) designed around the deletion, to score the polymorphism in the segregating population and directly visualise it on agarose gels. The analysis of 100 DH plants allowed positioning BBR on the high-density linkage map. BBR is on chromosome 4 in sub-linkage group 38 (Fig. 3.8.1).
As it will be discussed later BBR co-maps with *short awn* (*lk5*) a recessive locus responsible for the awn elongation.

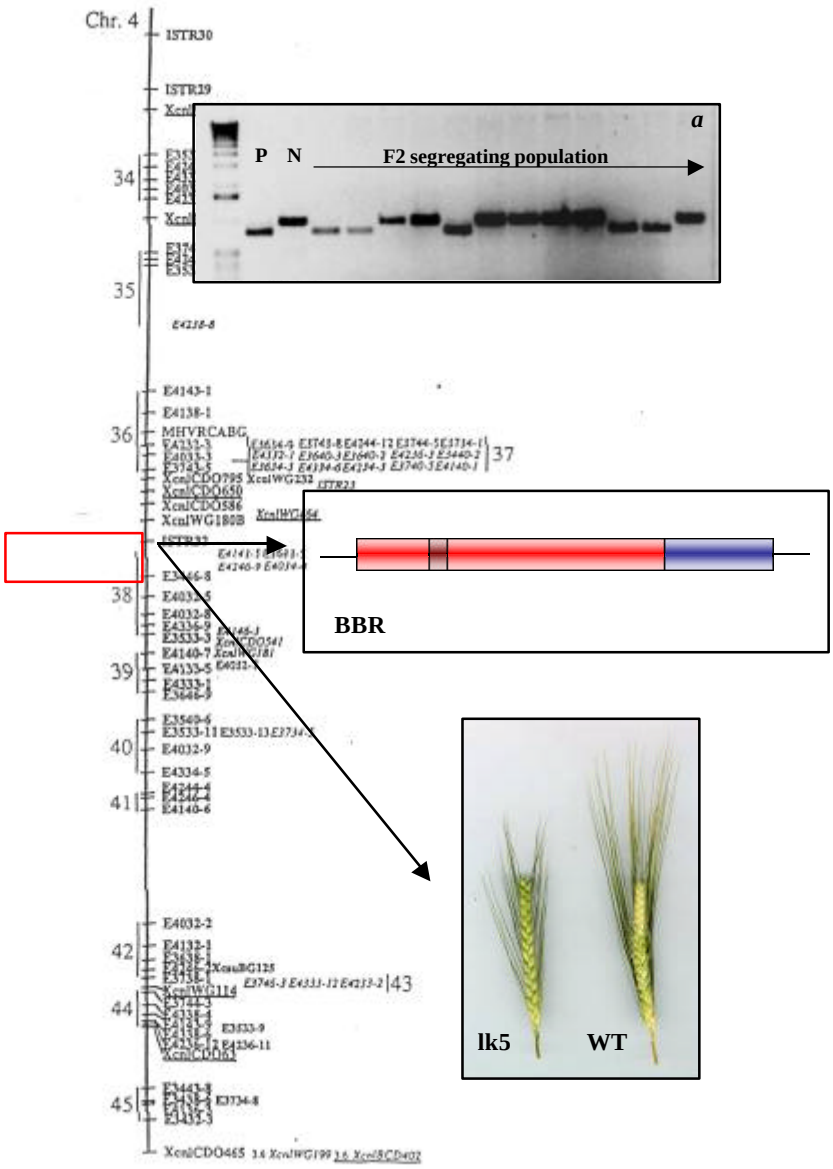


Fig. 3.8.1 BBR is located on chromosome 4, sub-linkage group 38 of the Proctor x Nudinka high density map. The recessive *lk5* mutant maps to the same chromosomal region as BBR. The polymorphism in the promoter region, detected between Proctor and Nudinka, was used to analyse the segregating population as shown in (a).

4 Discussion

Homeobox genes, and class I KNOX genes in particular, play a central role in the maintenance of the meristematic state and in organ fate determination.
This is supported by mutant phenotypes, expression studies and protein-protein interaction, which have revealed a complex network of regulation. Here a possible mechanism of regulation at the transcriptional level of *Bkn3*, a typical example of the class I KNOX genes, is discussed.

4.1 *Bkn3* promoter studies

The tobacco phenotype overexpressing *Bkn3* under the control of the cauliflower mosaic virus (CaMV35S) promoter has been analysed in detail by Lin *et al.*, (2001). Here, a comparison in tobacco of the native *Bkn3* expression pattern with that induced by the 35S is addressed. A minimal *Bkn3* promoter (655bp) capable of directing tissue specific expression was defined. In a fusion with the β-glucoronidase gene (GUS), the minimal element was capable to support the expression of GUS in meristematic regions of transgenic tobacco plants. The same minimal element was not active in SRI tobacco leaf protoplasts.

4.1.1 Expression of *Bkn3* under the control of its own promoter region

In order to ascertain the biological activity of the *Bkn3* promoter, a 4Kb portion the genomic sequence upstream the ATG start codon was fused to the *Bkn3* ORF and the resulting phenotypes of transgenic tobacco plants were analysed. These phenotypes were very similar but not identical to the CaMV35S overexpressing lines. Overexpression in heterologous systems has been widely used for testing the functionality of different class I and II of the KNOX genes. In the study of *Bkn3*, this is of particular significance because *Hooded*, the mutant carrying a mutated version of *Bkn3* is a gain of function dominant mutation responsible for an ectopic expression and the tobacco *Bkn3* overexpressing lines phenocopy, in a dicotyledonous context, the barley *Hooded* mutation (Muller *et al.*, 1995).
The CaMV35S::*Bkn3* lines differed substantially in the overall plant morphology from the wild type tobacco plants. They showed a pronounced dwarfism, a strongly modified leaf and flower phenotypes and the presence of different kind of epiphillous structures growing from the subepidermal layer of the adaxial leaf surface. The reduction in plant high varied between the transgenic lines, with extreme cases where the plant assumed a pronunced bushy appearance.

Dwarfism is a constant characteristic of the overexpression of the class I KNOX genes in tobacco (Lin *et al.*, 2001).

Leaves were reduced in length, more rounded and the lateral veins diverged earlier from the midrib almost at its base. For this reason, the whole leaf assumed a heart shape. The overexpression of *Bkn3* in potato resulted in more compounded leaves and a similar effects were described for the tomato *Kn1* transgenic plants (Sinha *et al.*, 1993). In tobacco, two kinds of epiphyllous structures were found: a) epiphyllous vegetative shoots growing on the lower level of heart-shaped leaves b) inflorescences or isolated flowers on the upper level of lanceolate leaves. Flowers of the main inflorescence carried petals that were fused only at their lower part and with a tubular shape at the tip. Ectopic malformations were present as outgrowths from the petal margin.

The *Bkn3*-promoter::*BKN3*-gene fusion transformed lines resembled the 35S overexpressing lines showing the same unmistakable leaf phenotype. However, even if reduced on average to half the size of wild type SRI tobacco, the *Bkn3*-promoter::*Bkn3* lines were never so extreme as the 35S plants. In the B*kn3*-promoter::*BKN3* lines, the epiphyllous structures were also present but reduced in number. The epiphyllous structures found were always of the first type that is epiphyllous shoots on lower heart shaped leaves. Epiphyllous flowers were not present on any kind of leaf and the main inflorescence flower phenotype was normal. Apart for an overall milder activity of the *Bkn3* promoter compared to the CaMV35S promoter, which was somehow expected, based on what is known from plant constitutive promoters, the similarities of the abnormal phenotypes in the vegetative part of the plant and the absence of abnormalities after inflorescence commitment, suggests that the *Bkn3* promoter activity is restricted to the non-reproductive part of the plant. This is also confirmed by results discussed later.

Considering the molecular nature and the phenotype of the *Hooded* mutation, which results in ectopic flower formation on vegetative like organs, together with the fact that the transgenic tobacco lines expressing *Bkn3* under its own promoter did not fully phenocopy the *Hooded* like syndrome of the 35S *Knox3* transgenics, support the conclusion that *Bkn3* expression is under the control of a multifaceted regulatory system. For example, the *Bkn3*-promoter *Bkn3*-ORF fusion were deprived of *Bkn3* resident intron, which can have regulatory functions (Muller *et al.*, 1995).

4.1.2 The *Bkn3* minimal promoter drives GUS expression in vegetative meristems

Using the reporter GUS fusion approach, the definition of a minimal functional promoter element of 655bp and the spatial and temporal regulation of *Bkn3*were approached.

Artefactual expressions not reflecting the *in vivo* regulation of the gene of interest were reported when using promoter-reporter gene fusion analysis (Thoma *et al.*, 1994). In spite of this the result that can be obtained are informative especially if compared with *in situ* hybridisation patterns in their homologous system. Moreover, the *Bkn3* promoter construct did not only contain 345bp of the proximal *Bkn3* promoter, but also 310bp of the 5' UTR.

Comparison of the nucleotide sequence of the *Bkn3* minimal promoter to databases of *cis* regulatory elements, identified a putative TATA box in the proximity of the transcriptional initiation site, and interestingly, a $(GA)_9$ repeat in the 5' leader sequence (Fig. 4.1).

```
          SalI
          |
  1  GTCGACATTGTTGGACGCACCTGACGGGCAAATAAAAAAAGTGTCTATTAGGCCGATCCA

 61  AACATACAGCCAAGGAAACAAACCATACGTCCGGTCGGCCGATTGAAGTTGCTCTAGGTG

121  AAGCGAATATATTTTCAGAATCCAATCAAAAACTGAAGCAAACATGTCCCCTTCCAAAAA

181  AACAGAAAAAAACCCAACAAAATCAGCTTCGTCTCCACCGGCGATTCTACAAAAAATTCC

241  CTACAAAGCCACTCCAAGCCCTGTGATTTGACCCCAAAAAGAAAGGAGCTCCCTCAGACA

301  CAGCACTATACATTGCAGTACAACAGCGCACCATCCTACACTACACTGTCTCTCTCTCCA

361  AAGGGCGAGGTGTGTACTGCAGTATTGCCCTCCAGCCTTTTCTCTCGATCGTCCCCATCC

421  CCTTCCCTCTAATGATAGCTTAGAGAGAGAGAGAGAGAGAAGGAGAAAAGGCTATTCCGC

481  TCAAGGCAAGGCAGGCAGGTACCCTTCTCCTTCCCCAAACCCAAACCCTTTTTCTCTTTC

541  TCTCCGCCCAAGAACTTCACCTCAAGCTCAATCCGTCTCCTTTTGCGCAGGAGCTAGCTC

                                                          NcoI
                                                           |
601  GGGGACAGTGTCGCCGGGAGCTCGATCGCTGCACCACTACATCACTAGAGATAACCCATG
                                                            M
```

Fig. 4.1 Sequence of the *Bkn3* minimal promoter (655bp) its predicted TATA box, the GAGA box and the starting methionine are indicated in red.

Strong *Bkn*3 promoter activity was detected in the shoot apex soon after germination. At the cotyledon stage of development, the expression domain extended from the shoot apical meristem to the vascular tissue. At later stages of development the expression was localised in lateral vegetative meristems and in the vascular system, but the expression level was lower compared both to the main shoot apex and to the level observed for the gene GUS driven by other known promoters. In

the analysed lines GUS activity was in fact it was detectable only after overnight incubation. In lateral vegetative meristems the expression was confined to the meristem. Signals detected by *in situ* hybridisation studies of different KNOX genes in meristematic tissues support these results, even if expression patterns are to some extent different in the different species. In barley *Bkn1* and *Bkn3* are expressed within the shoot apical meristem in early stages of leaf primordia, including the P_0 (Muller *et al.*, 2001). Similar results were obtained for the tomato genes *Tkn1* and *Tkn2* (Parnis *et al.*, 1997). These findings contrast the absence of expression of major class 1 KNOX genes in leaf primordia of maize, rice, *Arabidopsis*, pea and other species (Jackson *et al.*, 1994; Sentoku *et al.*, 1999). Different molecular mechanisms have been postulated for the definition of borders of KNOX gene domain. On the basis of mRNA detection and of protein immulocalisation of *KN 1* in maize, posttranscriptional regulation and mRNA stabilization have been proposed (Jackson *et al.*, 1994), and backed up by *in situ* data using *Bkn3* as probe, on CaMV35S::*Bkn3* overexpressing tobacco plants (Muller *et al.*, unpuplished). In the latter case despite being under the control of a constitutive promoter, *Bkn3* transcripts was detected specifically in the epiphylls and nowhere else, as if the mRNA was protected from degradation in regions of high meristematic activity. mRNA stabilisation could explain the presence of *Bkn3* transcripts in leaf primordia, associated to the lack of activity of *Bkn3* promoter in the same regions.

These results indicate that KNOX gene product *in planta* interacts with factors present in tissues of special and variable competence for KNOX gene action. Even in closely related species, the expression domain of KNOX genes in the SAM, may or may not include leaf primordia. In this context, for plants like barley, a reduced tissue competence of leaves to respond to over expressed KNOX genes products with a KNOX specific phenotype must be postulated. This view is strongly supported by results from mutant and transgenic grasses which ectopically express the orthologous KNOX genes *Kn1*, *Osh1* and *Bkn3*. Maize and rice upon ectopic expression of the *Kn1* and *Osh1* genes in leaves develop a knotted leaf phenotype (Vollbrecht *et al.*, 1991; Matsuoka *et al.*, 1995). In barley, on the contrary, *Knotted1* expression driven by the constitutive ubiquitin 1 promoter, did not generate any leaf phenotype, but rather developed a phenocopy of the *Hooded* mutation (William-Carrier *et al.*, 1997). The result of the latter experiment and the molecular phenotype of the *Hooded* mutant imply that the only tissue that responded to an up-regulation of *Bkn3* in barley was the lemma awn transition zone (Muller et al., 1995). As suggested by the *Bkn3*-promoter::*Bkn3*-gene studies, also in this case the promoter didn't show any activity in the reproductive part the plant. Moreover the minimal promoter was not active in SRI tobacco leaf protoplasts in transient assay experiments, confirming what was already found in stable transformants the expression was totally absent in leaf primordia and adult leaf organs.

4.1.3 The 305bp element in the intron IV of *Bkn3* is a functional enhancer

The 305bp duplication in the large intron IV of *Bkn3* is responsible, at the molecular level, for the *Hooded* phenotype. The duplication causes the ectopic expression of *Bkn3*, as revealed by RNA gel blot analysis and *in situ* hybridisation experiments (Muller *et al.*, 1995). This, together with the fact that *Bkn3* promoter seems to function only in the vegetative parts of the plant, while the phenotype of the *Hooded* plants is inflorescence specific, suggest that the intron acts like an intragenic enhancer, and that the duplication of the 305bp element inside it interferes with its normal role.

Apart for the short first intron that is only present in *Bkn3*, the genomic organisation of *Bkn3* from barley and *Kn1* from maize is very similar. The intron IV of *Bkn3* corresponds in size and position and shares homology with the third intron of *kn1*. For this reasons it has been stated that the forth intron of *Bkn3* corresponds to the third of *Kn1*. Moreover, the *Knotted 1* mutant, which is dominant as well as *Hooded*, has different alleles generated by several transposons insertions all in this large intron. Apparently even in maize the large intron defines the correct expression pattern of *Kn1*.

