The EMBO Journal vol.10 no.6 pp.1459-1467, 1991

HBP-1a and HBP-1b: leucine zipper-type transcription factors of wheat

Tetsuya Tabata[1], Takuya Nakayama, Koji Mikami and Masaki Iwabuchi

Department of Botany, Faculty of Science, Kyoto University, Kyoto 606, Japan

[1]Present address: Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143, USA

Communicated by M.L.Birnstiel

Wheat transcription factors HBP-1a and HBP-1b bind to the hexamer motif, ACGTCA, of wheat histone gene promoters. HBP-1b also binds to the hexamer motif in the promoter of the 35S RNA gene of cauliflower mosaic virus, whereas HBP-1a does not. A cDNA clone encoding HBP-1b was isolated on the basis of its binding specificity to the hexamer motif. The deduced amino acid sequence indicates that HBP-1b, like HBP-1a, belongs to a leucine zipper class of transcription factors. Mutational analyses of the HBP-1a and -1b encoded cDNAs revealed that truncated polypeptides containing the leucine zipper and basic regions are sufficient for DNA binding. HBP-1a and -1b form homodimers, as expected from earlier studies on this class of transcription factors, but did not form heterodimers. Although the hexamer motif or its homologs exist in several plant genes, HBP-1a and -1b exhibited the highest binding affinity to the hexamer motif in the histone promoters, suggesting that both DNA binding proteins are involved in transcriptional regulation of wheat histone genes.

Key words: HBP-1a/HBP-1b/leucine zipper/transcription factor/wheat histone gene

Introduction

Transcription is regulated through sequence-specific binding of transcription factors to *cis*-acting DNA elements within promoter and enhancer sequences. A given *cis*-acting element generally binds a single protein, but there are notable examples of sites that are recognized by several distinct factors or of proteins which interact with different binding sites, thereby modulating variable transcriptional effects. For instance, the animal octamer motif (ATGCAAAT) was characterized originally as part of a well-conserved sequence of animal histone H2B promoters (Harvey *et al.*, 1982) but has thereafter been found in a number of different promoters. The octamer motif has been shown to interact with both ubiquitous (Oct-1) and lymphoid-specific (Oct-2) transcription factors. Oct-1 is implicated in S-phase regulation of human H2B genes (LaBella *et al.*, 1988; O'Neill *et al.*, 1988) and regulation of small nuclear RNA transcription (Ares *et al.*, 1987; Tanaka *et al.*, 1988), and Oct-2 is implicated in cell type-specific regulation of immunoglobulin genes (Clerc *et al.*, 1988). Another well characterized instance is the ATF binding site (GTGACGTCA) and its cognate transcription factor ATF in mammals (Lin and Green, 1988). E1A-inducible adenovirus E2, E3 and E4 promoters contain one or more ATF sites. In the E4 promoter, the ATF sites have been shown to be required for E1A inducibility (Lee and Green, 1987). In addition to viral genes, several cellular promoters also contain ATF sites. The best studied among them are the cAMP-inducible promoters in which the ATF site is referred to as a cAMP-response element (CRE) and the binding factor thereof is referred to as CREB (Roesler *et al.*, 1988). There are several cellular promoters that contain ATF sites, yet are not regulated by either E1A or cAMP; for instance, heat-inducible HSP 70 and serum-inducible c-*fos* gene promoters (Hai *et al.*, 1989). Recently, more than 10 independent cDNA clones encoding distinct ATF or CREB binding factors have been isolated (Hoeffler *et al.*, 1988; Gonzalez *et al.*, 1989; Maekawa *et al.*, 1989; Hai *et al.*, 1989; Berkowitz and Gilman, 1990). It has been thought that the different ATF proteins bind selectively to ATF sites present in various promoters, thereby mediating distinct signal transduction.

