DNA binding sites recognised in vitro by a knotted class 1 homeodomain protein encoded by the *hooded* gene, *k*, in barley (*Hordeum vulgare*)

Lene Krusell, Inge Rasmussen, Kirsten Gausing *

Department of Molecular and Structural Biology, University of Aarhus, DK-8000 Aarhus C, Denmark

Received 24 January 1997; revised version received 26 March 1997

Abstract The homeodomain of the knotted classes of transcription factors from plants differs from the well characterized Antp/En type homeodomains from *Drosophila* at key amino acid residues contributing to the DNA binding. A cDNA, *Hvh21*, derived from the *hooded* gene and encoding a full length homolog of *knotted1* from maize was isolated from barley seedlings and expressed as a maltose binding protein fusion in *E. coli*. The purified HvH21-fusion protein selected DNA fragments with 1–3 copies of the sequence TGAC. Gel shift experiments showed that the TGAC element was required for binding and the results further indicate that the HvH21-fusion protein binds DNA as a monomer.

Key words: Homeobox gene; Transcription factor; cDNA sequence; Gel shift; Barley; Maize

1. Introduction

The *knotted1* (*kn1*) gene from maize was the first homeobox gene isolated from plants [22]. The encoded protein fulfills the criteria for the expected function as a transcription factor since it localises to the nucleus [19] but it is not known which genes it controls. The *kn1* gene from maize was isolated via one of several dominant *kn1* mutations that confer distorted growth of the leaves and stunting, subsequently demonstrated to be caused by ectopic expression of the wild type protein in lateral veins [6,19]. The precise function of the wild-type *kn1* gene is not known, but extensive in situ localisation experiments showed that both the *kn1* transcript and the Kn1 protein are confined to the apical meristem and the sub-apical developing vascular strands during vegetative growth [6,19]. During embryogenesis, the transcript is found in the embryo from an early stage [20]. Homologs of the *knotted*, class 1 genes, defined as a subfamily of *knotted* genes based on sequence homology criteria as well as gene expression pattern, have been isolated from rice, soybean, *Arabidopsis*, and tomato [5,9–11,13]. The encoded proteins of this class of genes have very homologous homeodomains, a conserved basic domain N-terminal to the homeodomain and some sequence similarity outside these regions; the proteins are about 360 amino acids long. The expression in vegetative tissue is largely as described for *kn1* in maize but minor differences in expression patterns in different plants are evident [5,10,14].

Two interesting morphological mutations have been correlated with *knotted* genes, the dominant mutation causing the *hooded* phenotype in barley [15] and the recessive shoot meristemless mutation, *stm*, in *Arabidopsis* [10]. The *hooded* mutation causes the homeotic transformation of the awn into a mirror image secondary floret [21] and is accompanied by overexpression of the *hooded* gene in the lemma which leads to the formation of a secondary floret meristem [15]. The phenotypes of *kn1*, *hooded*, *stm*, and the results of ectopic expression experiments in transgenic plants [13,18] serve to suggest a role of *knotted* class 1 genes in meristem formation and/or maintenance.

Here *hvh21*, a barley homolog of the maize *kn1*, is described. It probably corresponds to the same gene as *hooded* judged by the high nucleotide sequence homology of coding and non-coding regions but the encoded proteins are not identical. In order to elucidate the functions of the Hooded protein as a transcription factor and thereby to clarify its roles in plant development we wish to identify genes that are regulated by Hooded, and as a start have determined the preferred DNA binding site in vitro.

2. Materials and methods

2.1. Biological material

Barley (*Hordeum vulgare* cv Bomi) was grown and selected tissues were harvested on day 7 as described [1].