In addition to the presence of regulatory *cis*-elements in promoter sequences, more studies emphasize the requirement of other regulatory elements present in the 5'-UTR or 3'-UTR regions of genes (Dietrich *et al.*, 1992; Caspar and Quail, 1993;Lubberstedt *et al.*, 1994). However, much less described are regulatory sequences present in introns or in coding regions (Sieburth *et al.*, 1997). The role of the of intron may mediated by: a) tissue specific differential splicing which would modulate mRNA expression in certain cell types (tissue specific alternative splicing has already been reported in plants; Kopriva *et al.*, 1995; Marillonnet *et al.*, 1997), b) tissue specific stabilisation of the mRNA (Williams-Carrier *et al.*, 1997), c) translational efficiency as described in *Xenopus* oocytes (Matsumoto *et al.*, 1998) d) presence of *cis*-elements in the intronic sequence, as demonstrated in the first intron of the *Arabidopsis* eEF-1 β gene, which contains an enhancer-like sequence able to bind specific nuclear extracts (Gidekel *et al.*, 1996). The most cited example reported in plants, concerns the *Arabidopsis* MADS box gene *AGAMOUS* (AG), which needs, for its correct expression, not only 6kp of the proximal promoter region but also 3.8 kb of its intron sequence. The fine regulation of *AGAMOUS* has been demonstrated to be crucial for flower morphogenesis. At least two expression regulators of this gene have been identified, *APETALA2* (*AP2*) and *CURLY LEAF* (*CLF*). The last is one of the few *polycomb* genes described in plants and is not required for the initial specification of the *AG* expression, but rather to maintain repression during the later stages of development (Yanofsky *et al.*, 1990; Drews *et al.*, 1991; Goodrich *et al.*, 1997).

In order to address the definition of the enhancer sequence modified in *Hooded*, transgenic tobacco lines with different constructs carrying GUS as reporter gene were generated and their activity analysed. First, the intron IV was tested for qualitative and quantitative differences in front of the *Bkn3* promoter (0.655Kb) with one and two copies of the 305bp element inside the sequence, like present in wild type genotypes and in the mutant *Hooded* respectively. For both constructs the temporal and spatial expression among individual transgenic tobacco plants was broadly similar, but GUS enzyme activity varied widely between transformants. The difference in expression might be correlated to the number of integrated loci, to site of chromosomal integration (positional effect); to DNA methylation or to somatic mutations. We included our transgenic plants into 3 groups according to the minimal time for GUS production: 1, 3 to 6h and overnight staining. The analysed lines showed additional domains of expression in respect of the *Bkn3* promoter pattern: the expression was clearly localised in the shoot apex; in vegetative lateral meristems, GUS expression was visible after 3h in almost all the lines analysed, indicating an enhanced activity in comparison to plants harbouring constructs carrying only the *Bkn3* promoter.

Noteworthy was the finding that the intron demonstrated to drive gene expression in the reproductive part of the plant: GUS activity was detected in the inflorescence meristem and at the base of the flower, in particular at the base of petals and carpels. Similar to the vegetative domain, also in this case gene expression was excluded from floral organ primordia and from fully developed flower organs. Relevant differences due to the duplication of the 305bp element inside the intron sequence were not detected. However, it is not possible to exclude an enhanced activity of the duplicated element, because the high variability of GUS expression of the different lines transformed with the same construct made comparisons extremely difficult.

In order to assign to the 305bp fragment the responsability for the enhancer activity, other GUS reporter constructs were transformed in tobacco. Here, the 305bp fragment was located in one, two and three copies upstream of the CaMV 35S minimal promoter. The pattern of expression of these transformants was undistinguishable when compared to the one of the previous intron IV::*Bkn3* GUS lines. GUS coloration was present in both vegetative and inflorescence meristems. Also in this case, differences depending on the 305bp copy number could not be reliably detected.

In conclusion we assumed that both the *Bkn3* promoter and intron sequences regulate *Bkn3* expression in the major shoot apical meristem, while the intron IV of *Bkn3* additionally activates transcription within the floral meristems and at the base of the flower. The intron sequence seems also responsible for an increased level of transcription in the lateral vegetative meristems and the 305bp element acts like a *cis*-regulatory element inside the intron sequence. Taken together these data suggest that the intron IV of *Bkn3* harbour enhancer sequences capable of support gene action

adding organs to the plant body in barley and in transgenic tobacco. In this sense the ectopic expression of *Bkn3* at the tip of the barley awn, before the formation of the hood, is the response to an enhancer located within the duplication of the 305bp sequence. This enhancer was activated by factors that define such sites of organ additions at the tip of the barley awn.

4.2 Factors regulating *Bkn3* expression

As previously discussed, the presence of proteins functionally interacting with *cis*-acting elements located, within the regulatory region of *Bkn3*, should be responsible, with *Bkn3* itself, for new plant tissue competence. The identification of these factors that bind the enhancer sequence and regulate *Bkn3* expression is of primary interest. To address this problem Dr. Kai Muller adopted the yeast one-hybrid system approach. The one hybrid system is an *in vivo* genetic approach, which allows isolating novel genes encoding proteins that bind DNA target elements.

The core of the yeast one-hybrid screening method is not only the protein DNA interaction but also the gene transcription activation, which is the biological consequence that derives from it. In this case, the disadvantage lies in the heterologus character of the system because yeast factors could take part in the interaction and lead to the so called "yeast artefacts". If, for example, in the yeast genetic background is present a sequence coding for a protein (bridge protein) that is able to bind the target DNA on one side and, on the other side, one of the proteins encoded by the cDNA library, it is possible that this new complex activates the reporter system leading to the identification of a protein that does not present any target DNA binding properties. To overcame this problem, useful information can be found searching for homologies to proteins present in databases. Already described DNA binding domains or putative nuclear localisation signals can be reassuring about the reliability of the screen. Anyway, a protein-DNA interaction *in vitro* by means of an electro mobility shift assay is considered to be the best proof.

The 305bp element was cloned in one and three copies in the reporter vector used for the screen. The assembly of the 305bp sequence 5´to the yeast minimal promoters *his3* and *cyc1* abolished gene expression when transformed into the yeast strain. Therefore, it was assumed that one or more factors encoded by the yeast genome interacted with the 305bp element repressing the basal promoter activity present in the controls. This fact suggested that the 305bp sequence acts as a functional element in yeast acting as a silencer. To establish a yeast reporter strain suitable for the screening of a *Hooded* expression library a strain with the 305bp element 5´ to the *his3* reporter gene was mutagenized by EMS. The use of the mutagenised strain led to the identification of four proteins: BEIL (barley ethylene insensitive 3 like), BGRF (barley growth factor), BBR and BAPL

(barley apetala 2 like). For BEIL, BGRF and BAPL, homologies to putative DNA interacting proteins were found while for BBR the *in vitro* proof of binding was achieved (discussed in chapter 4.3).

Since the four proteins were identified for the ability to interact with the 305bp sequence, a possible interaction at the protein level of these factors, to mediate transcriptional regulation, was tested using the two-hybrid assay. The proteins didn't appear to interact in any of the combinations tested. A possible interaction cannot however be excluded, since none of the proteins tested with the two-hybrid system had the complete open reading frame. While BBR was characterized in more details, the function of other three proteins was only be derived from speculations based on their homologies with proteins of known function, and from their putative capability to interact with the 305bp element responsible for the barley *Hooded* mutant phenotype.

BEIL, which encodes for a peptide of 550 a.a., was isolated twice. Data from one and two hybrid systems suggest that it contains an activation domain, functional in yeast, located in the first 40 amino acids of his sequence. The protein is predicted to contains five clusters of basic a.a. dispersed throughout the sequence and between the second and third an α-helix structure. This structural feature supports the fact, already found in yeast, that this protein has DNA binding properties. It shares a high degree of homology with the protein ETHYLENE INSENSITIVE LIKE 3 (EIL3) from *A. thaliana.* In *Arabidopsis*, three EIL proteins have been isolated, all of them share homology with ETHYLENE-INSENSITIVE3 (EIN3), which has been demonstrated to severely limit the plant response to the gaseous hormone. EIN3 and the three EIL proteins take part, acting in the nucleus of the cell, in the ethylene cascade that starts from ETHYLENE RECEPTOR1 (ETR1) on the plasma membrane. Interestingly, it has been postulated that for gene regulation, *eil 3* and *eil* work together with transcription factors belonging to the EREBP subfamily (Chao *et al.*, 1997). BAPL, one of the other barley factors isolated, belongs to this subfamily.

Not much can be said for BGRF. Using the 385 a.a. full lenght protein sequence, a BLAST search was able to retrieve a highly homologous sequence from *O. sativa*, as well as two from *A. thaliana*. The function of these homologues is unknown, even if the rice gene is annotated to have a putative function in the regulation of stem growth depending on giberellin induction and the property to bind RNA molecules.

BAPL contains a well-defined DNA binding domain. This protein belongs to AP2/EREBP family of transcription factors and, having a single AP2 conserved domains, it falls in the EREBP subfamily. This is a large family of transcription factors found exclusively in plants. Data from *Arabidopsis* genome indicated the existence of 140 members of this family, 20 belonging to the AP2 subfamily and 120 belonging to the EREBP subfamily (Riechmann *et al.*, 2000). The

subdivision of the AP2/EREBP multigene family on the basis of molecular criteria may reflect a functional dichotomy. The AP2 subfamily genes, whose functions have been determined by mutant analyses (*Arabidopsis AP2* and *ANT*, and maize Gl15) act as key regulators in developmental processes, whereas the EREBP subfamily members, so far characterised, appear to be involved in responses to biotic and environmental stress. As for the other members of the EREBP subfamily and in contrast with the AP2 subfamily members, the AP2 DNA binding domain sequence of *BAPL* is not interrupted by introns (Riechmann and Meyerowitz, 1998). EREBP proteins bind to the GCC box, an 11bp consensus sequence (TAAGAGCCGCC) that has been shown to function as an ethylene responsive element (Yamamoto *et al.*, 1999). Even if three GCC repeats are present in a 15bp fragment of the 305bp sequence (97-112), no perfect GCC box was found. This information allows hypothesising for the role of the plant hormone ethylene in development, which could connect KNOX gene action with physiological hormone response, even if the evidence available for this kind of connection is very weak.

4.3 BBR

The use as a probe for the screen of the *K*-Atlas λ cDNA library of the longest partial sequence of BBR, identified using the one-hybrid method, allowed the definition of the full length sequence and consequently, the full open reading frame of the protein, which consists of 350 a.a.. Four BBR sequences, of different lengths, were cloned independently with the yeast system, a fact alone that already gives a certain degree of reliability to the screen.

All the four sequences share identity at C-terminal end of the encoded proteins, but extend differently towards the N-terminal end. The shortest sequence encodes a peptide of 128 a.a. and the longest a protein of 325 a.a., which is 25 a.a. shorter than the full length BBR. In barley, BBR is a single copy gene and no introns are present in the genomic sequence. Unfortunately, sequence homologies with proteins present in databases don't help in defining a clear function of BBR. The two highest homologues from *Arabidopsis* (74% and 49%) encode for "putative proteins" and to them any function is assigned. A homologous protein from rice (60%) has been reported to have nucleic acids binding properties (unpublished). Homology among these sequences is clearly clustered at the last 120 a.a. at the C-terminal end of the protein. This suggests a functional role of this region maintained during the speciation processes. Moreover, considering that the shortest functional protein in yeast consists of the last 128 a.a at the C-terminal end, and that this portion of the protein contains several basic amino acids, it is easy to presume, that this region is the DNA binding domain or that, at least, the DNA binding domain is contained in this region.

An additional feature of BBR is the presence of a glutamine rich domain that consists of a long stretch of QH repeats (from 63 to 123). This domain is absent in the *Arabidopsis* homologues and is just partially present in the rice sequence. Glutamine rich domains have been described in the literature in the contest of factors able to remodel chromatin structures. Different functions have been assigned to this domain, from promoter distortion and protein multimerization to complex fibre formation (Wilkins and Lis, 1999). In *A. thaliana*, BBR consists of a small gene family of seven members. At1g68120 (AtBBR1 on chr.1), At2g01930 (atBBR2 on chr.2), At1g14685 (atBBR3 on chr. 1), At5g42520 (atBBR4 on chr. 5), At2g21240 (atBBR5 on chr.2), At4g38910 (atBBR6 on chr. 4), At2g35550 (atBBR7 on chr.2). The genomic organisation of these genes is very similar to BBR; none of them has the coding region interrupted by introns. For Atbbr 2, 3, 4 and 5 also information on the 5' UTR is reported and interestingly all these clones contain several TC repeats in this region. As discussed later, we demonstrated that BBR binds these repeats leading to the assumption of a possible self-regulation mechanism of these *Arabidopsis* genes. Sequence comparison of the peptides, derived from the genomic information of the seven genes, was performed, revealing that also in this case the homology was mostly grouped in the putative DNA binding domain of BBR (Fig.4.2).

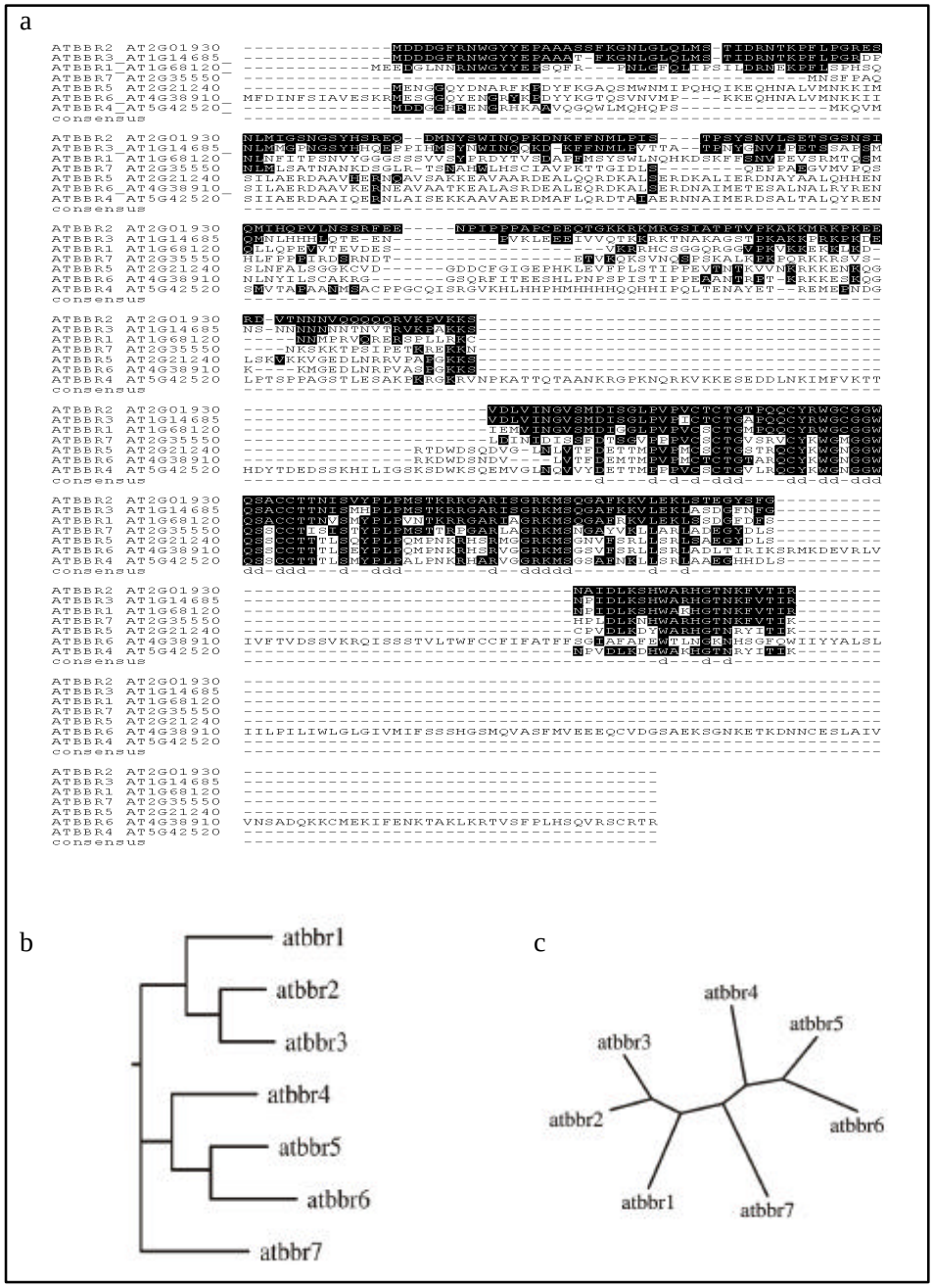


Fig. 4.2 a) Alignment of the seven predicted BBR protein homologues from *Arabidopsis*. High homology is clustered in the region of BBR that contains the DNA binding domain. b) rooted and c) unrooted trees based on the amino acidic comparison of the seven proteins.