We have been studying wheat transcription factors HBP-1a and HBP-1b, which were identified by binding to a *cis*-acting element of wheat histone genes with the hexameric core sequence, ACGTCA (Mikami *et al.*, 1987, 1989a; Tabata *et al.*, 1989). The hexamer motif is also present in the promoters of several plant cellular and viral genes and genes in T-DNAs of *Agrobacterium*, and is similar to the ATF/CREB binding motifs. This sequence has been shown to be required for efficient transcription of the wheat H3 gene (Nakayama *et al.*, 1989) and the 35S RNA gene of cauliflower mosaic virus (CaMV) (Kawata *et al.*, 1989; Lam *et al.*, 1989). Our previous study showed that HBP-1b could bind to the hexamer (ACGTCA) in the promoters of the CaMV 35S RNA and the nopaline synthase (NOS) genes, whereas HBP-1a could not (Mikami *et al.*, 1989a). This finding suggests that the difference in the binding specificity between them is due to the difference in their abilities to recognize the flanking sequences of the hexamer. We hereafter, for convenience, simply refer to the gene-specific hexamer as the H3 hexamer, the CaMV hexamer, the NOS hexamer, etc., but it must be remembered that these hexamers have their particularly unique flanking sequences. Nuclear proteins that bind to the CaMV hexamer or the hexamer homologies in other genes have also been identified in several plants. A maize transcription factor, OCSTF (octopine synthase transcription factor), binds to the OCS (octopine synthase) element that is conserved within the promoter regions of seven T-DNA genes and three plant viral genes, including CaMV (Bouchez *et al.*, 1989). The OCS element includes two hexamer homologies, and OCSTF has been shown to bind to each hexamer homology in the OCS element (Tokuhisa *et al.*, 1990). We have found that HBP-1b also binds to the hexamers (ACGTCA) in the OCS element of the CaMV and the NOS gene (Mikami *et al.*, 1989b). It is likely that the OCS element can be viewed as a

composite of two hexamer homologies. Chua and his co-workers have identified a tobacco nuclear factor, ASF-1, that binds to the CaMV hexamer (they have referred to this sequence as as-1) in the OCS element (Lam *et al.*, 1989). They have also shown that ASF-1 binds to the H3 hexamer and the OCS element in the OCS and NOS promoters (Fromm *et al.*, 1989). These lines of evidence suggest that wheat HBP-1b, tobacco ASF-1 and maize OCSFT may belong to a family of transcription factors bearing the same binding properties.

We now report the isolation and structural analysis of the cDNA encoding HBP-1b. Sequence analysis of this cDNA showed that HBP-1b contains both basic and leucine zipper motifs, which are the functional domains for DNA binding and dimer formation of a group of eukaryotic transcription factors including yeast GCN4, and mammalian ATF/CREB, Jun, Fos and C/EBP (for review, see Busch and Sassone-Corsi, 1990). We have previously isolated a cDNA (λ17), encoding a wheat DNA binding protein which also interacts with the H3 hexamer and belongs to a leucine zipper class of transcription factors (Tabata *et al.*, 1989). The DNA binding properties of the λ17-encoded protein, which we previously called HBP-1, demonstrate that this cDNA clone encodes HBP-1a. Using the two cDNA clones, we have investigated the DNA binding domains and the dimerization properties of HBP-1a and HBP-1b. In addition, DNA mobility shift experiments using the hexamer homologies in various plant genes were performed to find possible target genes of HBP-1a and -1b. The binding affinities of HBP-1a and -1b to those sequences were considerably lower compared with those to the H3 hexamer. This finding, as well as other data, suggests that HBP-1a and HBP-1b may be involved in cell cycle-dependent transcriptional regulation of plant histone genes.

Results

Isolation of cDNA encoding HBP-1b

The wheat nuclear proteins HBP-1a and HBP-1b bind to the hexamer motif of the wheat histone genes (Mikami *et al.*, 1989a). In addition, HBP-1b also binds to the hexamer motif in the CaMV 35S promoter, whereas HBP-1a does not (Figure 1). We have already isolated a cDNA clone (λ17) encoding a protein that specifically binds to the hexamer motif in the wheat histone H3 gene (Tabata *et al.*, 1989). The cDNA-encoded protein was expressed under the control of the T7 phage promoter in *Escherichia coli*, and was examined by gel mobility shift assay (Figure 1). The formation of the probe−protein complex was not competed by the CaMV hexamer, and the relative mobility of the complex was the same as that of native HBP-1a extracted from wheat nuclei, indicating that the λ17 clone encodes HBP-1a. Therefore, we have named this clone 1a-17.