2.2. cDNA cloning and DNA sequence analysis

A portion of a cDNA library (3×10^5 recombinants) in lambda ZAP1 prepared from poly(A+) RNA from the basal 2.5 cm of 7-day-old barley seedlings [12] was screened by low stringency hybridization [3] with all solutions including 4×SSC. The probe was a copy of 60 bp of the maize *kn1* cDNA sequence corresponding to amino acid 302–321 of the highly conserved region of the homeodomain [22]. Recombinant phages were purified and the pBluescript SK(−) plasmids were rescued according to the Stratagene lambda ZAP manual. Inserts were subsequently transferred to the higher yielding pBS(−) vector. Sequencing was performed with double-stranded DNA using the Sequenase 2 kit (United States Biochemicals). cDNA clones encoding the N-terminal of HvH21 were isolated as follows. Total RNA from the basal 2.5 cm of 7-day-old barley seedlings was purified with the RNAeasy kit for plant tissue (Qiagen Inc.) and the poly(A+) fraction was isolated with oligo-dT coated magnetic Dynabeads (Dynal) according to the supplier's instructions. The poly(A+) RNA was reverse transcribed for 45 min at 42°C in a 20 μl reaction containing 2 μg poly(A+) RNA, 900 ng oligo-dT, 50 units AMV reverse transcriptase (Boehringer Mannheim) and 16 units RNasin (Promega) in 50 mM Tris-HCl, pH 8.5, 8 mM $MgCl_2$, 30 mM KCl, 1 mM DTT, 4 mM $Na_4P_2O_7$, 500 μM each of dNTP. The RNA-DNA hybrids were purified on Sephadex G-50 and the RNA strand was removed with 2.5 units RNAse H (Boehringer Mannheim) in 20 μl 10 mM Tris-HCl, pH 8.0, 1mM EDTA in a 30 min reaction at 37°C. The cDNA (10 ng) was tailed with dG in a 20 μl reaction containing 50 units terminal deoxynucleotide transferase (Boehringer Mannheim), 5 μM dGTP, 200 mM K-cacodylate, 25 mM Tris-HCl, pH 6.6, 0.75 mM $CoCl_2$, 1.25 mg/ml BSA for 30 min at 37°C. *Hvh21* specific amplification was carried out with a primer (5′-GCGGAATTCT-GCAGCACCGGGTTGCTACAG-3′) complementary to the *Hvh21*

*Corresponding author. Fax: 45 86 196500.

Abbreviations: ITOG, isopropyl β-D-thiogalactoside

PII S0014-5793(97)00382-7

sequence coding for amino acid 73–79 (Fig. 1A) and the unspecific primer (5′-GCCGGAATTCCCTCGAGC$_{13}$-3′) complementary to the oligo-dG tail (all oligonucleotides were synthesized by DNA-technology ApS, Aarhus). The PCRs were carried out in 100 μl with 100 pmol of each of the two primers and 5 units Taq polymerase (Stratagene). The 30 PCR cycles were at 95°C for 1 min, 55°C for 1 min, 72°C for 2 min finishing with 10 min at 72°C. Two identical amplifications were carried out, the first round with 0.5 ng tailed cDNA and the second round with a DNA fragment of the expected size of 370 bp synthesized in the first round and isolated as a faint band from an agarose gel. The resultant 370 bp DNA fragment was digested with *Eco*RI and the fragment ligated into pBS(−). The inserts of five independent recombinants were sequenced as described above.

2.3. Expression and purification of HvH21-fusion protein

The coding regions of the original partial Hvh21 clone was combined with the coding region of the 5′ fragment obtained by the RT-PCR procedure using an *Nco*I restriction site at the start codon, a common *Pst*I site and a *Stu*I site located 10 bp downstream from the termination codon. The two fragments were inserted in the expression vector, pMAL-c2 (New England Biolabs) between the *Eco*RI and *Hin*dIII sites and cloned in *E. coli* JM109 [24]. To preserve an open reading frame between the maltose binding protein coding region and HvH21 coding region, the *Eco*RI and *Nco*I sites were filled in. The correct insertion of the Hvh21 coding region was verified by DNA sequence analysis. Synthesis of the HvH21-fusion protein was induced in an exponentially growing culture at 1.5×10^8 cells/ml with IPTG (0.3 mM final concentration) and the cells were harvested after 3 h. The cells were concentrated 10-fold in lysis buffer (20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM EDTA, 1 mM DTT), opened by sonication, the soluble HvH21-fusion protein was purified on an amylose resin column as described in the New England Biolabs manual and the maltose removed by dialysis against lysis buffer. The procedure was monitored by SDS-PAGE and the protein concentration was determined according to Bradford [2].