4.3.1 BRR binds to the 305bp element

Using the electro-mobility shift assay (EMSA) we demonstrated the ability of BBR to bind the 305bp element *in vitro* was demostrated. This excluded yeast artefacts in the yeast one-hybrid screen. A first shift was found using BBR transcribed and translated *in vitro.* In separate experiments, two differently generated 305bp fragments were used as hot probes. A purified *GST*-fusion of BBR was able to clearly retard the target DNA as well. The experiment was repeated several times and controls were included. The binding was extremely reproducible and the BBR-DNA interaction was very specific as competition experiments revealed. The presence in the retardation assay of additional bands is commonly interpreted as a sign of homodimerisation. Confirming what was already found using the two-hybrid assay, BBR does not seem to dimerise, because no additional supershifts were detected. However, it has to be noted that the sequence used for the two hybrid assays was not complete, lacking 25 a.a. at the N-terminal end.

Transient expression studies in tobacco protoplasts gave additional evidence of binding, which can be considered an *in planta* proof of interaction. Using a GUS reporter system, where the 305bp target element was located upstream of the CaMV35S minimal promoter, the ability of BBR to influence the levels of reporter gene expression was tested. Indeed, BBR was able to activate reporter gene expression up to six folds on average using two kinds of normalisation strategies, based on the total protein level or on the activity of an internal LUC control. A chimeric fusion of BBR together with the etherologous VP16 AD gave the same level of activation. In protoplasts BBR acts as an activator mediated by the binding to the 305bp element. However this cannot be directly extended to anticipate a function for BBR, because in barley, as discussed before, other factors, specific for the lemma-awn transition zone, may take part in the regulation of *Bkn3* and in the establishment of a local addition of new organs.

4.3.2 BBR binds specifically a 16bp $(GA)_8$ sequence contained in the 305bp element

Using the reproducibility and specificity of the binding of BBR to the 305bp fragment, it was possible to define and delimit the exact DNA binding site within this sequence. The 305bp sequence can be considered an upper limit for bandshift experiments. Sequences used for this purpose are usually much smaller and the minimum sequence for protein DNA interactions can just be a few base pairs long. One of the many known examples is the homeodomain-leucine zipper transcription factor DNA binding site that consists of a sequence of only 9bp (Johannesson *et al.*, 2001). The 305bp sequences was reduced to smaller fragments and the specific binding of BBR allowed the

definition of a repetitive $(GA)_8$ repeat acting as the smaller double stranded DNA binding site. BBR was able to shift sequences containing the described 16bp sequences, but not fragments contained in other regions of the 305bp sequence, showing once more the reliability of the experimental procedure. The $(GA)_8$ repeat is located at the 3' end of the 305bp element from position 249 to position 264 (Fig 4.3). Interestingly, as mentioned, a perfect $(GA)_8$ repeat was found in the 5' UTR of *Bkn3* (Fig. 4.1).

This particular repetition is common in regulatory regions of different genes implied in developmental processes in plants and animals. In database searches, GA repeats have been found in the cDNA 5' UTR of the *SUPERMAN* gene from *Arabidopsis*, as well as in the 5' UTR of the *OCTANDRA* gene from *Anthirrinum.* Also in the *Arabidopsis* homologues of BBR, repeats of such kind were found in the 5' UTR. However in all these cases a functional activity of the GAGA box has not jet been defined.

Evidence of the involvement of this sequence in gene regulation in general, and in regulation of developmental processes in particular, derives from analogies in the animal kingdom. The vasopressin V1b receptor gene promoter from rat showed to be functional in transient expression experiments only when coupled with its first intron sequence, which contains a GAGA box, in this case consisting of 25 TC repeats. (Rabadan-Diehl *et al.*, 2000). Furthermore, in *Xenopus* a GAGA box, 11 TC repeats, is involved in the transcriptional regulation of the stromelysin-3 gene. *In vitro* DNA binding and mutational studies, using nuclear protein extracts, have provided strong evidence for the participation of GAGA or GAGA-like binding factors in this type of regulation (Li *et al.*, 1998). In mouse the presence of a GAGA box, in this case just a GAGAG sequence, in the promoter region of the *hsp70.1* gene, is required for the correct expression of the gene. Hypothesising the presence and the action of GAGA binding factors, the authors speculate on novel mechanism of gene derepression in early mouse embryos (Bevilacqua *et al*., 2000).

In all the previously mentioned cases the presence of a factor that could bind the (GAGA) box and through this interaction regulate gene expression was assumed from the analysis of the full nuclear protein extracts, but in none of these cases had the gene encoding for the factor been cloned or the resulting protein characterised. However in *D. melanogaster*, one GAGA binding factor, GAF, has been cloned and the activity of the GAGA box has been directly related to the regulation of gene expression in developmental processes. TheGAF factor has been shown to be involved in several distinct aspects of chromosome dynamics (Granok *et al.*, 1995; Read and Driscoll, 1997;Wilkins and Lis, 1997). The GAGA binding factor is encoded by the essential *Trithorax-like* (*Trl*) gene (Farkas *et al.*, 1994; Bhat *et al.*, 1996). Immunofluorescent staining of polytene chromosomes has revealed that many euchromatic genes include binding sites for GAF, suggesting a general role in

transcription control (Tsukiyama *et al.*, 1994; Benyajati *et al.*, 1997). A number of studies have shown that GAGA-binding elements coincide with Dnase I-hypersensitive sites at promoters, indicative of *in vivo* chromatin remodelling (Wu, 1980; Costlow and Lis, 1984; Cartwright and Elgin, 1986; Lis and Wu, 1993; Lu *et al.*, 1993; Shopland *et al.*, 1995). Biochemical studies indicate that GAF activates RNA polymerase II transcription by counteracting chromatin repression (Croston *et al.*, 1991;Tsukiyama *et al.*, 1994; Okada and Hirose, 1998). This is likely to involve the disruption of nucleosomes harbouring GAGA-binding sites by GAGA acting together with the energy-dependent chromatine remodelling factor NURF (Tsukiyama *et al.*, 1994; Tsukiyama and Wu, 1995; Tsukiyama and Wu, 1996).

The function of GAGA is not limited to gene-specific transcriptional activation. *Trl* mutations are dominant enhancers of position effect variegation, indicating that GAF counteracts heterochromatic silencing (Farkas *et al.*, 1994). GAGA has also been implicated in the functioning of the polycomb response elements (Strutt *et al.*, 1997). Thus, GAF is a multipurpose protein that mediates gene-specific regulation but also plays a global role in chromosome function. As mentioned before a direct link between GAGA binding factors and regulation of homeotic genes has been demonstrated in *D. melanogaster*. The *trithorax* group of genes (*trxG*) is required for mantainence of the active transcriptional state of homeotic genes (Kennison, 1995). In contrast, *Polycomb* genes form a group (*PcG*) that are responsible for maintainence of the repression or silencing. To date, thirteen *PcG* genes have been molecularly characterized. All the characterized *PcG* genes, with the exception of *pleiohomeotic* (*pho*), encode chromatin-associated proteins with motifs characteristic of chromatin-bound proteins or suggestive of protein-protein interactions. PHO is the only PcG protein so far characterised that has specific DNA binding properties. The consensus DNA binding sites (ATGGC) have been identified (Brown *et al.*, 1998).

GAF takes part in the regulation mediated by *trxG* and *PcG* genes and is thought to be required for the mantainence of the activation of homeotic genes. Consistent with this, *Trl* mutations show phenotypes indicative of the loss of activity of homeotic genes and enhanced loss of function phenotypes (Farkas *et al.*, 1994). Nevertheless, other studies clearly implicate *Trl* in the repression of homeotic gene transcription. *Trl* therefore should be classified as belonging to both the *trxG* and *PcG* groups. *TRL* would not be the first gene to be classified as belonging to both positive and negative regulatory groups (Bustura *et al.*, 2001).

The fact that BBR is a putative regulator of *Bkn3* and binds to two GAGA boxes within *Bkn3* sequence, one in the fourth intron and one in the 5' UTR, underlines strong analogies between the two systems. Moreover, in plants, as described for *D. melanogaster*, *polycomb* genes take part in the regulation of homeotic genes (Goodrich *et al.*, 1997). In the 305bp sequence, 3' to the GAGA box,

a perfect consensus binding site (ATGGC) for the *polycomb* gene PHO from *D. melanogaster* is present, suggesting that a similar factors might also take part in this kind of regulation in barley (Fig 4.3).

BBR has no homology with the only GAGA factor cloned up to now, GAF, except for the Q rich region described before. In GAF, this region is responsible for promoter distortion, single strand binding and multimerisation activities. The Q domain may interact with melted regions as evident by its relatively strong affinity for single stranded DNA. It has been proposed that glutamine residues may function to facilitate strand separation, as amide groups in amino acids side chains have been shown to have destabilising effects on DNA through interaction with bases in single stranded regions (Wilkins and Lis, 1999).

```
  1  GAGTTCTGAC CAGTTGCAGT GCAGTGGCAT GGACTGCTAG ACACTGGTCT
 51  ATGTGCCTCG AGAACCATAG AAACAGTCAA GTGCTTGCCA CAGCCTACAG
101  CCACCATTAT TTGATGTGAG ATCTCCAGCA ATTTATGCGT GATCTCCTTG
151  GCAGCAATAT TATATGCATG CTAGCTCCAT CTACTAGTGT AGGCATGTAG
201  CTAGACTAGA TACCATGTTG CTGTATTTTG CGACAGCCTT ATCTCTGGTC
251  TCTCTCTCTC TCTCATGAAT AATGGCGGTC AAGAGACGTT GGATGCTTTC
301  CAAGG
```

Fig. 4.3 The 305bp element sequence. The $(TC)_8$ repeat (red) represents the GAGA box , which is the BBR binding site. At the 3', the (ATGGC) (blue) sequence defines a polycomb binding site.

4.3.3 Overexpression of BBR in *N. tabacum*

BBR was overexpressed in transgenic tobacco under the control of the CaMV 35S promoter.

A first set of transformations involved a truncated version of BBR alone and in fusion with the yeast GAL4 AD. This partial BBR is 25 a.a. shorter than the full length BBR .The putative DNA binding domain and the Q-rich region were both included in these constructs. The expression of the transgene in the selected plant lines was detected by RNA blot analysis and the phenotype was investigated over three generations. Several lines harbouring and expressing the fusion between BBR and the etherologous yeast activation domain showed a clear leaf phenotype, while plants expressing the partial BBR alone were undistinguishable to wild type SRI tobacco plants.

The striking phenotype of the lines overexpressing GAL4::BBR was reproducible over the generations analysed. The plants all had leaves showing a marked elongated shape. Measurements of the length and width of the major leaf axis, at two different stages of development, revealed that differences in both the anticlinal axis of growth were involved in the process. The reduction in growth along the trasversal axis was much more pronounced than the increase in longitudinal growth. The phenotype was not only restricted to the vegetative stage of the plant, but was also detectable in the reproductive organs, where the sepals and petals of the flower were clearly longer. The sepals were measured and compared to wild type. In average they showed an increase of one third of the length of the organ.

The overespression of the Gal4AD::BBR altered the leaf morphology from very early stages of development. Already at the cotyledon stage, the leaflets presented a curled appearance that could be interpret as an increased tension along the major longitudinal axis of the leaf. It has already been reported that a single alteration in the frequency or orientation of mitoses at an early stage can have many effects on the adult phenotype (Stebbins and Price, 1971). Two different mechanisms could lead to this phenotype, an increased frequency of transversal mitotic divisions and/or an increase in cell elongation along the longitudinal axis. The fact that the leaves, compared to wild type, were more affected in width than in length, suggests that an alteration in the frequency and direction of cell division, rather than an increase in cell elongation, is likely to be responsible of this morphological effect. From this point of view the described phenotype is interesting because it shows characteristics that are opposite to the *Bkn3* overexpressing tobacco plants. In fact, plants that ectopically express class I KNOX genes present heart shaped leaves as a consequence of a reduction in longitudinal growth and an increase along the transversal axis. This might suggest that BBR should work in the down regulation of KNOX genes but the fact that the constructs leading to the morphogenetic effect are just the ones that present an etherologous activation domain make it very difficult to formulate such a kind of speculation. However, a second set of transformed plants, overexpressing the full length BBR, exhibit the same altered leaf morphology. This last result suggests a functional role of the last 25 a.a. at the N-terminal end of BBR, since none of the plants overexpressing the truncated version of BBR was able to lead to the described altered leaf shape.

From the etherologous overexpression strategy, it is difficult to state the real function of BBR in barley and to deduce the effect on *Bkn3* expression. However, we can conclude that, when overexpressed, BBR leads to morphogenetic effects involving leaves and flower structures of leaf origin. Moreover, the overexpression that drives the elongation of the leaves recalls by contrast the phenotype of plants overexpressing *Bkn3*, confirming that both genes take part in the same developmental program

4.3.4 The search for function: a candidate gene approach

Differences in developmental processes between monocots and dicots and the use of the 35S CaMV promoter, which is a strong constitutive promoter, make both stable and transient overexpression studies in tobacco an unclear system for assigning gene function. This is particularly true considering the peculiarities of the *Hooded* mutation. In fact, as previously discussed, it is the lemma-awn transition zone that has the competence to respond to the up-regulation of *Bkn3*. The lemma represents a reduced vegetative leaf comparable to a leaf sheath, while the awn is comparable to the leaf blade (Clifford 1988).