In order to determine the structure of the other factor, HBP-1b, we screened a wheat cDNA library in λgt11 for β-galactosidase fusion proteins using a probe containing the CaMV hexamer that binds to HBP-1b but not to HBP-1a. Of 3 000 000 clones screened, three clones remained after three rounds of screening. A gel mobility shift assay with a lysogen extract from one clone, λc8, gave a specific retarded band that was competed out by both the H3 and CaMV hexamers (data not shown). Therefore, we proposed that this clone may encode HBP-1b. Sequence analysis of

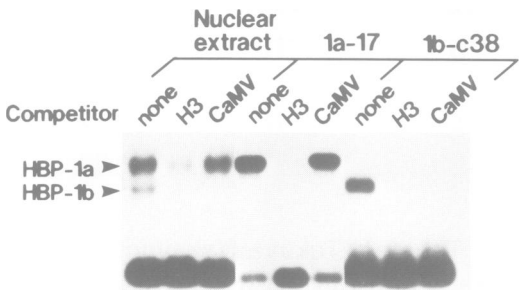


Fig. 1. DNA mobility shift assay of wheat nuclear extracts, and 1a-17- and 1b-c38-encoded proteins. The ^{32}P-labeled probe containing the H3 hexamer was incubated with wheat nuclear extracts, or bacterially expressed 1a-17- or 1b-c38-encoded protein, in the absence or the presence of the 10 ng of competitors. The sequences of the H3 and CaMV competitors are shown in Table I.

the insert revealed that λc8 did not encode the amino-terminal region of HBP-1b. The cDNA library was re-screened by plaque hybridization using a synthetic oligonucleotide corresponding to the 5′ terminal sequence of the λc8 insert. After three rounds of purification, the clone (1b-c38) containing the longest insert was isolated. The nucleotide sequence of 1b-c38 (Figure 2) was the same as that of λc8, except for 11 differences in nucleotides within the overlapping regions. The most proximal AUG codon in the open reading frame of 1b-c38 is preceded by a stop codon (position −9). Beginning at this putative initiation codon, the translated open reading frame apparently encodes a protein of 332 amino acids, with a predicted molecular mass of 36 800 daltons. The protein was expressed in *E.coli*, then subjected to gel mobility shift assay (Figure 1). The relative mobility and competition patterns were indistinguishable from those obtained with native wheat nuclear HBP-1b. To examine the binding properties of the protein encoded by 1b-c38 more precisely, the bacterially expressed protein was subjected to methylation interference analysis. The interference pattern was centered over the hexamer motif (Figure 3) and was quite similar to that obtained with native wheat HBP-1b (Mikami *et al.*, 1987; T.Tabata, K.Mikami and M.Iwabuchi, unpublished results). We therefore conclude that 1b-c38 encodes HBP-1b.

The predicted amino acid sequence reveals that HBP-1b belongs to a member of the leucine zipper family of transcription factors, which includes a highly basic region and a stretch containing leucines arranged as a heptad repeat (Busch and Sassone-Corsi, 1990). HBP-1a (1a-17) also has a leucine zipper domain, but significant sequence homology between 1a-17 and 1b-c38 is restricted to the basic and leucine zipper regions. Overall, the HBP-1b clone, 1b-c38, resembles the tobacco nuclear protein TGA-1a (ASF-1) that has also been identified by binding to the H3 hexamer and the CaMV hexamer (Katagiri *et al.*, 1989): there is 90% similarity in the amino acid sequences of the basic and leucine zipper regions, and 51% similarity in the overall regions between 1b-c38 and TGA-1a. In addition, HBP-1a contains a proline-rich domain near the amino terminus, and HBP-1b contains a glutamine-rich domain near the carboxy terminus. Recent studies have suggested that the mammalian transcription factor Sp1 activates transcription through a distinct motif characterized by a high glutamine content