2.4. DNA binding site selection

The experiments were carried out with HvH21-fusion protein immobilised on nitrocellulose filters (200 μg/1 cm^3 filter) and reacted with a mixture of 65 bp DNA fragments with a central core of 20 bp random sequence, amplified by PCR and reacted with fresh filters additionally 1 or 3 times, exactly as described by Nørby et al. [17]. The sequences of the DNA fragments were 5′-CGGGCTGAGAT-CAGTCTAGATCTNNNNN-NNNNNNNNNNNNNNNGGATC-CGAGACTGAGCGTCGTC-3′, and DNA fragments bound by the HvH21-fusion protein were amplified by PCR using primers matching the common sequences at the 5′ and 3′ end. Experiments were performed with HvH21-fusion protein and with BSA. No DNA was selected by BSA and one of the Hvh21-fusion protein experiments was negative. DNA fragments from the remaining experiments were inserted directly in the TA-cloning vector, pCRII (Invitrogen), using INVαF′ as host. The inserts of recombinant plasmids, verified by *Bgl*II digestion were sequenced as described above.

2.5. The gel shift experiments

DNA fragments containing one, two, or three HvH21-fusion protein binding sites or no binding site were taken from the collection of plasmids isolated in the binding site selection experiment (the precise sequence of the fragments is shown in Fig. 3A). The fragments were excised with *Eco*RI, purified on 5% acrylamide gels, labelled with

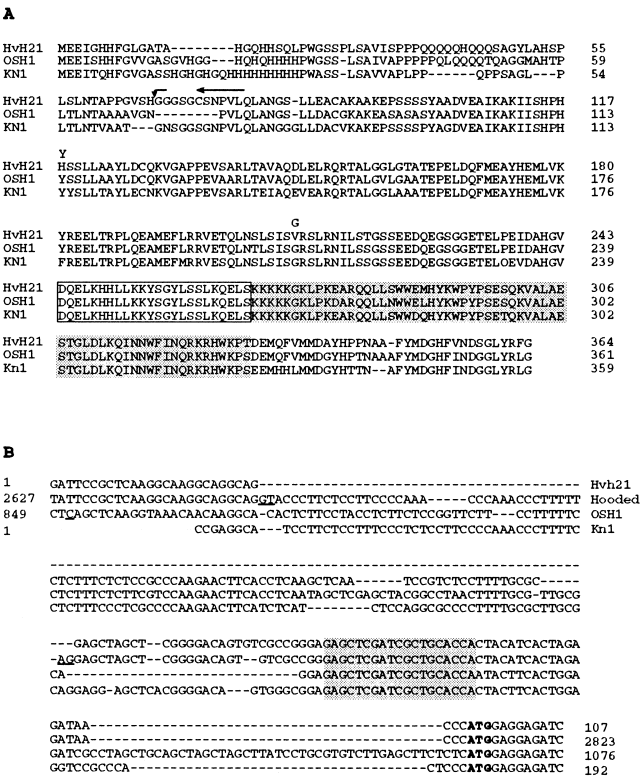

Fig. 1. Comparison of derived amino acid sequences of Knotted class 1 (Kn1) proteins and of 5′ non-translated leader sequences from corresponding genes and cDNAs. A: HvH21 sequence derived from the present work. OSH1 is the homologous sequence derived from the rice gene/cDNA (accession number D16507). Kn1 is the sequence derived from the maize gene, *knotted1* (accession number X61308). The amino acid differences between HvH21 and the sequence derived from the barley *hooded* gene (accession number X83518) are shown above the sequences. The homeodomain is shown in grey and the ELK region characteristic of the Knotted homeodomain proteins is shown in the box. The bent arrow indicates the extent of the primary *Hvh21* cDNA clone, and the leftward arrow indicates the position of the *Hvh21* specific primer employed to isolate sequences from the 5′ end of the *Hvh21* messenger RNA. B: alignment of the nucleotide sequences of the *Hvh21* cDNA, the *hooded* gene, the rice *OSH1* gene (the transcription start site is underlined) and the maize *knotted1* sequence. The grey box indicates a conserved sequence element in the non-translated leader sequences. The underlined GT-AG shows the location of the leader intron in the barley gene.