Finding a mutant in the homologous system and analysing its phenotype is the only clear way to deduce a function for these genes.

Due to the large genome size and to the low transformation efficiency of barley, genomic tools such as reverse genetics or complete database information are not yet fully available for this plant. On the other hand, the diploid chromosomal state allowed an easier construction of a genetic map (Castiglioni et al., 1998), to which part of a large collection of developmental mutants could be mapped (Pozzi *et al.*, 2002 submitted). Most of the mutations mapped had abnormalities involving the lemma awn transition zone (*calcaroides*, *suppressor of Hooded*, *leafy lemma*, *short awn*). Following a candidate gene approach, the objective is to associate gene DNA sequences and mutant phenotypes based on their map position. For genes involved in the *Bkn3* network, candidate phenotypes of particular interest are the suppressors of *Hooded* (*suK*). *Hooded* is dominant and homozygous viable which made it possible to screen for second site mutations in order to identify suppressors of the *Hooded* phenotype. 41 *suK* were identified and grouped in 5 *suK* recessive loci: *b*, *c*, *d*, *e*, and *f*. The suppressor loci *suKb*, *c*, *e* and *f* all mapped on linkage group 1, in a short interval represented by sub-linkage groups 5 to 7. *suKd* was assigned to chromosome 7 between sub-linkage groups 66 and 67 (Roig *et al.*, 2002 submitted) (Fig. 4.4).

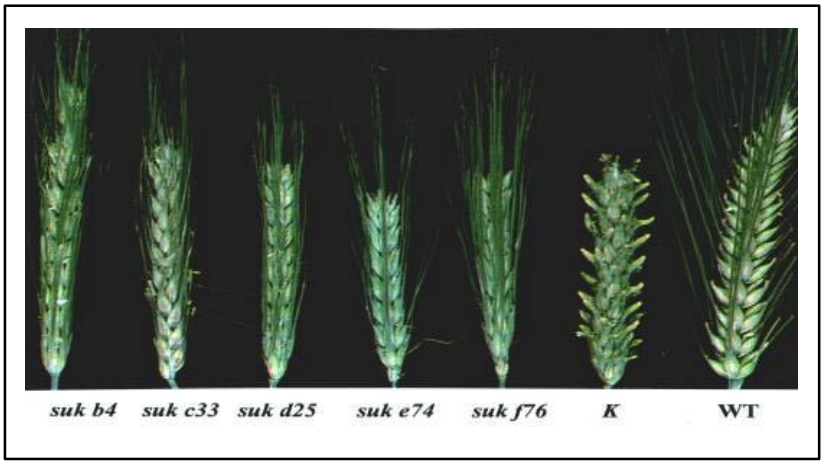


Fig. 4.4 *suppressors of Hooded* (*suK*) compared to *Hooded* (*K*) and wild type (WT). The suppressor loci *suKb, c, e* and *f* all map on chromosome 1 in a short interval represented by sub-linkage groups 5 to 7 , while *suKd* maps on chromosome 7 between sub-linkage groups 66 and 67 (Roig *et al.*, 2002; submitted).

BEIL has been mapped using a single nucleotide polymorphism detected in a 420bp fragment within the intron sequence (Wang, 2001 *Ph D thesis* University of Cologne). Interestingly, BEIL falls in the same sub-linkage groups on chromosome 1, where the four suppressors are located. Expression analyses of BEIL in the four different suppressor background and parasequencing strategies are on the way to test the association of this gene with one of the 4 putative *suk* loci.

In the case of BBR, a polymorphism between Proctor and Nudinka consisting of a deletion of 100bp (this thesis) in the promoter region, allowed to position the locus on chromosome 4, sub-linkage group 38. In the same sub-linkage group, in linkage with the same AFLP markers (E4034-4, E4743-5 and E4246-9) the barley *short awn* 5 (*lk5*) mutant was located. The fact that the *lk5* recessive mutation is responsible for a reduced development of the awn of barley, which is comparable to the leaf blade of monocots or to the leaf lamina of dicots, together with the fact that the over expression of *BBR* in transgenic tobacco plants leads to the elongated leaf phenotype, makes the *lk5* mutant a good candidate for the *BBR* loss of function phenotype in the homologous system. Also in the case of *BBR*, expression analysis and sequence information will help to validate the candidate gene-phenotype association.

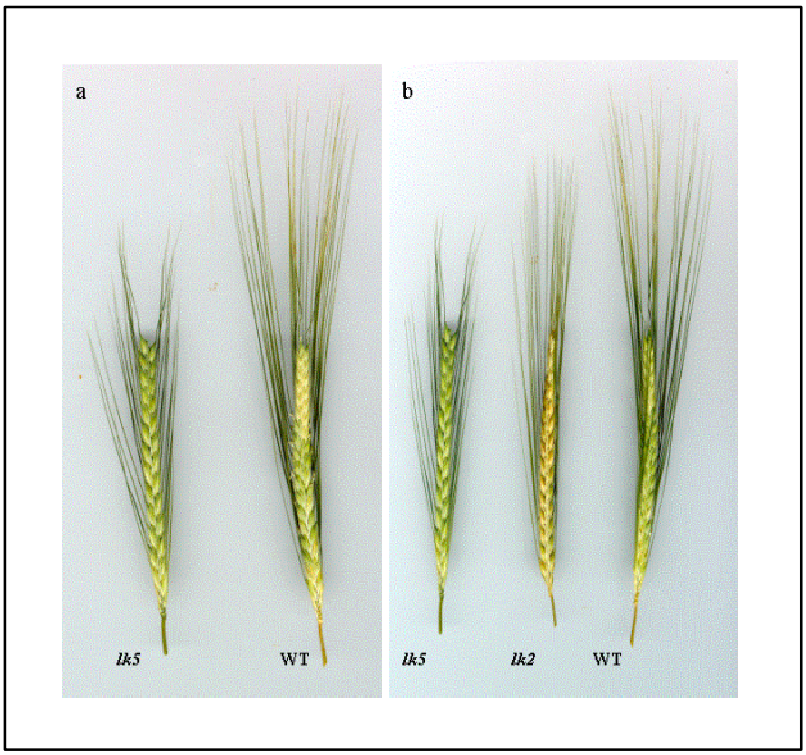


Fig. 4.5 Comparison of a) *short awn 5* (*lk5)* and b) *short awn* 2 (*lk2)* to wild type (WT). *lk5* is located on chromosome 4 sublinkage group 38.

Summary

The class I of the plant KNOX homeobox genes play a central role in the maintenance of the meristematic state and in leaf cell fate determination. *Bkn3*, a typical example of this family, was associated with the barley *Hooded* mutant. The difference at the molecular level between the mutant and the WT alleles is the duplication of 305bp in the forth intron of *Bkn3*.

In this thesis a possible mechanism of regulation at the transcriptional level of *Bkn3* is discussed. The biological activity of the *Bkn3* promoter was demonstrated analysing the phenotype of transgenic tobacco plants, stably transformed with *Bkn3* promoter *Bkn3* gene fusion construct, in comparison to transgenic tobacco plants overexpressing *Bkn3* under the constitutive CaMV 35S promoter. Using the reporter GUS fusion approach, a minimal functional promoter element of 655bp and the enhancer like function of the 305bp element -responsible in barley for the mutated phenotype- were defined in transgenic tobacco lines. The *Bkn3* promoter drives expression in vegetative meristems. The 305bp element showed an enhanced activity in the same expression domain but in addition it is capable to drive expression in the reproductive part of the plant at the base of the flower. A cDNA library, generated from the *K*-Atlas barley mutant, was screened to acquire further information on the predicted full open reading frames and the 3' and 5' untranslated regions (UTR) of four genes BEIL, BGRF, BBR and BAPL, previously cloned using the 305bp as target sequence in a one hybrid system (K. Muller). BBR was characterised in details. The binding property was confirmed using EMSA and tobacco leaf protoplasts transient transformation. In the latter case BBR was able to activate the GUS reporter gene activity up to six folds on average. Several deletions were used for band shift experiment and this allowed to define a $(GA)_8$ repeat as the precise binding site within the 305bp element. Moreover, BBR showed to have morphogenetic effects when overexpressed as fusion with the GAL4AD in stable tobacco transformants. BBR maps on chromosome 4 in sub-linkage group 38 of the Proctor X Nudinka high-density map. In the same sub-linkage group co-maps the recessive *lk5* mutant that is phenotypically characterised by short awns. The fact that BBR acts, in transient assay, as an activator and that the stable overexpressing lines present an elongated leaf phenotype, makes the *lk5* mutant a good candidate for the *BBR* loss of function phenotype in the homologous system.

Zusammenfassung

KNOX-Homöobox-Gene der Klasse I spielen eine zentrale Rolle in der Aufrechterhaltung des meristematischen Zustandes und der zukünftigen Determination von Blattzellen. BKN3 assoziiert als typischer Vertreter dieser Familie mit der Hooded-Mutante in Gerste. Auf molekularer Ebene besteht der Unterschied zwischen Mutanten- und WT-Allel in einer Duplikation von 305bp im vierten Intron.

Thema der vorliegenden Arbeit ist der Regulationsmechanismus von BKN3 auf transkriptioneller Ebene.

Die biologische Aktivität des Bkn3 Promotors wurde anhand phänotypischer Analyse in transgenen Tabakpflanzen gezeigt, die stabil transformiert waren mit dem Bkn3-Promotor - Bkn3-Gen Fusionskonstrukt und mit transgenen Tabakpflanzen verglichen, die Bkn3 unter dem konstitutiven CaMV 35S Promotor überexprimierten. Durch die Verwendung eines GUS-Reporter Konstruktes konnte ein minimal-funktionelles Promotorelement von 655bp und eine Enhancer-ähnliche Funktion für das erwähnte 305bp Element in transgenem Tabak identifiziert werden.

Der Bkn3 Promotor treibt die Expression in vegetativen Meristemen. In der gleichen Expressionsdomäne zeigt das 305bp Element ebenfalls eine verstärkte Aktivität. Außerdem ist es jedoch in der Lage, die Expression im reproduktiven Teil der Pflanze an der Blütenbasis zu steuern. Eine cDNA Bank, die aus der K-Atlas Gersten Mutante hergestellt wurde, wurde durchmustert, um weitere Informationen zu gewinnen über die vorhergesagten Vollängen-ORFs (offene Leseraster) und die 3´ und 5´ untranslatierten Regionen (UTR) der vier Gene BEIL, BGRF, BBR und BAPL, die zuvor unter Verwendung des 305bp Fragments als Zielsequenz im one-hybrid-System kloniert wurden (K. Müller).

BBR wurde im Detail charakterisiert. Die DNA-Bindungseigenschaften wurden anhand von EMSAs (electric phoretic mobility shift assays) und transienter Transformation von Blattprotoplasten bestätigt. Bei der Transformation konnte BBR die Aktivität des GUS-Reporters um den Faktor 6 steigern. Verschiedene Deletionen wurden in Gelretardations-Experimenten eingesetzt. Dies ermöglichte die Identifizierung einer $(GA)_8$-Wiederholung als exakter Bindestelle innerhalb des 305bp Elements. Weiterhin zeigte die Überexpression von BBR als Fusion mit der GAL4AD (GAL4Aktivierungsdomäne) in stabil transformiertem Tabak morphogene Effekte. BBR kartiert in der Proctor x Nudinka high-density Karte auf Chromosom 4 in der Unterkopplungsgruppe 38. In der gleichen Kopplunggruppe kartiert

ebenfalls die rezessive *lk5* Mutante, die phänotypisch kurze Grannen zeigt. Die Tatsache, daß BBR im transienten Expressionsversuch als Aktivator wirkt, und daß stabil überexprimierende Linien einen elongierten Blattphänotyp aufweisen, machen die *lk5* Mutante zu einem guten Kandidaten für den BBR "loss of function" Phänotyp im homologen System.

Acampora, D.; D'Esposito, M.; Faiella, A.; Pannese, M.; Migliaccio, E.; Morelli, F.; Stornaiuolo, A.; Nigro, V.; Simeone, A. and Boncinelli, E. (1989). "The human HOX gene family." Nucleic Acids Res **17**(24): 10385-402.

Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W. and Lipman, D. J. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-402.

Alvarez-Venegas, R. and Avramova, Z. (2001). "Two *Arabidopsis* homologs of the animal trithorax genes: a new structural domain is a signature feature of the trithorax gene family." Gene **271**(2): 215-21.

Ansari, A. Z.; Mapp, A. K.; Nguyen, D. H.; Dervan, P. B. and Ptashne, M. (2001). "Towards a minimal motif for artificial transcriptional activators." Chem Biol **8**(6): 583-92.

Avivi, Y.; Lev-Yadun, S.; Morozova, N.; Libs, L.; Williams, L.; Zhao, J.; Varghese, G. and Grafi, G. (2000). "Clausa, a tomato mutant with a wide range of phenotypic perturbations, displays a cell type-dependent expression of the homeobox gene LeT6/TKn2." Plant Physiol **124**(2): 541-52.

Bardwell, V. J. and Treisman, R. (1994). "The POZ domain: a conserved protein-protein interaction motif." Genes Dev **8**(14): 1664-77.

Bartels, D. and Thompson, R. D. (1983). "The characterization of cDNA clones coding for wheat storage proteins." Nucleic Acids Res **11**(10): 2961-77.

Becker, D.; Kemper, E.; Schell, J. and Masterson, R. (1992). "New plant binary vectors with selectable markers located proximal to the left T-DNA border." Plant Mol Biol **20**(6): 1195-7.

Becraft, P. W. and Freeling, M. (1994). "Genetic analysis of Rough sheath1 developmental mutants of maize." Genetics **136**(1): 295-311.

Bellaoui, M.; Pidkowich, M. S.; Samach, A.; Kushalappa, K.; Kohalmi, S. E.; Modrusan, Z.; Crosby, W. L. and Haughn, G. W. (2001). "The *Arabidopsis* BELL1 and KNOX TALE homeodomain proteins interact through a domain conserved between plants and animals." Plant Cell **13**(11): 2455-70.

Belmont, A. S.; Dietzel, S.; Nye, A. C.; Strukov, Y. G. and Tumbar, T. (1999). "Large-scale chromatin structure and function." Curr Opin Cell Biol **11**(3): 307-11.

Benyajati, C.; Mueller, L.; Xu, N.; Pappano, M.; Gao, J.; Mosammaparast, M.; Conklin, D.; Granok, H.; Craig, C. and Elgin, S. (1997). "Multiple isoforms of GAGA factor, a critical component of chromatin structure." Nucleic Acids Res **25**(16): 3345-53.

Bevilacqua, A.; Fiorenza, M. T. and Mangia, F. (2000). "A developmentally regulated GAGA box-binding factor and Sp1 are required for transcription of the hsp70.1 gene at the onset of mouse zygotic genome activation." Development **127**(7): 1541-51.

Bhat, K. M.; Farkas, G.; Karch, F.; Gyurkovics, H.; Gausz, J. and Schedl, P. (1996). "The GAGA factor is required in the early *Drosophila* embryo not only for transcriptional regulation but also for nuclear division." Development **122**(4): 1113-24.