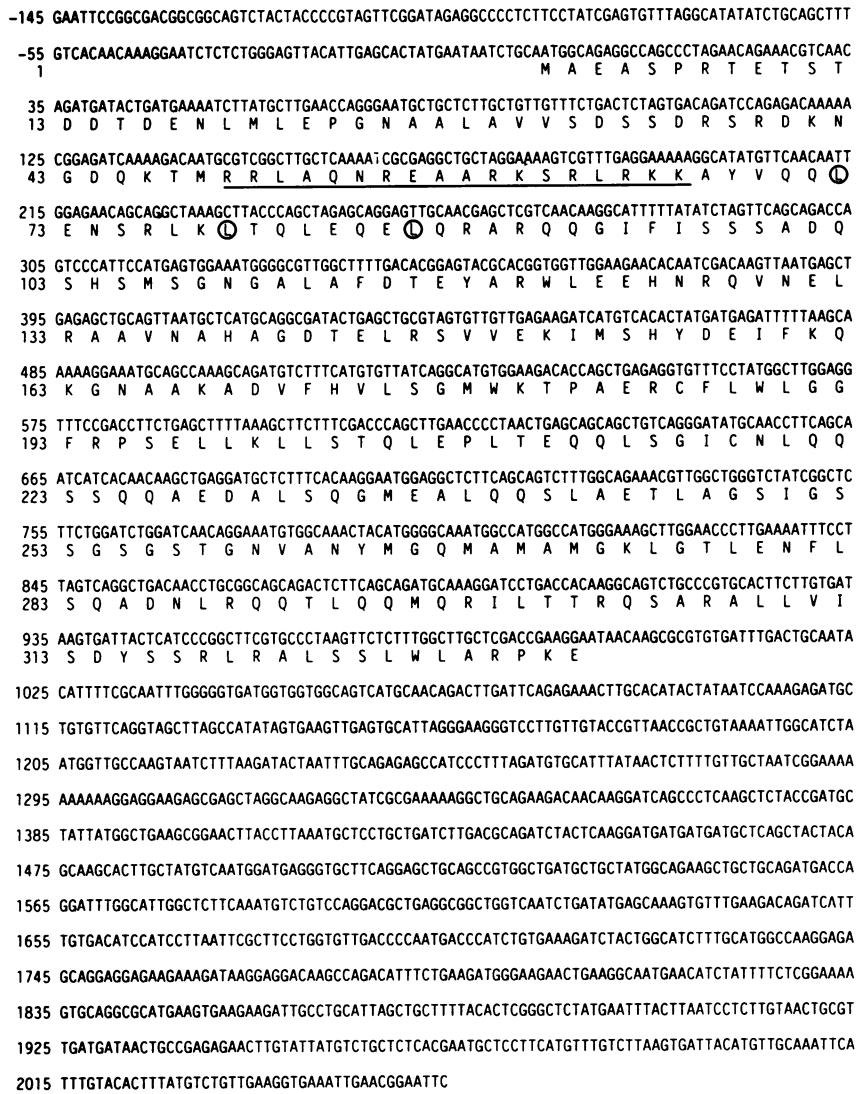


Fig. 2. The nucleotide sequence of 1b-c38 (HBP-1b) and the deduced amino acid sequence thereof. Numbering of nucleotides and amino acids begins with the first ATG codon. Amino acid sequence is shown using single letters. The basic motif is underlined and periodically appearing leucines in the leucine zipper structure are circled.

(Courey and Tjian, 1988). The factor CTF/NF1 (CAAT binding transcription factor/nuclear factor 1) has an active domain rich in proline residues (Mermod *et al.*, 1989). When considering these facts, it can be imagined that these domains of HBP-1a and -1b may also function as transcription activating regions (Figure 4A).

DNA binding domain and dimer formation of HBP-1a and HBP-1b

The basic and leucine zipper domains of transcription factors, such as GCN4, Jun/Fos and C/EBP, have been shown to mediate DNA binding and protein dimerization, respectively (Busch and Sassone-Corsi, 1990). To examine whether these domains of HBP-1a and -1b play similar roles in the protein−DNA interaction, two cDNA clones for truncated HBP-1a and HBP-1b, termed HBP-1aBL and HBP-1bBL, which contain both the basic and leucine zipper domains, were constructed and expressed in *E.coli*. Bacterially expressed proteins were analyzed by gel mobility shift assay (Figure 4). As shown in Figure 4B, both truncated proteins could bind to the DNA fragment containing the H3 hexamer motif, indicating that the truncated proteins containing the basic and leucine zipper regions of HBP-1a and -1b are sufficient for DNA binding, like those of other leucine zipper class of transcription factors. This, however, is not clear evidence about the DNA binding activity of HBP-1bBL (in Figure 4B), probably due to inefficient expression of HBP-1bBL in our system using transformed bacterial cells. Alternatively, the amino acid residues 1−127 of HBP-1bBL may be essential but not sufficient for DNA binding. When bacterial cell lysates containing HBP-1a and -1aBL were mixed, a third protein−probe complex of an intermediate electrophoretic mobility appeared (Figure 4B). This intermediate complex, due to the dimer formation between HBP-1a and -1aBL, suggests that HBP-1a binds DNA as a homodimer. However, such an intermediate complex was not formed between HBP-1b and -1-bBL, although they bound to the probe (Figure 4B).