[^{32}P]dATP in the presence of 0.2 mM dTTP using the Sequenase DNA polymerase and buffer, and purified on Sephadex G-50 columns. The specific activity was 4000–8000 dpm/fmol fragment. The radioactive DNA fragments (5 fmol) were reacted for 30 min at room temperature with 0–500 ng HvH21-fusion protein or up to 500 ng maltose binding protein fused to paramyosin (New England Biolabs protein expression kit) in the presence of 2 μg poly d(I-C)-poly d(I-C) (Pharmacia) in a total volume of 10 μl binding buffer used in the DNA binding site selection experiments (25 mM HEPES, pH 7.9, 40 mM KCl, 3 mM $MgCl_2$, 1 mM DTT) supplemented with 10% by vol. glycerol. The samples were electrophoresed on 0.7% agarose/3% acrylamide gels and the gels dried as described [16]. Several autoradiograms were prepared for each gel. They were scanned on an Apple scanner and semi-quantitative analysis of band intensities was carried out on a Macintosh using the Image program from NIH.

3. Results

3.1. *Hvh21 cDNA cloning and sequence analysis*

The barley homolog of maize *kn1* was isolated from a barley seedling cDNA library constructed from polyA(+) mRNA from the basal 2.5 cm of 7-dday-old seedlings, using a copy of the most conserved region of the *kn1* homeobox as probe. Eight partial cDNA clones encoding proteins from the *knotted* family were isolated, four of which coded for the same close homolog of maize Kn1. The length of the inserts was 1208 bp starting at codon 68 (Fig. 1A) and terminating in a polyA tail. cDNA clones including the 5′ non-translated leader (Fig. 1B) and the N-terminal coding region were isolated with an RT-PCR procedure. The complete amino acid sequence derived from the original cDNA clone and the RT-PCR clones, designated HvH21, is shown in Fig. 1A. The barley amino acid sequence is 74% identical to the maize *kn1* sequence [22] and 81% identical with the rice *OSH*1 sequence [13] and the similarities increase to 94% when comparing the homeodomains. Northern analysis of *Hvh21* gene expression in different regions of young seedlings was positive only in the most basal region including the shoot apical meristem which is similar to the region of expression of *kn1* in maize seedlings (not shown). However, it is premature to conclude that the presumed orthologous genes *kn1*, *OSH1*, and *Hvh21* have precisely the same functions in seedling development.

Hvh21 most likely corresponds to the gene correlated with the *hooded* mutation in barley [15]. The complete gene isolated

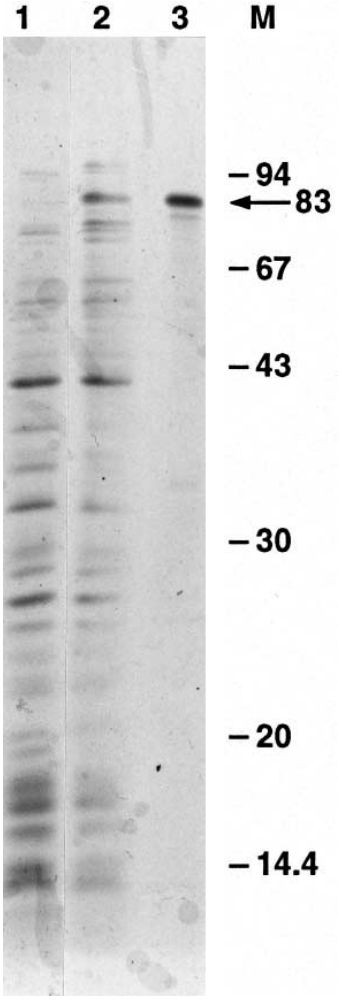

Fig. 2. SDS-acrylamide gels of the HvH21-fusion protein. Lane 1, extract of uninduced *E. coli* harbouring the Hvh21-pMALc2 plasmid; lane 2, soluble fraction of sonicated *E. coli* cells harvested 3 h after IPTG induction; lane 3, the HvH21-fusion protein after purification on amylose resin; lane M, molecular weight markers.