Bradford, M. (1976). Anal. Biochem. **72**, 248.

Brockdorff, N. and Duthie, S. M. (1998). "X chromosome inactivation and the Xist gene." Cell Mol Life Sci **54**(1): 104-12.

Brown, J. L.; Mucci, D.; Whiteley, M.; Dirksen, M. L. and Kassis, J. A. (1998). "The *Drosophila* Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1." Mol Cell **1**(7): 1057-64.

Busturia, A.; Lloyd, A.; Bejarano, F.; Zavortink, M.; Xin, H. and Sakonju, S. (2001). "The MCP silencer of the *Drosophila* Abd-B gene requires both Pleiohomeotic and GAGA factor for the maintenance of repression." Development **128**(11): 2163-73.

Busturia, A.; Wightman, C. D. and Sakonju, S. (1997). "A silencer is required for maintenance of transcriptional repression throughout *Drosophila* development." Development **124**(21): 4343-50.

Cartwright, I. L. and Elgin, S. C. (1986). "Nucleosomal instability and induction of new upstream protein-DNA associations accompany activation of four small heat shock protein genes in *Drosophila* melanogaster." Mol Cell Biol **6**(3): 779-91.

Caspar, T. and Quail, P. H. (1993). "Promoter and leader regions involved in the expression of the *Arabidopsis* ferredoxin A gene." Plant J **3**(1): 161-74.

Castiglioni, P.; Pozzi, C.; Heun, M.; Terzi, V.; Muller, K. J.; Rohde, W. and Salamini, F. (1998). "An AFLP-based procedure for the efficient mapping of mutations and DNA probes in barley." Genetics **149**(4): 2039-56.

Chan, C. S.; Rastelli, L. and Pirrotta, V. (1994). "A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression." Embo J **13**(11): 2553-64.

Chao, Q.; Rothenberg, M.; Solano, R.; Roman, G.; Terzaghi, W. and Ecker, J. R. (1997). "Activation of the ethylene gas response pathway in *Arabidopsis* by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins." Cell **89**(7): 1133-44.

Chen, J. J.; Janssen, B. J.; Williams, A. and Sinha, N. (1997). "A gene fusion at a homeobox locus: alterations in leaf shape and implications for morphological evolution." Plant Cell **9**(8): 1289-304.

Chen, W.; Zollman, S.; Couderc, J. L. and Laski, F. A. (1995). "The BTB domain of bric a brac mediates dimerization in vitro." Mol Cell Biol **15**(6): 3424-9.

Cheng, X.; Kay, B. K. and Juliano, R. L. (1996). "Identification of a biologically significant DNA-binding peptide motif by use of a random phage display library." Gene **171**(1): 1-8.

Chua, Y. L.; Brown, A. P. and Gray, J. C. (2001). "Targeted histone acetylation and altered nuclease accessibility over short regions of the pea plastocyanin gene." Plant Cell **13**(3): 599-612.

Clifford, H.T. (1988)."Spikelet and floral morphology." " Grass systematics and evolution." Smithsonian Institution Press, Washington DC, London.

Cockell, M. and Gasser, S. M. (1999). "Nuclear compartments and gene regulation." Curr Opin Genet Dev **9**(2): 199-205.

Coen, E. S. and Meyerowitz, E. M. (1991). "The war of the whorls: genetic interactions controlling flower development." Nature **353**(6339): 31-7.

Costlow, N. and Lis, J. T. (1984). "High-resolution mapping of DNase Ł-hypersensitive sites of *Drosophila* heat shock genes in *Drosophila melanogaster* and *Saccharomyces cerevisiae.*" Mol Cell Biol **4**(9): 1853-63.

Croston, G. E.; Kerrigan, L. A.; Lira, L. M.; Marshak, D. R. and Kadonaga, J. T. (1991). "Sequence-specific antirepression of histone H1-mediated inhibition of basal RNA polymerase II transcription." Science **251**(4994): 643-9.

Di Cristina, M.; Sessa, G.; Dolan, L.; Linstead, P.; Baima, S.; Ruberti, I. and Morelli, G. (1996). "The *Arabidopsis* Athb-10 (GLABRA2) is an HD-Zip protein required for regulation of root hair development." Plant J **10**(3): 393-402.

Dietrich, R. A.; Radke, S. E. and Harada, J. J. (1992). "Downstream DNA sequences are required to activate a gene expressed in the root cortex of embryos and seedlings." Plant Cell **4**(11): 1371-82.

Dillon, N. and Sabbattini, P. (2000). "Functional gene expression domains: defining the functional unit of eukaryotic gene regulation." Bioessays **22**(7): 657-65.

Drews, G. N.; Bowman, J. L. and Meyerowitz, E. M. (1991). "Negative regulation of the *Arabidopsis* homeotic gene AGAMOUS by the APETALA2 product." Cell **65**(6): 991-1002.

Emili, A.; Greenblatt, J. and Ingles, C. J. (1994). "Species-specific interaction of the glutamine-rich activation domains of Sp1 with the TATA box-binding protein." Mol Cell Biol **14**(3): 1582-93.

Endrizzi, K.; Moussian, B.; Haecker, A.; Levin, J. Z. and Laux, T. (1996). "The SHOOT MERISTEMLESS gene is required for maintenance of undifferentiated cells in *Arabidopsis* shoot and floral meristems and acts at a different regulatory level than the meristem genes WUSCHEL and ZWILLE." Plant J **10**(6): 967-79.

Fang, R. X.; Nagy, F.; Sivasubramaniam, S. and Chua, N. H. (1989). "Multiple cis regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants." Plant Cell **1**(1): 141-50.

Farkas, G.; Gausz, J.; Galloni, M.; Reuter, G.; Gyurkovics, H. and Karch, F. (1994). "The Trithorax-like gene encodes the *Drosophila* GAGA factor." Nature **371**(6500): 806-8.

Felsenfeld, G.; Boyes, J.; Chung, J.; Clark, D. and Studitsky, V. (1996). "Chromatin structure and gene expression." Proc Natl Acad Sci U S A **93**(18): 9384-8.

Foster, T.; Yamaguchi, J.; Wong, B. C.; Veit, B. and Hake, S. (1999). "Gnarley1 is a dominant mutation in the knox4 homeobox gene affecting cell shape and identity." Plant Cell **11**(7): 1239-52.

Fowler, J. E.; Muehlbauer, G. J. and Freeling, M. (1996). "Mosaic analysis of the liguleless3 mutant phenotype in maize by coordinate suppression of mutator-insertion alleles."

Genetics **143**(1): 489-503.

Fraenkel, E. and Pabo, C. O. (1998). "Comparison of X-ray and NMR structures for the Antennapedia homeodomain- DNA complex." Nat Struct Biol **5**(8): 692-7.

Fried, M. and Crothers, D. M. (1981). "Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis." Nucleic Acids Res **9**(23): 6505-25.

Frisch, D. A.; Harris-Haller, L. W.; Yokubaitis, N. T.; Thomas, T. L.; Hardin, S. H. and Hall, T. C. (1995). "Complete sequence of the binary vector Bin 19." Plant Mol Biol **27**(2): 405-9.

Frischauf, A. M.; Lehrach, H.; Poustka, A. and Murray, N. (1983). "Lambda replacement vectors carrying polylinker sequences." J Mol Biol **170**(4): 827-42.

Frugis, G.; Giannino, D.; Mele, G.; Nicolodi, C.; Chiappetta, A.; Bitonti, M. B.; Innocenti, A. M.; Dewitte, W.; Van Onckelen, H. and Mariotti, D. (2001). "Overexpression of KNAT1 in lettuce shifts leaf determinate growth to a shoot-like indeterminate growth associated with an accumulation of isopentenyl-type cytokinins." Plant Physiol **126**(4): 1370-80.

Garner, M. M. and Revzin, A. (1981). "A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system." Nucleic Acids Res **9**(13): 3047-60.

Gidekel, M.; Jimenez, B. and Herrera-Estrella, L. (1996). "The first intron of the *Arabidopsis* thaliana gene coding for elongation factor 1 beta contains an enhancer-like element." Gene **170**(2): 201-6.

Goodrich, J.; Puangsomlee, P.; Martin, M.; Long, D.; Meyerowitz, E. M. and Coupland, G. (1997). "A Polycomb-group gene regulates homeotic gene expression in *Arabidopsis*." Nature **386**(6620): 44-51.

Granok, H.; Leibovitch, B. A.; Shaffer, C. D. and Elgin, S. C. (1995). "Chromatin. Ga-ga over GAGA factor." Curr Biol **5**(3): 238-41.

Grasser, K. D. (1995). "Plant chromosomal high mobility group (HMG) proteins." Plant J **7**(2): 185-92.

Grossniklaus, U.; Vielle-Calzada, J. P.; Hoeppner, M. A. and Gagliano, W. B. (1998). "Maternal control of embryogenesis by MEDEA, a polycomb group gene in *Arabidopsis*." Science **280**(5362): 446-50.

Hebbes, T. R.; Clayton, A. L.; Thorne, A. W. and Crane-Robinson, C. (1994). "Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain." Embo J **13**(8): 1823-30.

Henikoff, S. (1990). "Position-effect variegation after 60 years." Trends Genet **6**(12): 422-6.

Henikoff, S. (2000). "Heterochromatin function in complex genomes." Biochim Biophys Acta **1470**(1): 1-8.

Hertzberg, M. and Olsson, O. (1998). "Molecular characterisation of a novel plant homeobox

gene expressed in the maturing xylem zone of Populus tremula x tremuloides." Plant J **16**(3): 285-95.

Hodgson, J. W.; Argiropoulos, B. and Brock, H. W. (2001). "Site-specific recognition of a 70-base-pair element containing d(GA)(n) repeats mediates bithoraxoid polycomb group response element-dependent silencing." Mol Cell Biol **21**(14): 4528-43.

Ish-Horowicz, D.; Pinchin, S. M.; Ingham, P. W. and Gyurkovics, H. G. (1989). "Autocatalytic ftz activation and metameric instability induced by ectopic ftz expression." Cell **57**(2): 223-32.

Jackson, D.; Veit, B. and Hake, S. (1994). "Expression of the maize Knotted-1 related homeobox genes in shoot apical meristem predicts patterns of morphogenesis in vegetative shoot." Development **120**: 405-423.

Jofuku, K. D.; den Boer, B. G.; Van Montagu, M. and Okamuro, J. K. (1994). "Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2." Plant Cell **6**(9): 1211-25.

Johannesson, H.; Wang, Y. and Engstrom, P. (2001). "DNA-binding and dimerization preferences of *Arabidopsis* homeodomain- leucine zipper transcription factors in vitro." Plant Mol Biol **45**(1):63-73.

Jones, R. S. and Gelbart, W. M. (1993). "The *Drosophila* Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax." Mol Cell Biol **13**(10): 6357-66.

Katsani, K. R.; Hajibagheri, M. A. and Verrijzer, C. P. (1999). "Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology." Embo J **18**(3): 698-708.

Kehle, J.; Beuchle, D.; Treuheit, S.; Christen, B.; Kennison, J. A.; Bienz, M. and Muller, J. (1998). "dMi-2, a hunchback-interacting protein that functions in polycomb repression." Science **282**(5395): 1897-900.

Kennison, J. A. (1995). "The Polycomb and trithorax group proteins of *Drosophila*: trans-regulators of homeotic gene function." Annu Rev Genet **29**: 289-303.

Kerstetter, R. A.; Laudencia-Chingcuanco, D.; Smith, L. G. and Hake, S. (1997). "Loss-of-function mutations in the maize homeobox gene, knotted1, are defective in shoot meristem maintenance." Development **124**(16): 3045-54.

Kissinger, C. R.; Liu, B. S.; Martin-Blanco, E.; Kornberg, T. B. and Pabo, C. O. (1990). "Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions." Cell **63**(3): 579-90.

Kopriva, S.; Cossu, R. and Bauwe, H. (1995). "Alternative splicing results in two different transcripts for H-protein of the glycine cleavage system in the C4 species Flaveria trinervia." Plant J **8**(3): 435-41.

Kuhn, R. M.; Caspar, T.; Dehesh, K. and Quail, P. H. (1993). "DNA binding factor GT-2 from *Arabidopsis*." Plant Mol Biol **23**(2): 337-48.

Kuroiwa, A.; Hafen, E. and Gehring, W. J. (1984). "Cloning and transcriptional analysis of the segmentation gene fushi tarazu of *Drosophila*." Cell **37**(3): 825-31.

Kyba, M. and Brock, H. W. (1998). "The *Drosophila* polycomb group protein Psc contacts ph and Pc through specific conserved domains." Mol Cell Biol **18**(5): 2712-20.

Lee, T. I. and Young, R. A. (2000). "Transcription of eukaryotic protein-coding genes." Annu Rev Genet **34**: 77-137.

Li, J.; Liang, V. C.; Sedgwick, T.; Wong, J. and Shi, Y. B. (1998). "Unique organization and involvement of GAGA factors in transcriptional regulation of the Xenopus stromelysin-3 gene." Nucleic Acids Res **26**(12): 3018-25.

Lin, J. and Mueller, K.J. (2001). "Structure and development of epiphylly in KNOX-transgenic tobacco." Planta. *Published online.*

Lis, J. and Wu, C. (1993). "Protein traffic on the heat shock promoter: parking, stalling, and trucking along." Cell **74**(1): 1-4.

Locke, J.; Kotarski, M. A. and Tartof, K. D. (1988). "Dosage-dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effect." Genetics **120**(1): 181-98.

Long, J. A.; Moan, E. I.; Medford, J. I. and Barton, M. K. (1996). "A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of *Arabidopsis*." Nature **379**(6560): 66-9.

Lu, Q.; Wallrath, L. L.; Granok, H. and Elgin, S. C. (1993). "(CT)n (GA)n repeats and heat shock elements have distinct roles in chromatin structure and transcriptional activation of the *Drosophila* hsp26 gene." Mol Cell Biol **13**(5): 2802-14.

Lubberstedt, T.; Bolle, C. E.; Sopory, S.; Flieger, K.; Herrmann, R. G. and Oelmuller, R. (1994). "Promoters from genes for plastid proteins possess regions with different sensitivities toward red and blue light." Plant Physiol **104**(3): 997-1006.

Mahon, K. A.; Westphal, H. and Gruss, P. (1988). "Expression of homeobox gene Hox 1.1 during mouse embryogenesis." Development **104**(Suppl): 187-95.

Marillonnet, S. and Wessler, S. R. (1997). "Retrotransposon insertion into the maize waxy gene results in tissue- specific RNA processing." Plant Cell **9**(6): 967-78.

Masucci, J. D.; Rerie, W. G.; Foreman, D. R.; Zhang, M.; Galway, M. E.; Marks, M. D. and Schiefelbein, J. W. (1996). "The homeobox gene GLABRA2 is required for position-dependent cell differentiation in the root epidermis of *Arabidopsis* thaliana." Development **122**(4): 1253-60.

Matsumoto, K.; Wassarman, K. M. and Wolffe, A. P. (1998). "Nuclear history of a pre-mRNA determines the translational activity of cytoplasmic mRNA." Embo J **17**(7): 2107-21.