In order to obtain additional evidence for protein dimerization, *in vitro* translated proteins were cross-linked with glutaraldehyde and subjected to SDS−polyacrylamide gel electrophoresis (Figure 4C). *In vitro* synthesized HBP-1a

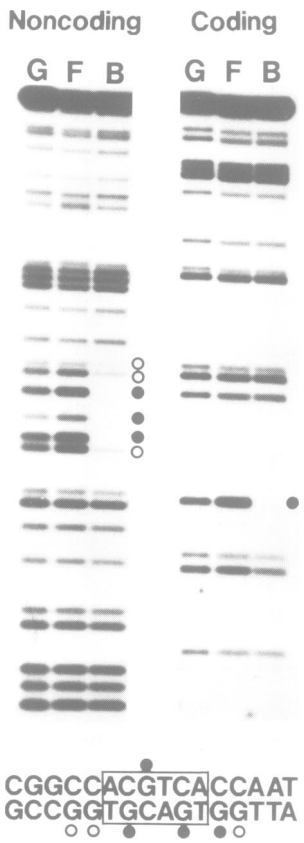


Fig. 3. Methylation interference analysis of 1b-c38-encoded protein. Methylation interference analysis was carried out with the probe containing the H3 hexamer, which had been end-labeled with [γ-^{32}P]ATP on either strand with T4 polynucleotide kinase. The partially methylated probe was incubated with bacterially expressed 1b-c38-encoded protein and separated by electrophoresis into a protein-bound fraction (lane B) and unbound fraction (lane F). Each fraction was cleaved at methylated G residues and analyzed by electrophoresis on a 6% sequencing gel. Lane G is a reference ladder generated by cleavage of the probe at G residues. The positions of methylated G residues that weakly and strongly interfere with binding are indicated by open and closed circles, respectively, and are depicted at the bottom.

migrated on gels with a molecular mass of 50 kd, which was larger than the value (37 kd) estimated on the basis of the deduced amino acid sequence from its cDNA. This phenomenon is probably due to a high content (12%) of proline residues. On the other hand, the apparent molecular mass of *in vitro* translated HBP-1b, as determined by PAGE, exhibited almost the same value as the predicted one (37 kd) from the deduced amino acid sequence. When *in vitro* translated proteins were treated with glutaraldehyde to cross-link anticipated multimers, bands of HBP-1a and -1b, with a molecular mass approximately twice that of the respective monomer, were observed in addition to the monomer bands. Thus, we conclude that HBP-1a and -1b each can form homodimers. As mentioned above, HBP-1b did not form a mixed dimer between the normal and truncated proteins. Taken together with the above observations, the results of the protein cross-linking experiments suggest that the preformed HBP-1b dimers are very stable.

Formation of the homodimers raises a possibility that a heterodimer between HBP-1a and HBP-1b may also form. To test this point, the two proteins, which were co-translated *in vitro* in the same test tube, were treated with glutaraldehyde. As shown in Figure 4C, no bands of the predicted mol. wt of heterodimer were observed; however, protein bands corresponding to the HBP-1a or HBP-1b homodimer appeared, suggesting that heterodimer between HBP-1a and HBP-1b is not formed *in vivo*. This result is consistent with observations that only protein–DNA complex of HBP-1a or -1b was found in DNA mobility shift assays with wheat nuclear extracts (cf. Figure 1). But, more detailed experiments using pure proteins of high concentrations are needed to rule out the possibility of the heterodimerization between HBP-1a and HBP-1b.