A

```
F2 B4 4                      tctagatctGACAGGGGCCGTCGTGATGGgatcc
F2 B4 5          tctagatcttTGATTGACAGGTGCTTGACggatcc
F2 B4 7             ggatcccGCCAGTCTCTGATGATGACagatctaga
F2 B4 9                      tctagatctGTGACACCTGTCGTGCAGTGgatcc
F2 B4 11               ggatccTCATTCCTGACGTGACACGGagatctaga
F2 B4 12             tctagatctACAGGGACGCGTGACGCTCGggatcc
F2 B4 13                     tctagatctGTGACAGTGGAGTTGACCGGgatcc
F2 B4 15                       ggatccAACGACCGATTTTTCTCGGAagatctaga
F2 B4 16           ggatccCAACATCAACCCCCATGACTagatctaga
F2 B4 18           tctagatctGATGTATGACAGCTGACGTggatcc
F2 B4 19                     tctagatctGTGACAGGAATGTGTGTGATggatcc
F2 B4 20                        ggatccGACGTTTGCTTACTGCTAGagatctaga
F2 B4 21           tctagatctATGACTGATGGGTGACAGGGgatcc
F2 B4 23                 ggatccTCCGACGTGTGACGCTCCGGagatctaga
F3 B2 1               tctagatctACGGTCCCTTGACGTTGTGTggatcc
F3 B4 1                      ggatccGACCCTTGCTACTATACAGCagatctaga
F3 B4 3                 tctagatctAGGACTGACAAACAGCTGATggatcc
F3 B4 4                    ggatccGAGACTGACCTGTTAGAGACagatctaga
F3 B4 6                 tctagatctGGTGACGTGACAGTTTTATAggatcc
F3 B4 7                 tctagatctAAGCTGACTGACGCATGACggatcc
F3 B4 8                      tctagatctGACCGCTCCTGTACGGGCGTggatcc
F3 B4 9              tctagatctGATTGACTGGTGACCGGCTCggatcc
F3 B4 10                     tctagatctGACGTTTGACGTGGACTGATggatcc
F3 B4 11                  tctagatctGACCTGACGTATGACACTTTggatcc
F3 B4 12           tctagatctTGTGCTTGACAGGTGACAGggatcc
F3 B4 13         tctagatctTGATGTAACGACGTGTGACAggatcc
F3 B4 14            ggatccACCTTCCTCCACCCTAACCAagatctaga
F3 B4 15                  tctagatctAAGTGATAGAGGAGGACAAggatcc
F3 B4 16           ggatccTGGCCATCAATTTCGTGACAagatctaga
F3 B4 18         tctagatctATATGTCCTGACATGTGACAggatcc
F3 B4 19            ggatccAACTGGAACTCCTGCACGCCagatctaga
F3 B4 21                     ggatccGGTGACCGCTGTCATACATCagatctaga
F3 B4 22                     tctagatctGACGTTGAGTGACTTGTGACggatcc
F3 B4 23         tctagatcttTTGTTGATTGGTCGATGACggatcc
F3 B4 24           tctagatctGGTCATAAGGGACGTGACAGggatcc
F3 B4 25                     tctagatctCTGACCGCTGTCACTTCCAGggatcc
F5 B2 2                   ggatccCCCCGCCTGAAGCCACTAACagatctaga
```

B

SEQUENCE	N	N	N	N	T	G	A	C	N	N	N	N
G	11	9	7	16	1	**35**			13	18	13	11
A	3	14	5	4		1	**37**	1	15	4	9	7
T	11	3	13	3	**33**			1	1	7	9	15
C	12	11	12	14	3	1		**35**	8	8	6	4
CONSENSUS	n	n	t/c	g/c	**T**	**G**	**A**	**C**	g/a	g/c	n	t/g

Fig. 3. Sequences of the central regions of DNA fragments bound by HvH21-fusion protein and identification of the preferred binding sequence. A: the conserved TGAC sequence is underlined. The separate binding site selection-amplification series are indicated by the F number, the number of cycles performed before cloning of the DNA fragment is given by the B number, and the number given to individual clones is shown last. The name of the four DNA fragments employed in the gel-shift analyses is underlined. B: simple count of the four deoxynucleotides at different positions in the alignment in A.