Matsuoka, M.; Ichikawa, H.; Saito, A.; Tada, Y.; Fujimura, T. and Kano-Murakami, Y. (1993). "Expression of a rice homeobox gene causes altered morphology of transgenic plants." Plant Cell **5**(9): 1039-48.

Meisel, L. and Lam, E. (1996). "The conserved ELK-homeodomain of KNOTTED-1 contains two regions that signal nuclear localization." Plant Mol Biol **30**(1): 1-14.

Mikami, K.; Sakamoto, A. and Iwabuchi, M. (1994). "The HBP-1 family of wheat basic/leucine zipper proteins interacts with overlapping cis-acting hexamer motifs of plant histone genes." J Biol Chem **269**(13): 9974-85.

Milot, E.; Strouboulis, J.; Trimborn, T.; Wijgerde, M.; de Boer, E.; Langeveld, A.; Tan-Un, K.; Vergeer, W.; Yannoutsos, N.; Grosveld, F. and Fraser, P. (1996). "Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription." Cell **87**(1): 105-14.

Mizukami, Y. and Ma, H. (1992). "Ectopic expression of the floral homeotic gene AGAMOUS in transgenic *Arabidopsis* plants alters floral organ identity." Cell **71**(1): 119-31.

Muller, J.; Wang, Y.; Franzen, R.; Santi, L.; Salamini, F. and Rohde, W. (2001). "In vitro interactions between barley TALE homeodomain proteins suggest a role for protein-protein associations in the regulation of Knox gene function." Plant J **27**(1): 13-23.

Muller, K. J.; Romano, N.; Gerstner, O.; Garcia-Maroto, F.; Pozzi, C.; Salamini, F. and Rohde, W. (1995). "The barley Hooded mutation caused by a duplication in a homeobox gene intron." Nature **374**(6524): 727-30.

Muller, M.; Hagstrom, K.; Gyurkovics, H.; Pirrotta, V. and Schedl, P. (1999). "The mcp element from the *Drosophila* melanogaster bithorax complex mediates long-distance regulatory interactions." Genetics **153**(3): 1333-56.

Munemitsu, S.; Innis, M. A.; Clark, R.; McCormick, F.; Ullrich, A. and Polakis, P. (1990). "Molecular cloning and expression of a G25K cDNA, the human homolog of the yeast cell cycle gene CDC42." Mol Cell Biol **10**(11): 5977-82.

Murashige, T. and Skoog, F. (1962). "A revised medium for rapid groth and bioassays with tobacco tissue cultures." Physiol. Plant **15**: 473-497.

Nardelli, J.; Gibson, T. J.; Vesque, C. and Charnay, P. (1991). "Base sequence discrimination by zinc-finger DNA-binding domains." Nature **349**(6305): 175-8.

Ni, M.; Dehesh, K.; Tepperman, J. M. and Quail, P. H. (1996). "GT-2: in vivo transcriptional activation activity and definition of novel twin DNA binding domains with reciprocal target sequence selectivity." Plant Cell **8**(6): 1041-59.

Okada, M. and Hirose, S. (1998). "Chromatin remodeling mediated by *Drosophila* GAGA factor and ISWI activates fushi tarazu gene transcription in vitro." Mol Cell Biol **18**(5): 2455-61.

Parnis, A.; Cohen, O.; Gutfinger, T.; Hareven, D.; Zamir, D. and Lifschitz, E. (1997). "The dominant developmental mutants of tomato, Mouse-ear and Curl, are associated with distinct modes of abnormal transcriptional regulation of a Knotted gene." Plant Cell **9**(12): 2143-58.

Paro, R. and Harte, J. (1996). "The role of *Polycomb* group and *trithorax* group chromatin complexes in the manteinance of determined cell states." " Epigenetics mechanisms of gene regulation". *Cold Spring Harbor Laboratory Press.*

Pascal, E. and Tjian, R. (1991). "Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism." Genes Dev **5**(9): 1646-56.

Paule, M. R. and White, R. J. (2000). "Survey and summary: transcription by RNA polymerases I and III." Nucleic Acids Res **28**(6): 1283-98.

Pellegrini, L.; Tan, S. and Richmond, T. J. (1995). "Structure of serum response factor core bound to DNA." Nature **376**(6540): 490-8.

Pinto, M. and Lobe, C. G. (1996). "Products of the grg (Groucho-related gene) family can dimerize through the amino-terminal Q domain." J Biol Chem **271**(51): 33026-31.

Pirrotta, V. (1999). "Polycomb silencing and the maintenance of stable chromatin states." Results Probl Cell Differ **25**: 205-28.

Pisarchik, A. V. and Kartel, N. A. (2000). "[Simple repeating sequences and gene expression]." Mol Biol (Mosk) **34**(3): 357-62.

Plesch, G.; Stormann, K.; Torres, J. T.; Walden, R. and Somssich, I. E. (1997). "Developmental and auxin-induced expression of the *Arabidopsis* prha homeobox gene." Plant J **12**(3): 635-47.

Pozzi, C.; Faccioli, P.; Terzi, V.; Stanca, A. M.; Cerioli, S.; Castiglioni, P.; Fink, R.; Capone, R.; Muller, K. J.; Bossinger, G.; Rohde, W. and Salamini, F. (2000). "Genetics of mutations affecting the development of a barley floral bract." Genetics **154**(3): 1335-46.

Rabadan-Diehl, C.; Lolait, S. and Aguilera, G. (2000). "Isolation and characterization of the promoter region of the rat vasopressin V1b receptor gene." J Neuroendocrinol **12**(5): 437-44.

Raff, J. W.; Kellum, R. and Alberts, B. (1994). "The *Drosophila* GAGA transcription factor is associated with specific regions of heterochromatin throughout the cell cycle." Embo J **13**(24): 5977-83.

Read, C. M. and Driscoll, P. C. (1997). "GAGA over the nucleosome." Nat Struct Biol **4**(2): 87-9.

Reiser, L.; Sanchez-Baracaldo, P. and Hake, S. (2000). "Knots in the family tree: evolutionary relationships and functions of knox homeobox genes." Plant Mol Biol **42**(1): 151-66.

Riechmann, J. L.; Heard, J.; Martin, G.; Reuber, L.; Jiang, C.; Keddie, J.; Adam, L.; Pineda, O.; Ratcliffe, O. J.; Samaha, R. R.; Creelman, R.; Pilgrim, M.; Broun, P.; Zhang, J. Z.; Ghandehari, D.; Sherman, B. K. and Yu, G. (2000). "*Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes." Science **290**(5499): 2105-10.

Riechmann, J. L. and Meyerowitz, E. M. (1998). "The AP2/EREBP family of plant transcription factors." Biol Chem **379**(6): 633-46.

Sainz, M. B.; Goff, S. A. and Chandler, V. L. (1997). "Extensive mutagenesis of a transcriptional activation domain identifies single hydrophobic and acidic amino acids important for activation in vivo." Mol Cell Biol **17**(1): 115-22.

Sainz, M. B.; Grotewold, E. and Chandler, V. L. (1997). "Evidence for direct activation of an anthocyanin promoter by the maize C1 protein and comparison of DNA binding by related Myb domain proteins." Plant Cell **9**(4): 611-25.

Sakamoto, A.; Minami, M.; Huh, G. H. and Iwabuchi, M. (1993). "The putative zinc-finger protein WZF1 interacts with a cis-acting element of wheat histone genes." Eur J Biochem **217**(3): 1049-56.

Sambrook, J.; Fritsch, E.F. and Maniatis, T. (1989). "Molecular cloning". *Cold Spring Harbor Laboratory Press.*

Sato, N.; Ohshima, K.; Watanabe, A.; Ohta, N.; Nishiyama, Y.; Joyard, J. and Douce, R. (1998). "Molecular characterization of the PEND protein, a novel bZIP protein present in the envelope membrane that is the site of nucleoid replication in developing plastids." Plant Cell **10**(5): 859-72.

Schindler, U.; Beckmann, H. and Cashmore, A. R. (1993). "HAT3.1, a novel *Arabidopsis* homeodomain protein containing a conserved cysteine-rich region." Plant J **4**(1): 137-50.

Schneeberger, R.; Tsiantis, M.; Freeling, M. and Langdale, J. A. (1998). "The rough sheath2 gene negatively regulates homeobox gene expression during maize leaf development." Development **125**(15): 2857-65.

Scott, M. P. and Weiner, A. J. (1984). "Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of *Drosophila*." Proc Natl Acad Sci U S A **81**(13): 4115-9.

Sentoku, N.; Sato, Y.; Kurata, N.; Ito, Y.; Kitano, H. and Matsuoka, M. (1999). "Regional expression of the rice KN1-type homeobox gene family during embryo, shoot, and flower development." Plant Cell **11**(9): 1651-64.

Shao, Z.; Raible, F.; Mollaaghababa, R.; Guyon, J. R.; Wu, C. T.; Bender, W. and Kingston, R. E. (1999). "Stabilization of chromatin structure by PRC1, a Polycomb complex." Cell **98**(1):37-46.

Sharp, G. J.; Brayan, J. G.; Dilli, S.; Haddad, P. R. and Desmarchelier, J. M. (1988). "Extraction, clean-up and chromatographic determination of organophosphate, pyrethroid and carbamate insecticides in grain and grain products. A review." Analyst **113**(10): 1493-507.

Shimell, M. J.; Peterson, A. J.; Burr, J.; Simon, J. A. and O'Connor, M. B. (2000). "Functional analysis of repressor binding sites in the iab-2 regulatory region of the abdominal-A homeotic gene." Dev Biol **218**(1): 38-52.

Shopland, L. S.; Hirayoshi, K.; Fernandes, M. and Lis, J. T. (1995). "HSF access to heat shock elements in vivo depends critically on promoter architecture defined by GAGA factor, TFIID, and RNA polymerase II binding sites." Genes Dev **9**(22): 2756-69.

Sieburth, L. E. and Meyerowitz, E. M. (1997). "Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically." Plant Cell **9**(3): 355-65.

Sinha, N. R.; Williams, R. E. and Hake, S. (1993). "Overexpression of the maize homeo box gene, KNOTTED-1, causes a switch from determinate to indeterminate cell fates." Genes Dev **7**(5): 787-95.

Spena, A. and Salamini, F. (1995). "Genetic tagging of cells and cell layers for studies of plant development." Methods Cell Biol **49**: 331-54.

Spiker, S.; Murray, M. G. and Thompson, W. F. (1983). "DNase I sensitivity of transcriptionally active genes in intact nuclei and isolated chromatin of plants." Proc Natl Acad Sci U S A **80**(3): 815-9.

Stott, K.; Blackburn, J. M.; Butler, P. J. and Perutz, M. (1995). "Incorporation of glutamine repeats makes protein oligomerize: implications for neurodegenerative diseases." Proc Natl Acad Sci U S A **92**(14): 6509-13.

Strahl, B. D. and Allis, C. D. (2000). "The language of covalent histone modifications." Nature **403**(6765): 41-5.

Strom, A. C.; Forsberg, M.; Lillhager, P. and Westin, G. (1996). "The transcription factors Sp1 and Oct-1 interact physically to regulate human U2 snRNA gene expression." Nucleic Acids Res **24**(11): 1981-6.

Strutt, H.; Cavalli, G. and Paro, R. (1997). "Co-localization of Polycomb protein and GAGA factor on regulatory elements responsible for the maintenance of homeotic gene expression." Embo J **16**(12): 3621-32.

Thoma, S.; Hecht, U.; Kippers, A.; Botella, J.; De Vries, S. and Somerville, C. (1994). "Tissue-specific expression of a gene encoding a cell wall-localized lipid transfer protein from *Arabidopsis*." Plant Physiol **105**(1): 35-45.

Tie, F.; Furuyama, T.; Prasad-Sinha, J.; Jane, E. and Harte, P. J. (2001). "The *Drosophila* Polycomb Group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3." Development **128**(2): 275-86.

Topfer, R.; Matzeit, V.; Gronenborn, B.; Schell, J. and Steinbiss, H. H. (1987). "A set of plant expression vectors for transcriptional and translational fusions." Nucleic Acids Res **15**(14): 5890.

Tsiantis, M.; Schneeberger, R.; Golz, J. F.; Freeling, M. and Langdale, J. A. (1999). "The maize *rough sheath2* gene and leaf development programs in monocot and dicot plants." Science **284**(5411): 154-6.

Tsukiyama, T.; Becker, P. B. and Wu, C. (1994). "ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor." Nature **367**(6463):525-32.

Tsukiyama, T. and Wu, C. (1995). "Purification and properties of an ATP-dependent nucleosome remodeling factor." Cell **83**(6): 1011-20.

Tsukiyama, T. and Wu, C. (1996). "Purification of GAGA factor of *Drosophila* and its role in nucleosome disruption." Methods Enzymol **274**: 291-9.

Udvardy, A.; Maine, E. and Schedl, P. (1985). "The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains." J Mol Biol **185**(2): 341-58.

Ulmasov, T.; Hagen, G. and Guilfoyle, T. J. (1997). "ARF1, a transcription factor that binds to auxin response elements." Science **276**(5320): 1865-8.

Varagona, M. J. and Raikhel, N. V. (1994). "The basic domain in the bZIP regulatory protein Opaque2 serves two independent functions: DNA binding and nuclear localization." Plant J **5**(2): 207-14.

Vollbrecht, E.; Veit, B.; Sinha, N. and Hake, S. (1991). "The developmental gene Knotted-1 is a member of a maize homeobox gene family." Nature **350**(6315): 241-3.

Wallrath, L. L. (1998). "Unfolding the mysteries of heterochromatin." Curr Opin Genet Dev **8**(2): 147-53.

Wang, Y. (2001). "Expression analysis of barley KNOX homeobox genes and characterisation of putative upstream regulators of *Bkn3*, the *Hooded* gene." *Ph D. thesis* , University of Cologne.

Waterborg, J. H. (1990). "Sequence analysis of acetylation and methylation in two histone H3 variants of alfalfa." J Biol Chem **265**(28): 17157-61.

Weigel, D. and Doerner, P. (1996). "Cell-cell interactions: taking cues from the neighbors." Curr Biol **6**(1): 10-2.

Wilde, R. J.; Cooke, S. E.; Brammar, W. J. and Schuch, W. (1994). "Control of gene expression in plant cells using a 434:VP16 chimeric protein." Plant Mol Biol **24**(2): 381-8.

Wilkins, R. C. and Lis, J. T. (1999). "DNA distortion and multimerization: novel functions of the glutamine- rich domain of GAGA factor." J Mol Biol **285**(2): 515-25.

Williams-Carrier, R. E.; Lie, Y. S.; Hake, S. and Lemaux, P. G. (1997). "Ectopic expression of the maize kn1 gene phenocopies the Hooded mutant of barley." Development **124**(19): 3737-45.

Wu, C. (1980). "The 5' ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I." Nature **286**(5776): 854-60.

Yamamoto, S.; Suzuki, K. and Shinshi, H. (1999). "Elicitor-responsive, ethylene-independent activation of GCC box- mediated transcription that is regulated by both protein phosphorylation and dephosphorylation in cultured tobacco cells." Plant J **20**(5): 571-9.

Yanofsky, M. F.; Ma, H.; Bowman, J. L.; Drews, G. N.; Feldmann, K. A. and Meyerowitz, E. M. (1990). "The protein encoded by the *Arabidopsis* homeotic gene agamous resembles transcription factors." Nature **346**(6279): 35-9.