Binding specificity of HBP-1a and HBP-1b to hexamer motifs in various plant genes

Although HBP-1a binds only to the H3 hexamer, HBP-1b can be associated with the CaMV and NOS hexamers as well as the H3 hexamer, implying that 5′ and/or 3′ flanking sequences of the hexameric motif are also involved in an interaction between proteins and the DNA (Mikami *et al.*, 1989a). Gel mobility shift assays were performed using double-stranded synthetic DNA oligomers (H3m1 to H3m7), with base substitutions within the hexamer or its flanking sequences, as competitors to identify the critical nucleotides involved in sequence-specific DNA binding of HBP-1a and -1b. Figure 5 shows the competition patterns of various synthetic oligomers on the DNA–protein complex formation between the H3 hexamer and HBP-1a or -1b. The relative DNA-binding affinities were determined and are summarized in Table I. All mutations greatly reduced the ability of the oligonucleotide to compete binding of protein to the H3 hexamer although there were slight differences among them: alterations of two Cs 5′ or 3′ to the hexamer motif affected binding of HBP-1a more severely than binding of HBP-1b, whereas mutations within the hexamer motif itself reduced the ability to bind HBP-1b more than HBP-1a. Thus, it seems unlikely that all nucleotides within the hexamer and its surrounding sequences contribute to the HBP-1a and -1b binding in the same way.

The hexamer homologs are found in the upstream, presumably regulatory, regions of nuclear, plasmid and viral genes in plants. The gene *dbp* of *Arabidopsis thaliana*, encoding a lysine-rich DNA binding protein (Alliotte *et al.*, 1989), contains an 11 bp sequence GCCACGTCACC (DBP hexamer) in its 5′ flanking region that has perfect homology to the corresponding sequence of the H3 hexamer. A similar sequence, GCCACGTCACA (NITR hexamer), is present in the rice nitrate reductase gene (Choi *et al.*, 1989). The genes of chlorophyll *a/b* binding protein (AB) of *Nicotiana plumbaginifolia* (Castresana *et al.*, 1988), spinach ribulose bisphosphate carboxylase activase (RBCA) (Werneke *et al.*, 1989) and pea legumin A (Leg A) (Lycett *et al.*, 1985) also contain the hexamer homologs in their 5′ flanking sequences (Table I). The 'G box sequence', with the consensus sequence CCACGTGG, has been shown to function as a light-regulatory element in many plant genes (Schulze-Lefert *et al.*, 1989; Block *et al.*, 1990). The sequence CCACGTGGCC, containing the G box of the parsley chalcone synthase (CHS) gene (Herrmann *et al.*, 1988), resembles the H3 hexamer and its surrounding sequence. The OCS element is conserved in the promoters of opine synthase genes such as OCS (De Greve *et al.*, 1982), NOS (Depicker *et al.*, 1982) and mannopine synthase