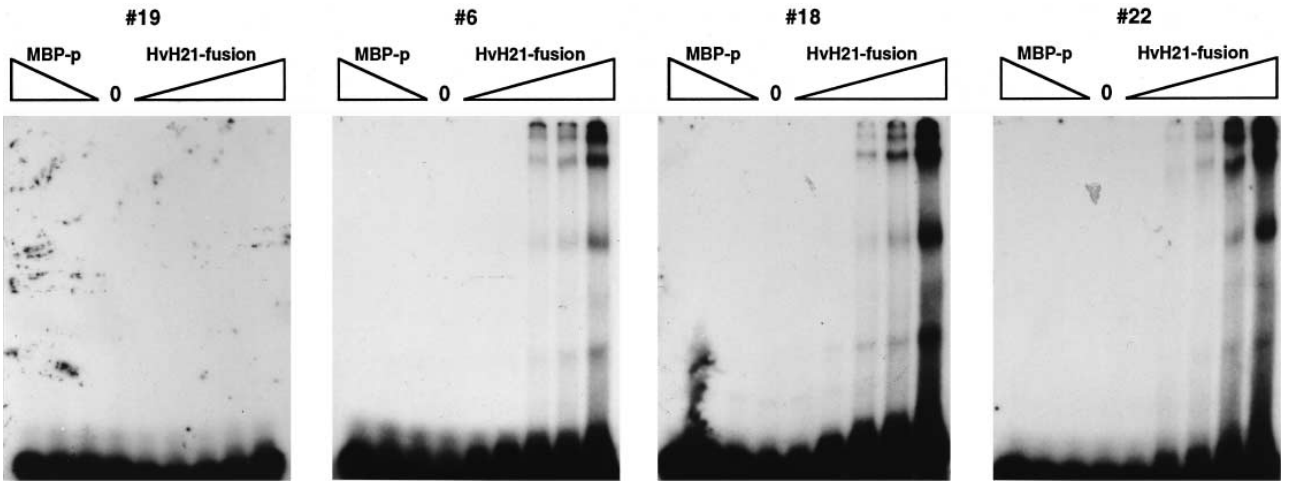

Fig. 4. Gel-shift analyses. All lanes contained 5 pmol ^{32}P-labeled DNA fragment and the identity of the DNA fragments is shown above each gel (the sequences are given in Fig. 3). The MBP-p (maltose binding protein-paramyosin fusion) control lanes contained 500, 200, and 100 ng protein, lane O contained no protein, and the remaining lanes contained 100, 200, 300, 400, and 500 ng HvH21-fusion protein. All binding reactions included 2 μg poly(dI-dC)-poly(dI-dC). The specific activity of the DNA fragments were the same within a factor two and the autoradiograms shown were exposed for the same period of time. The top band is material trapped at the origin; the bottom band is the free probe.

from cv C1 was sequenced (accession number X83518) and the 5′ non-translated leader sequence is aligned with the Hvh21 leader in Fig. 1B. The sequences are almost identical except for a region bounded by consensus splice sites. An intron was indicated [15] and Fig. 1B places this leader intron at location 2653–2744 in the *hooded* gene sequence. The sequence of the unspliced 5′ non-translated leader of *hooded* is very similar to the leader sequence in *kn1* and *OSH1* as illustrated by the completely conserved domain of 19 bp (Fig. 1B) but there is no intron in the leader sequence of the *kn1* and *OSH1* genes. The coding regions of *Hvh21* and the *hooded* gene are mismatched at two positions that result in the amino acid differences shown in Fig. 1A. The 3′ untranslated regions are mismatched at five positions and the alignment showed that one polyadenylation site is located at 10279 in the *hooded* sequence.

3.2. Expression of a HvH21-fusion protein in E. coli

The maltose binding protein fusion system was used for the expression of the complete HvH21 homeodomain protein. The coding region of Hvh21 was combined in the expression vector pMAL-c2 (New England Biolabs) and the protein expressed starts with the maltose binding protein bridged by a linker including a factor Xa protease site and fused to the HvH21 amino acid sequence. The N-terminal sequence of HvH21 after factor Xa cleavage was predicted to be ISEFM, adding four amino acids to the end of the protein. *E. coli* strain JM109 was used as host and the cultures were induced with IPTG at 1.5×10^8 cells/ml. The growth rate of the culture decreased gradually and an optimal yield of fusion protein was obtained after 3 h growth in the presence of IPTG. Fig. 2 shows the SDS-acrylamide gel profiles of cell extracts before and 3 h after IPTG induction. The HvH21-fusion protein did not form inclusion bodies and the protein was purified on amylose resin directly from the supernatant of a sonicated extract. The purified fusion protein analysed on SDS-acrylamide gel appears as a single band of the expected mass (83 kDa) with trace amounts of smaller peptides (Fig. 2). Factor Xa processing was not feasible since it resulted in complete breakdown of the modified HvH21 protein leaving only the maltose binding protein intact.