Zollman, S.; Godt, D.; Prive, G. G.; Couderc, J. L. and Laski, F. A. (1994). "The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in *Drosophila*." Proc Natl Acad Sci U S A **91**(22): 10717-21.

Appendix

Oligonucleotides

305bp element

name	Sequence 5'→ 3'
305H	TTGTGAAGAAGCTTCTGGGGAGTTCTGACC
305RV	CTTGGACTCGCGATATCAAGCATCCAACGT
305SA1	CTTGGAGTCGACTTGGAAAGCATCCAAACG
305a	TTGTGTGTGTTCTGGGGAGTTCTGACCAGT
305b	TCACATCAAATAATGGTGGCTGTAGGCTGT
305c	ACAGTCAAGTGCTTGCCACAGCTACAGCCA
305d	CATGGTATCTAGTCTAGCTACATGCCTACA
305e	GACTGCTAGCTCCATCTACTAGTGTAGGCA
305f	ACTTGGACTCGCCTTGGAAAGCATCCAACG
401	AAAACATCTACTAGTGTAGGCATGTAGCTA
402	AAAAATGGTATCTAGTCTAGCTACATGCCT
403	AAAATAGATACCATGTTGCTGTATTTTGCG
404	AAAAGAGATAAGGCTGTCGCAAAATACAG
405	AAAAAGCCTTATCTCTGGTCTCTCTCTCTC
406	AAAACATTATTCATGAGAGAGAGAGAGAG
407	AAAAATGAATAATGGCGGTCAAGAGACGT
408	AAAATTGGAAAGCATCCAACGTCTCTTGAC
261	TCTCTGGTCTCTCTCTCTCTCTCATG
262	AGAGACCAGAGAGAGAGAGAGAGAGTAC
161	TCTCTCTCTCTCTCTC
162	GAGAGAGAGAGAGA
171	TCTCTGGTCTCTCTCTC
172	AGAGACCAGAGAGAGAG

Bkn3

name	Sequence 5'→ 3'
KS1	ACCGCGCCGCCCGGGGTAAGCCACGGCG
KS2	GTCACCGGTAGGAGACGAGGCCCGGGCGAG

BBR

name	Sequence 5'→ 3'
FBBR1	AAAGAAGGATCCATTTCTCGTCCCAATCTCG ʒ
FBBR2	AGCGTTGGATCCTACAAGGTGATCAACATA
BBRVP1	AAAGAAGTCCGGATTTCTCGTCCCAATCTCG ʒ
BBRVP2	AGCGTTCCGGACAAGGCGATCAACATA
77LU2	GGGTCGTCAGCTAAGGTTC
77LU5	CATTTCCACAGGCTTAAGGGG
PBBR1	GCTCCGTATCGTAGTCTATAG
PBBR2	GAGGTTTCCTTTCATCGTCTCG

Markers

name	Sequence 5'→ 3'
hph1	ATCACGCCATGTAGTGTATTGACC
hph2	CATTGTCCGTCAGGACATTGTTGG
KA1	TAGAAGGCGGCGGTGGAATCGAAATCT
KA2	AATCTGCACCGGATCTGGATCGTTTCG

cDNAs

BEIL

```
   1  ACTCACTATA GGGCGAATTG GAGCTCCACC GCGGTGGCGG CCGCTCTAGA
  51  ACTAGTGGAT CCCCCGGGCT GCAGGAATTC CGATTGAGGC TGAGGAGCTG
 101  ACCCGGCGAA TGTGGAAAGA TAAGGTCAGG CTCAAGAGGA TCAAGGAGAA
 151  GCAGCAGAGG CTTGCTTTGG AGCAGGCGGA ACTGGAGAAG TCTAATCCAA
 201  AGAAGTTGTC TGATCTAGCC CTTCGCAAGA AGATGGCAAG AGCCCAGGAT
 251  GGGATTTTGA AGTACATGCT CAAGTTGATG GAAGTATGCA ATGCGCAGGG
 301  TTTTGTTTAT GGGATCATTC CTGATAAAGG GAAGCCTGTC AGTGGAGCAT
 351  CGGAAAACAT TAGAGCTTGG TGGAAGGAGA AGGTTAAGTT TGATAAGAAT
 401  GGGCCAGCAG CAATTGCAAA ATATGAGGTT GAGAACTCTC TGTTGGTTAA
 451  TGGTCAGAGC AGTGGGACCA TGAATCAATA TAGCTTGATG GATCTCCAAG
 501  ATGGTACCCT GGGCTCATTG CTTTCTGCAT TGATGCAGCA TTGCAGCCCT
 551  CAGCAGCGCA AGTACCCACT GGATAAGGGT ATTCCGCCCC CGTGGTGGCC
 601  ATCAGGGAAT GAGGAGTGGT GGATTGCTTT AGGCCTTCCG AAGGGTAAAA
 651  CACCTCCATA CAAAAAACCT CATGATCTTA AGAAGTTTTG GAAGGTTGGT
 701  GTGCTGACGG CTGTGATCAA ACACATGTCT CCGCATTTTG ATAAGATAAG
 751  ATATCATGTA CGGAAGTCAA AGTGCTTGCA GGACAAAATG ACTGCAAAAG
 801  AGAGCTTGAT TTGGCTGGTT GTTTTGCAAA GAGAGGAGTA TGCTCACAGT
 851  ATTGATAACG GTGTATCAGA TACTCACCAT TGTGACCTAG GGGACAAAAA
 901  TGGGAGTTCA TACAGCAGCT GTGATGAGTA TGATGTTGAC TGTATGGAGG
 951  AGCCTCCTCA GTCTACAATA TCCAAAGACG ATGTGGGAGT TCGTCAGCCA
1001  ACTGTGCACA TCAGAGAAGA GAATGCCTCA AGTAGTGGGA ACAAAAAACG
1051  TCATGATAAA CGCTCTACTC AAACGCTGCC TAGTACTAAG GAAACTAAAA
1101  AACCACTAAA GCGAAGAAAA CATATCGGAC AGTTTTCCGT TGATGGGTCT
1151  GAGGTTGAAG GAACACAGAG AAATGATAAC ACGCCAGAGG TTTTGAGCAA
1201  CGCAATTCCT GATATGAATA GCAATCAGAT GGAGTTGGTC TGTGTTGCTG
1251  ACCTGTTGAC AAGCTTCAAT CATGTCAGTA CAAATGGAGG AGCTTTACAA
1301  CATCAAGGAG ATGTTCAAGG GAACTTTGTA CCCCCTGGTG TTGTTGTTAA
1351  TAATTACAGC CAGGCTGCAA ATATTGCTCC TTCCAGCATC TATATGGCCG
1401  ACCAGCCATT GGCTTCTGCA AGTAATGATT ATGCAAACTC CTGGCCTGGA
1451  AATACTTTTC AACCAAACGT TGGTCTTGGA TCTATTGGCT TTAGTTCTTC
1501  TTCACATGAT TACCAGTCTT CTTCTGCTGC AAAACACTCA TTGCCACTAT
1551  CTACGGATAA CCATGTGCCT GCCATGGGAA CAGGAGGTTT GAACAGTTCT
```

BBR

```
   1  ATGCAATGGT GGGGATAGCC TGACCGTCCA GAGCCCGAAA GAAGAGCACA
  51  TTTCTCGTCC CAATCTCGGT TTTCTTGGAT TTCGATTCGT TTGTTGGGCC
 101  GGCCGAGATG GACGACGACG GCAGCTTGAG CATTCGGAAT TGGGGCTTCT
 151  ACGAGACGAT GAAAGGAAAC CTCGGCCTGC AGCTGATGCC ATCTGTGACC
 201  GGCGGCCACC GGGACACGAA GCCGCTGCTC CCCAACGGTA CCTTCTTGCA
 251  GCACCACACC CCCCCGCACC ACCCGCCACA TTCGCACCAC CCCCGCGACT
 301  ATGGTAACGG CGAACCCTCT GGTGGCATGC CCGCCGAGCC GCCGGCTATT
 351  CACATGGACT TTGTGCGCAA TGAGGCCTGG ATGCACCCCT CGCAGCATCA
 401  ACATCAGCAT CAGCATCAGC ACCAGCATCA ACATCAGCAC CAGCATCAAC
 451  TTCAGCACCA GCATCAACAT CAACATTCCC GTGAGTTGAA GGTCCTTAAT
 501  GCTGTTCCTG TTGGGCCTGC TCCACACATT GGACATCCTG GACATGCTGT
 551  GCATCACCAC CCTACAGGTT TTGGGATGAT GCCAGATGCG CGTGGTGCGC
 601  ACACTCTCCA GATGATGCAG CCACAGGAGC CTCCTGTGCC TGATGAGGAA
 651  AAAATTACCC CACCGCTGGT TGAAGATCAT TCTGTGGTCG GAAGCAAGCC
 701  TCCTGTGAAG AAGAGGCAGC AGGGTCGTCA GCCTAAGGTT CCGAAGCCGA
 751  AGAAGCCCAA GAAGGATGCT ACCCCAGGGG AAGATGGGGC ACCCAAGGCC
 801  CGTGCACCCC GAAGCAGGGG TCCCCTTAAG CCTGTGGAAA TGGTAATTAA
 851  TGGTATTGAT TTTGACATTT CAAGGATACC AACACCTGTG TGCTCATGCA
 901  CTGGAGCTCC CCAGCAATGC TACCGGTGGG GTGCAGGTGG CTGGCAGTCT
 951  GCATGCTGCA CAACTTCTAT TTCGACATAT CCGCTGCCAA TGAACACAAA
1001  GCGCCGGGGT GCACGTATTG CTGGGAGGAA AATGAGCCAA GGTGCATTCA
1051  AGAAGGTTCT TGAGAAGCTA GCTGGTGAAG GGTACAACCT TAATAATCCA
1101  ATTGACTTGA AGACCTTCTG GGCAAAGCAT GGCACGAACA AGTTTGTAAC
1151  AATCAGGTAA AAGCCATGCT ATGTTGATCG CCTTGTAGGG TCCAACGCTG
1201  CAGTTTTGCC TGAGTTATTG CCTGTACCTG CTCTTGCCAA TCTGCAGATG
1251  TCTTTAGAAG TAGCTTGTAG CCATGTTCGG TTTCCTGGAC TTAATTTACA
1301  TGCCTATTTC GAGCTTCAGC TGTGAAGTAG GAAGTCTGTC TGTATCTGTC
1351  AATTTAGAAG TTGTAGCGTA ATGGCAACAG TTTTTCTAAT TAGTTAGCAT
1401  TTAATGCTTC AATGTCTGTT ATGACCTGGA TGTTTATGGC TTTATGTAGA
1451  GCTTTTCTCC ATTGAAAGTT GAAAGATAAT CACCACCTCT TCTAGGGCTA
```

BGRF

```
   1  ACTATAGGGC GAATTGGAGC TCCACCGCGG TGGCGGCCGC TCTAGAACTA
  51  GTGGATCCCC CGGGCTGCAG GAATTCCGGG ACCGGACCGG AGCCAAGCAG
 101  CAGCCGCAGC CGCAGCCGCA GCAGAGGAGA GAGAGAGGGA GGGAGAAGCA
 151  TATATGGCGA TGCCCTTTGC CTCCCTGTCG CCGGCAGCCG ACCACCACCG
 201  CTCCTCCCCC ATCTTCCCCT TCTGCCGCTC CTCCCCTCTC TACTCGGTAG
 251  GGGAGGAGGC GGCGCATCAG CATCCTCATC CTCAGCAGCA GCAGCAGCAG
 301  CACGCGATGA GCGGCGCGCG GTGGGCGGCG AGGCCGGCGC CCTTCACGGC
 351  GGCGCAGTAC GAGGAGCTGG AGCAGCAGGC GCTCATCTAC AAGTACCTCG
 401  TCGCCGGCGT CCCCGTCCCG CAGGACCTCC TCCTCCCCAT CCGCCGCGGC
 451  TTCGAGACCC TCGCCTCGCG CTTCTACCAC CACCACGCCC TTGGGTACGG
 501  GTCCTACTTC GGGAAGAAGC TGGATCCGGA GCCGGGGCGG TGCCGGCGGA
 551  CGGACGGCAA GAAGTGGCGG TGCTCCAAGG AGGCCGCTCA GGACTCCAAG
 601  TACTGCGAGC GCCACATGCA CCGCGGCCGC AACCGTTCAN GAAAGCCTGT
 651  GGAAACGCAG CTCGTCGCCA GCTCCCACTC CCAGTCCCAG CAGCACGCCA
 701  CCGCCGCCTT CCACAACCAC TCGCCGTATC CGGCGATCGC CACTGGCGGT
 751  GGCTCCTTCG CCCTGGGGTC TGCTCAGCTG CACATGGACA CTGCTGCGCC
 801  TTACGCGACG ACCGCCGGTG CTGCCGGAAA CAAAGATTTC AGGTATTCTG
 851  CCTATGGAGT GAGGACGTCG GCGATCGAGG AGCACAACCA GTTCATCACC
 901  GCGGCCATGG ACACCGCCAT GGACAACTAC TCGTGGCGCC TGATGCCGTC
 951  CCAGGCCTCG GCATTCTCGC TCTCCAGCTA CCCCATGCTG GGCACGCTGA
1001  GCGACCTGGA CCAGAGCGCG ATCTGCTCGC TGGCCAAGAC TGAGAGGGAG
1051  CCACTGTCCT TCTTCGGCGG CGGCGGCGAC TTCGACGACG ACTCGGCTGC
1101  GGTAAAGCAG GAGAACCAGA CGCTGCGGCC CTTCTTCGAC GAGTGGCCCA
1151  AGGACAGGGA CTCGTGGCCG GAGCTGCAAG ACCACGACGC CAACAACAAC
1201  AGCAACGCCT TCTCAGCCAC CAAGCTGTCC ATCTCCATGC CGGTCACCAG
1251  CTCCGACTTC TCTGGCACCA CCGCCGGCTC CCGCTCGCCC AACGGTATAT
1301  ACTCCCGGTG AACGGCGTCG GCCGGCCTGA TCTCTGCTGA TTTGCCGTGG
1351  TCACGACGGA CGTCCTCAAA TCATCACAGA TGAGCGAACC GGCCGACCCG
1401  ATCGAATGTG TCTGTGAGCC GACTGCAGCT TGCTTGCTCA TTTTGTATGG
1451  ATCGTCGTGC AGCAGGAACG AAACACTACT CCTTTAATTT CCTTTCTTTA
1501  ATTTCACAAC GTTTTTTCTG GGTTTTGCCG TGTATCGGCC GGAACTGTAC
1551  TACCAAGTTT TCTATAGCCT CGATGGTCAT GCACGACATC GTTGACTGTT
1601  TCCCGCGCAA AAGCGGAATT CGATATCAAG CTTATCGATA CCGTCGACCT
1651  CGAGGGGGGG CCCGGTACCA GCTTTTGTTC CCTTTAGTGA GGGTTAATTT
1701  CGAGCTTGGC GTAATCATGG TCATAGCTGT TTCCTGTGTG AAATTGTTAT
1751  CCGCTCACAA TTCCACACAA CATACGAGCC GGAAGCATAA AGTGTAAAGC
1801  CTGGGGTGCC TAATGAGTGA GCTAACTCAC ATTAATTGCG TTGCGCTCAC
1851  TGCCCGCTTT CCAGTCGGGA AACCTGTCGT GCCAGCTGCA TTAATGAATC
1901  GGCCAACGCG CGGGGAGAGG CGGTTTGCGT ATTGGGCGCT CTTCCGCTTC
```