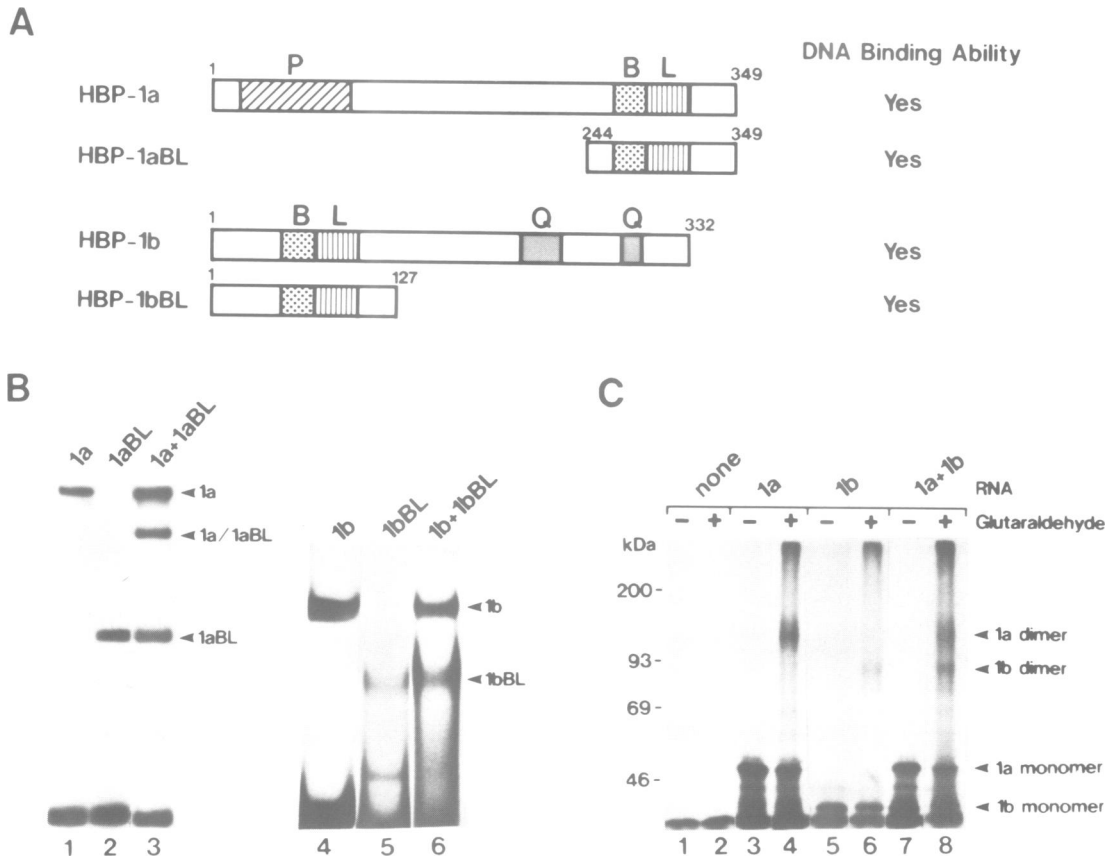


Fig. 4. DNA binding domain and dimer formation of HBP-1a and HBP-1b. (**A**) Schematic representation of HBP-1a and HBP-1b, and their truncated versions. The basic (B), leucine zipper (L), proline-rich (P) and glutamine-rich (Q) regions are depicted. (**B**) DNA mobility shift assay using bacterially expressed proteins. The ^{32}P-labeled probe DNA containing the H3 hexamer was incubated with bacterially expressed proteins, HBP-1a (lane 1), HBP-1aBL (lane 2), HBP-1a plus HBP-1aBL (lane 3), HBP-1b (lane 4), HBP-1bBL (lane 5), or HBP-1b plus HBP-1bBL (lane 6). Each probe–protein complex is indicated. (**C**) Cross-linking assay of HBP-1a and HBP-1b. HBP-1a and HBP-1b were separately translated (lanes 3–6) or co-translated (lanes 7 and 8) from mRNA, which had been synthesized *in vitro* from their cDNAs driven by T7 phage promoter, in the cell-free system with wheat germ extracts. These reactions were treated with glutaraldehyde (lanes 2, 4, 6 and 8), then run on 10% SDS–polyacrylamide gels.

(MAS) (Velten *et al.*, 1984) genes on T-DNAs of several *Agrobacterium* Ti plasmids and plant viral genes including CaMV (Franck *et al.*, 1980). This sequence has been shown to function as an active *cis*-element (Bouchez *et al.*, 1989) in an array of two hexamer homologs. To determine whether HBP-1a or -1b binds to these hexamer homologs, the competitive abilities of 20 bp synthetic oligomers, with hexamer homologs in the above described genes and their flanking sequences, were examined by DNA mobility shift assays (Figure 5). The binding affinities of these synthetic oligomers to HBP-1a and -1b are summarized in Table I.

Of the synthetic oligomers tested, the H3 hexamer exhibited the highest affinity for both HBP-1a and -1b. The DBP and NITR hexamers could bind to both HBP-1a and -1b, as expected, although their binding affinities were less than one-fifth that of the H3 hexamer, suggesting that flanking sequences of a perfectly matched 10–11 bp sequence, including the DBP or NITR hexamer, must also influence an interaction with the DNA binding proteins. The Leg A hexamer also bound to the two proteins with lower affinity (6–10% of the H3 hexamer affinity). The CHS G box could bind small amounts of HBP-1a,and no HBP-1b. In contrast, the RBCA hexamer and the hexamer homologs in the OCS elements bound very slightly to only HBP-1b. Unexpectedly the MAS and CaMV hexamers had higher affinities to HBP-1b than the DBP hexamer, yet the DBP hexamer apparently has a high sequence homology to the H3 hexamer, compared with the MAS and CaMV ones. Thus, it is supposed that surrounding sequences as well as the hexamer may also be implicated in an interaction between DNA and HBP-1b.