3.3. Determination of DNA binding sites preferred by the Hvh21-fusion protein in vitro

The binding site selection method has been described in detail [17]. Briefly, HvH21-fusion protein was immobilised on nitrocellulose filters and reacted with a mixture of 65 bp DNA fragments with a core of 20 random bp. The filters were washed and the DNA eluted and amplified by PCR. A portion of the amplified DNA was reacted with a fresh filter and this cycle was repeated 2 or 4 times. The samples of DNA fragments were cloned, individual recombinant colonies were collected, and the inserts of 37 plasmids were sequenced. The cores of the sequences (Fig. 3A) were all different, thus resulting from separate binding events. Inspection of the sequences revealed that the majority contain the tetramer TGAC and that it is present in two or three copies in several cases. The multiple TGACs are located as direct repeats except in two cases. Comparison of the sequences at positions upstream or downstream from the TGAC core does not indicate strict sequence requirements, albeit some preferences are noted (Fig. 3B). The consensus binding site is thus short and asymmetrical, indicating that the HvH21-fusion protein binds DNA as a monomer [8].

3.4. DNA binding analysed by gel-shift experiments

The specificity of HvH21-fusion protein binding to TGAC containing DNA fragments were tested in gel-shift experiments. DNA fragments were prepared from four different plasmids from the binding site selection experiments, and their sequences are shown in Fig. 3A. They have one, two or three binding sites or no complete binding site. To test the possible binding of the maltose binding protein region of the HvH21-fusion protein a fusion to paramyosin was reacted with each of the four DNA fragments. The results (Fig. 4) show that none of the DNA fragments were bound by this control-fusion protein. Increasing amounts of the HvH21-fusion protein were reacted with the four different DNA fragments and the

analyses of the complexes (Fig. 4) show that the TGAC core element is absolutely required for binding. A pattern of two main slow migrating bands (the top band is material trapped at the origin) and a faster migrating weaker band is seen with all three TGAC containing DNA fragments and the DNA fragment with three TGAC elements is bound marginally better than the DNA fragment with one TGAC element. The multiple bands are most likely the result of the instability of the HvH21-fusion protein, and it is noted that the distribution of material in the different bands is independent of the protein concentration, in agreement with the indication that the protein binds as a monomer. A semi-quantitative analysis of the HvH21-fusion protein binding to the different DNA fragments as function of the protein concentration was carried out (not shown). It was seen that the DNA fragments with one, two, or three TGAC elements bind to the HvH21-fusion protein with similar affinity. In addition, the binding was linearly dependent upon the protein concentration, suggesting that no more than one HvH21-fusion protein can bind per DNA fragment, even when more than one TGAC element is available.

4. Discussion

The *Hvh21* cDNA coding for a homeodomain protein of the Knotted class 1 was isolated and used to construct a plasmid for expression of the protein in *E. coli*. The expression construct codes for a maltose binding protein fusion which is a soluble protein in bacterial extracts and readily purified on amylose resin. Homeodomain proteins are transcription factors that interact directly with DNA and the preferred DNA binding site for Hvh21-fusion protein was found to have the core sequence, TGAC (Fig. 3B). The DNA binding sites recognised by the majority of homeodomains investigated have a core sequence of A-T base pairs (TAAT) where the first two base pairs contact the N-terminal arm (residue 1–7) and the remaining base pairs primarily contact the DNA recognition helix (residue 42–60) of a prototype 60 amino acid homeodomain [4,8]. The core sequence recognised by HvH21 (TGAC) is similar to the core sequence recognised by the yeast transcription factor, Matα2 (TTAC) assuming similar orientation of the DNA element in the protein complex [23]. The Matα2 DNA recognition is correlated with a variant N-terminal arm and the substitution of a conserved isoleucine residue at position 47 in the prototype homeodomain with an asparagine residue. HvH21 has an equivalent asparagine residue (residue 317 in HvH21 in Fig. 1A) but the N-terminal arm is not similar to that of Matα2 or the prototype homeodomain. The knotted family has a second unusual DNA contacting residue, an isoleucine at position 50 in the prototype sequence (residue 320 in HvH21 in Fig. 1A).