BAPL

```
   1  CGCCACCACC CTCCATCCCC CACTCCGCCC GCAAAGCTCG CCTCCTCCAT
  51  CACTGTGCCG CCCGTGCTCG CCTCTCCCGG CAGGCACCGA TCTAGCCCTC
 101  CCTCCCTTGA TCCACAGCTC GCCATGTGCG GCGGAGCCAT CCTCGCGGGA
 151  TTCATCCCGC CGTCGGCGGC CGCGGCGGCG GCCAAGGCAG CGGCGACGGC
 201  CAAGAAGAAG CAGCAGCAGC GCAGCGTGAC GGCAGACTCG CTCTGGACGG
 251  GCCTGCGGAA AAAGGCGGAC GAGGAGGACT TCGAGGCCGA CTTCCGCGAC
 301  TTCGAGCGGG ACTCCAGCGA GGAGGAGGAC GACGAGGTCG AGGAGGTCCC
 351  CCCTCCGCCG GCGCCGGCGA CGGCCGGGTT CGCCTTCGCC GCCGCGGCCG
 401  AGGTCGCGCT CAGGGCCCCT GCCCGCCGAG ATGCTGCTGT TCAACATGAT
 451  GGACCTGCTG CCAAACAAGT AAAGCGCGTT CGGAAGAATC AGTACAGAGG
 501  GATCCGCCAG CGTCCCTGGG GGAAATGGGC AGCTGAAATC CGTGACCCTA
 551  GCAAGGGTGT CCGGGTTTGG CTCGGGACAT ACGACACTGC TGAGGAGGCA
 601  GCCAGGGCAT ATGATGCTGA GGCCCGCAAG ATCCGTGGCA AGAAAGCCAA
 651  GGTCAATTTT CCTGAGGATG CTCCGACTGT TCAGAAGTCT ACCCTGAAGC
 701  CAACTGCTGC TAAATCAGCA AAGCTAGCTC CACCTCCGAA GGCCTGCGAG
 751  GATCAGCCTT TCAATCATCT GAGCAGAGGA GACAATGATT TGTTCGCGAT
 801  GTTTGCCTTC AGTGACAAGA AGGTTCCTGC AAAGCCAACT GACAGTGTGG
 851  ATTCCCTTCT TCCAGTGAAA CACCTTGCCC CCACCGAGGC ATTCGGAATG
 901  AACATGCTCT CTGACCAGAG CAGCAATTCA TTTGGCTCCA CTGACTTTGG
 951  GTGGGACGAC GAGGCCATGA CCCCGGACTA CACGTCCGTC TTCGTACCGA
1001  GTGCTGCTGC CATGCCGGCG TACGGCGAGC CCGCTTACCT GCAAGGCGGA
1051  GCTCCAAAGA GAATGAGGAA CAACTTTGGC GTAGCTGTGC TGCCTCAGGG
1101  AAATGGTGCA CAAGACATCC CTGCTTTTGA CAATGAGGAT GGGGCAAGCA
1151  GTGGGGATCT CTGGAGCCTC GATGAGCTGT TCATGGCAGC TGGTGGTTAT
1201  TGATGGTTCT TGTCAATGTG GTCTGCGGAC AGCACAAATG TCCCTTGCAT
1251  GTGGGCAAGA TGAAGAATTG GTGGTGCATG TGGCCAAGAT GAAAGAAGGA
1301  TCGGTGGCTT CTGCTACGTT CTGTAGCGGA TGAAACCATA GTTATGCTAA
1351  AGACTGTATG CTGCTAGCAG TGGGAACCGT ATGGTCATGT TTATATTCTG
```

Mesuraments of the two major leaf axis of WT and AD-partial BBR overexpressing lines. Values are expressed in cm.

leaf n°4	plant	lenght	width
	1 w.t.	19	9,5
	2 w.t	21	10,0
	AD1.1	20	4,0
	AD1.2	20	5,0
	AD1.3	24	6,0
	AD1.4	18	3,5
	AD1.5	20	6,0
	AD1.6	23	4,0
	AD2.1	24	4,5
	AD2.2	25	5,5
	AD2.3	20	3,0
	AD2.4	25	4,5
	AD2.5	20	2,0
	AD3.1	22	4,5
	AD3.2	20	4,5
	AD3.3	29	7,5
	AD3.4	22	4,5
	AD3.5	20	4,5
	AD3.6	20	4,5
	AD3.7	23	4,5
	AD3.8	25	5,0

leaf n°5	plant	lenght	width
	1 w.t.	15	7
	2 w.t	17	7,5
	AD1.1	14	4
	AD1.2	16	3
	AD1.3	13	3,5
	AD1.4	18	4
	AD1.5	20	4
	AD1.6	16	2
	AD2.1	12	2
	AD2.2	20	4
	AD2.3	20	3,5
	AD2.4	20	3,5
	AD2.5	25	2
	AD3.1	19	3
	AD3.2	11	2
	AD3.3	15	5
	AD3.4	18	4
	AD3.5	24	5,5
	AD3.6	17	3
	AD3.7	14	4
	AD3.8	14	3

lenght		average	**S.D.**
	WT	20	1,41
leaf n°4	AD-Pa.BBR	22,10	2,72
leaf n°5	WT	16	1,41
	AD-Pa.BBR	17,15	3,83

width		average	**S.D.**
leaf n°4	WT	9,8	0,35
	AD-Pa.BBR	4,6	1,17
leaf n°5	WT	7,25	0,35
	AD-Pa.BBR	3,42	0,98

Mesuraments of the sepals of WT and AD-partial BBR overexpressing lines. Values are expressed in cm.

Sepals	plant	lenght
	1 w.t.	1
	2 w.t	1,2
	AD1.1	2
	AD1.2	1,5
	AD1.3	1,7
	AD1.4	1,5
	AD1.5	1,7
	AD1.6	1,4
	AD2.1	1,8
	AD2.2	1,5
	AD2.3	1,8
	AD2.4	1,4
	AD2.5	1,5
	AD3.1	1,5
	AD3.2	1,3
	AD3.3	1,4
	AD3.4	1,7
	AD3.5	2
	AD3.6	1,5
	AD3.7	1,6
	AD3.8	1,6
		S.D.
WT	1,1	0,14
AD-Pa-BBR	1,55	0,19

Measurements of GUS activity in transient protoplasts expression experiments

GUS values after normalisation to the protein content are expressed in pmol 4-MUG /µg protein/min.

EXP. 1

		30min 1	1h 2	2h 3
	1	13643	30378	144108
	2	9206	14936	89893
	3	15629	15241	101838
K373	4	11478	11201	37627
	mean	12489	17939	93366,5
		1	2	3
	1	29434	71576	336500
	2	30290	83666	296966
	3	17104	29740	115920
Min1	4	38038	83076	462076
	mean	28716,5	67014,5	302865,5
		1	2	3
	1	77924	132320	1300000
	2	78956	187935	3574545
	3	85691	154789	1236956
Min1	4	121963	213687	2711111
BBR	mean	91133,5	172182,75	2205653
		1	2	3
	1	125836	296354	1548796
	2	297897	698745	3056987
	3	112547	243690	2215473
Min1	4	261875	395123	1987456
VP16BBR	mean	199538,75	408478	2202178

means referred to K 373

1	2	3	mean
1,00	1,00	1,00	1,00

means referred to K 373

1	2	3	mean
2,30	3,74	3,24	3,09

means referred to K 373

1	2	3	mean
7,30	9,60	23,62	13,51

means referred to K 373

1	2	3	mean
15,98	22,77	23,59	20,78

EXP. 2

		1	2	3
	1	9397	16250	43529
K373	2	5440	9218	18828
	mean	7418,5	12734	31178,5
		1	2	3
	1	23028	47672	97856
	2	8585	24828	79563
	3	16159	18369	89997
Min1	4	9940	15587	75530
	mean	14428	26614	85736,5
		1	2	3
	1	77144	87122	404444
	2	54265	82982	853097
	3	231521	425000	807608
Min1	4	92225	330000	443750
BBR	mean	113788,75	231276	627224,75
		1	2	3
	1	66987	90254	289745
	2	64253	165478	598214
	3	84457	106548	329998
Min1	4	39875	123578	389745
VP16BBR	mean	63893	121464,5	401925,5

means referred to K 373

1	2	3	mean
1,00	1,00	1,00	1,00

means referred to K 373

1	2	3	mean
1,94	2,09	2,75	2,26

means referred to K 373

1	2	3	mean
15,34	18,16	20,12	17,87

means referred to K 373

1	2	3	mean
8,61	9,54	12,89	10,35

EXP. 3

		1	2	3
K373	1	461	753	4692
	2	426	560	980
	3	739	866	1921
	mean	542	726	2531

means referred to K 373

1	2	3	mean
1,00	1,00	1,00	1,00

		1	2	3
Min1	1	2179	2585	5380
	2	4723	2728	21848
	3	1875	2483	3481
	4	1205	2048	3394
	mean	2495,5	2461	8525,75

means referred to K 373

1	2	3	mean
4,60	3,39	3,37	3,79

		1	2	3
Min1 BBR	1	22250	39654	35500
	2	15850	20456	29875
	3	4833	10231	11854
	4	8903	14306	19693
	mean	12959	21161,75	24230,5

means referred to K 373

1	2	3	mean
23,91	29,14	9,57	20,87

		1	2	3
Min1 VP16BBR	1	10305	7555	17277
	2	9358	9328	15731
	3	16538	23961	68115
	4	14236	17789	28289
	mean	12609,25	14658,25	32353

means referred to K 373

1	2	3	mean
23,26	20,18	12,78	18,74

Average of the folds of activation, in respect to K373, of the three different experiments reported. The standard deviation (S.D.) is indicated.

	K373	**Min1**	**Min1 + BBR**	**Min1 + VP16BBR**
	1,00	**3,05**	**17,42**	**16,62**
S.D.	**-**	**0,76**	**3,70**	**5,53**

GUS activity normalised to an internal LUC control.

GUS activity is expressed in pmol 4-MUG X 10000 RLU^{-1}.

Min1		**30min** 1	**1h** 2	**2h** 3
	1	0,0025	0,0063	0,0076
	2	0,0024	0,0059	0,0082
	3	0,0041	0,0087	0,0109
	4	0,0024	0,0061	0,0079
	5	0,0046	0,0079	0,0131
	6	0,0042	0,0077	0,0128
	7	0,0032	0,0056	0,0409
	8	0,0013	0,0115	0,0421
	9	0,0029	0,0061	0,0074
	10	0,0026	0,0056	0,0077
	mean	0,00302	0,00714	0,01586

1	2	3	mean
1,00	1,00	1,00	1,00

Min1+BBR		1	2	3
	1	0,0231	0,0564	0,0752
	2	0,0074	0,0551	0,1379
	3	0,0259	0,0975	0,1663
	4	0,0314	0,1256	0,2069
	5	0,0081	0,0177	0,0126
	6	0,0075	0,0086	0,0118
	7	0,0136	0,0095	0,0115
	8	0,0166	0,0545	0,1098
	9	0,0213	0,1136	0,3165
	10	0,0075	0,0531	0,1329
	mean	0,01624	0,05916	0,11814

1	2	3	mean	S.D.
5,38	8,29	7,45	7,04	1.5

Erklärung

"Ich versichere, daß ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit –einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; daß diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; daß sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht

worden ist sowie, daß ich eine solche Veröffentlichung vor Abschluß des Promotionsverfahrens nicht vornehmen werde.

Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Francesco Salamini betreut worden."

Köln, den 23.03.2002

Santi Luca

Teilpublikationen:

1 Müller, J., Müller, K., Pozzi, C., Santi, L., Wang, Y., Salamini, F., Rohde, W. (2000) Networking around the barley *Hooded* locus: molecular analysis of potential partners for epiphyllous flower formation. *Barley Genetics VIII.* 114-116.

2 Muller J., Wang Y., Franzen R., Santi L., Salamini F., Rohde W.. "In vitro interactions between barley TALE homeodomain proteins suggest a role for protein-protein associations in the regulation of Knox gene function". Plant J. 2001; 27(1): 13-23

3 Cristina Roig, Carlo Pozzi, Luca Santi, Judith Muller, Michele Stanca, Francesco Salamini. "The genetics of *Hooded* suppression". Submitted to Genetics.

Lebenslauf

Name

Santi Luca

Adresse

Max-Planck-Institut fuer Zuechtungsforschung
Carl-von-Linne Weg 10
D-50829
Tel: +49-221-5062-405
Fax: +49-221-5062-413
e-mail: santi@mpiz-koeln.mpg.de

Geburtsdatum und ort	Nationalitaet
11.11.1973 Genoa Italien	italienisch

Ausbildung

1979-1984	Grundschule
1984-1992	Gymnasium
1992	Abitur
1992-1998	Studium der Biologie an der Universitaet Genova, Italien
1995	Erasmus programme Galway Universitaet Irland
1996	Practicum an der IST Institut Genova
1997	Diplomarbeit an der Universitaet Genova, Prof. Cesarone Thema: "Modulation of the poly(Adp-ribose) polymerase by glucocorticoid.
08.10.97	Diplompruefung
04/1998	Beginn der Doktorarbeit am Max-Planck Institut fuer Zuechtungs-forschung, Koeln, Abt. Prof. F. Salamini, Arbeitsgruppe Prof. Salamini.

Publikationen

1 Müller, J., Müller, K., Pozzi, C., Santi, L., Wang, Y., Salamini, F., Rohde, W. (2000) Networking around the barley *Hooded* locus: molecular analysis of potential partners for epiphyllous flower formation. *Barley Genetics VIII.* 114-116.

2 Muller J., Wang Y., Franzen R., Santi L., Salamini F., Rohde W.. "In vitro interactions between barley TALE homeodomain proteins suggest a role for protein-protein associations in the regulation of Knox gene function". Plant J. 2001; 27(1): 13-23

3 Cristina Roig, Carlo Pozzi, Luca Santi, Judith Muller, Michele Stanca, Francesco Salamini. "The genetics of *Hooded* suppression". Submitted to Genetics.

Conferenzen

1 EGRAM Meeting June 1999, Montpellier France. "barley workshop".

2 TEBIO Meeting May 2000, Genoa Italy.

3 13th IIGB Meeting 21-24 October, 2000 Capri Italy: " Plant development: from cell fate to organ formation".

4 PAGMX Meeting 11-16 January, 2002 San Diego USA: "Plant, Animal and microbes genomes".

Acknowledgement

I would like to thank Prof. Dr. F. Salamini for providing me the opportunity to work in his department at the Max-Planck-Institut für Züchtungsforschung.

I wish to expecially thank Maria Rosaria, Cristina and Anna.

Finally, I would like to express my deep gratitude to my parents.