The homeodomains of the knotted class 1 proteins which in addition to the sequences aligned previously [7] include HvH21/hooded from barley, stm from *Arabidopsis* [10], and Tkn1 from tomato [5] have identical DNA recognition helix and nearly identical N-terminal arm (lysine in position 1, 2, or 3 can be an arginine). It is therefore predicted that this class of homeodomain proteins have very similar DNA binding sites, sharing the TGAC core sequence. As discussed for homeodomain transcription factors in other organisms the high in vivo specificity is most likely conferred by additional DNA contacts as well as interactions with other transcription factors [4,8]. On the other hand, the DNA recognition sequences recognised in vitro in the absence of other proteins are closely similar to the DNA elements bound in vivo in the cases where both have been investigated [4].

TGAC elements are present in the *OSH*1 and *hooded* promoters but without the slightly preferred neighbouring base pairs shown in Fig. 3B and they are not located in the substantial regions of high homology found when the two promoters are compared (not shown). Autoregulation of the expression of *Hvh21* and *OSH1* is not ruled out but this and other features of the protein-DNA interaction can now be investigated in greater detail.

Acknowledgements: We thank Peder Nørby for the oligonucleotide samples for the binding site selections. Expert technical assistance was provided by Lisbet Kjeldberg, Ane Kjeldsen, and Bent Sørensen. This work was supported by The Danish Agricultural and Veterinary Research Council and The Danish Natural Science Research Council.

References

[1] R.B. Barkardottir, B.F. Jensen, J.D. Kreiberg, P.S. Nielsen, K. Gausing, Dev. Genet. 8 (1987) 495–511.
[2] M.M. Bradford, Anal. Biochem. 72 (1976) 248–254.
[3] K. Gausing, R. Barkardottir, Eur. J. Biochem. 158 (1986) 57–62.
[4] W.J. Gehring, M. Affolter, T. Bürglin, Annu. Rev. Biochem. 63 (1994) 487–526.
[5] D. Hareven, T. Gutfinger, A. Parnis, Y. Eshed, E. Lifschitz, Cell 84 (1996) 735–744.
[6] D. Jackson, B. Veit, S. Hake, Development 120 (1994) 405–413.
[7] R. Kerstetter, E. Vollbrecht, B. Lowe, B. Veit, J. Yamaguchi, S. Hake, Plant Cell 6 (1994) 1877–1887.
[8] A. Laughon, Biochemistry 30 (1991) 11357–11372.
[9] C. Lincoln, J. Long, J. Yamaguchi, K. Serikawa, S. Hake, Plant Cell 6 (1994) 1859–1876.
[10] J.A. Long, E.I. Moan, J.I. Medford, M.K. Barton, Nature 379 (1996) 66–69.
[11] H.C. Ma, M.D. Mcmullen, J.J. Finer, Plant Mol. Biol. 24 (1994) 465–473.
[12] L.H. Madsen, J.D. Kreiberg, K. Gausing, Curr. Genet. 19 (1991) 417–422.
[13] M. Matsuoka, H. Ichikawa, A. Saito, Y. Tada, T. Fujimura, Y. Kanomurakami, Plant Cell 5 (1993) 1039–1048.
[14] M. Matsuoka, M. Tamaoki, Y. Tada, T. Fuyjimura, A. Tagiri, N. Yamamoto, Y. Kanomurakami, Plant Cell Rep. 14 (1995) 555–559.
[15] K.J. Müller, N. Romano, O. Gerstner, F. Garciamaroto, C. Pozzi, F. Salamini, W. Rohde, Nature 374 (1995) 727–730.
[16] P.S. Nielsen, K. Gausing, Eur. J. Biochem. 217 (1993) 97–104.
[17] P.L. Nørby, N. Pallisgaard, F.S. Pedersen, P. Jørgensen, Nucleic Acids Res. 20 (1992) 6317–6321.
[18] N.R. Sinha, R.E. Williams, S. Hake, Genes Dev. 7 (1993) 787–795.
[19] L.G. Smith, B. Greene, B. Veit, S. Hake, Development 116 (1992) 21.
[20] L.G. Smith, D. Jackson, S. Hake, Dev. Genet. 16 (1995) 344–348.
[21] G.L. Stebbins, E. Yagil, Genetics 54 (1966) 727–741.
[22] E. Vollbrecht, B. Veit, N. Sinha, S. Hake, Nature 350 (1991) 241–243.
[23] C. Wolberger, A.K. Vershon, B. Liu, A.D. Johnson, C.O. Pabo, Cell 67 (1991) 517–536.
[24] C. Yanisch-Perron, J. Vieira, J. Messing, Gene 33 (1985) 103–119.