Discussion

HBP-1 family in plants

HBP-1a and -1b were first identified as nuclear proteins that specifically bind to the hexamer motif in the wheat histone H3 and H4 genes (Mikami *et al.*, 1989a). In addition, it is possible to consider that there may be other target genes of HBP-1a and -1b because the hexamer and its homologs are present in the 5′ flanking regions of some plant nuclear genes. Competition analyses using a DNA mobility shift assay revealed that some hexamer homologs of other genes can bind to both HBP-1a and -1b, and the H3 hexamer had higher binding affinity to the two proteins than any other hexamer homologs.

Regulatory and activating transcription factors, which may

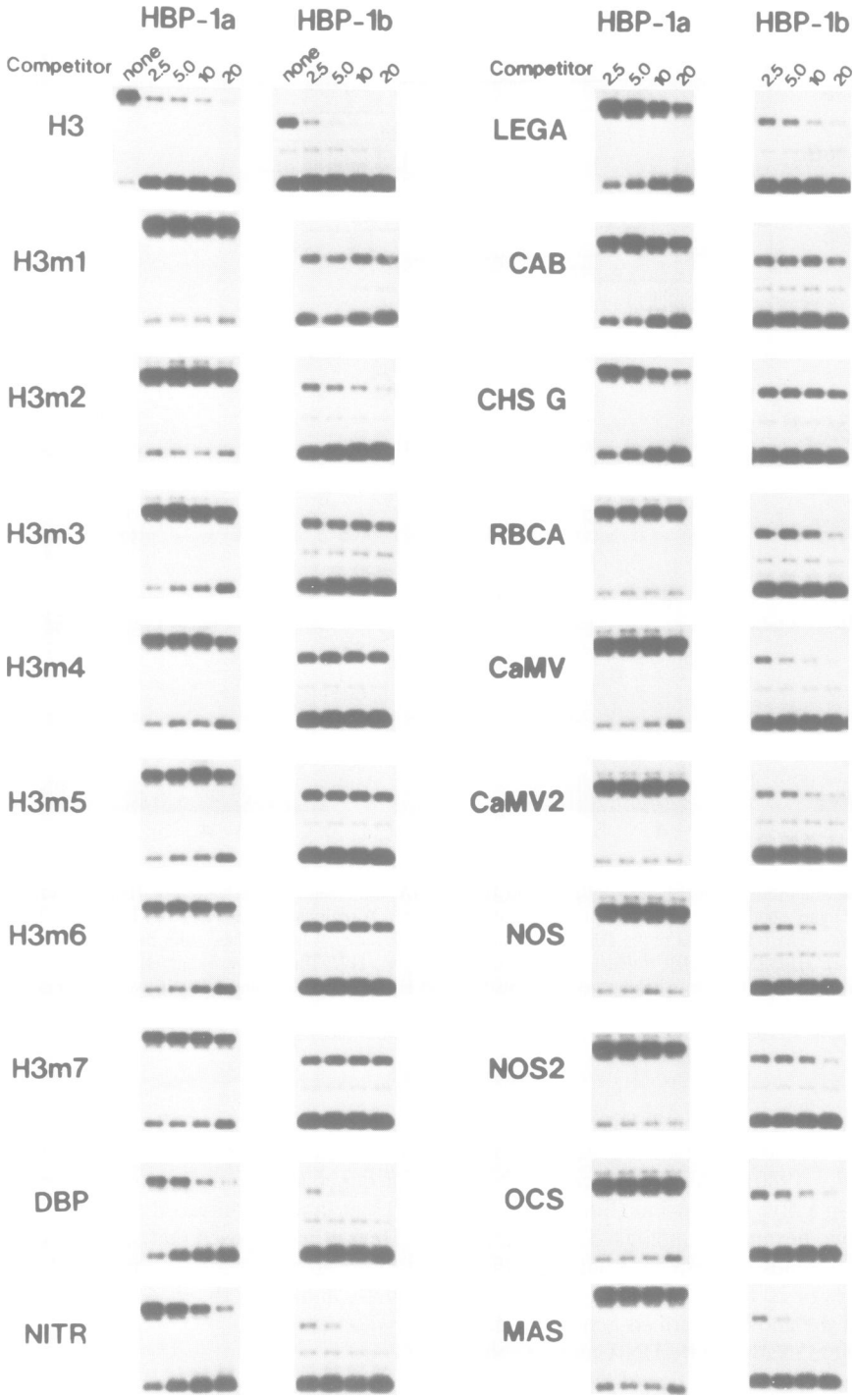


Fig. 5. DNA mobility shift assays of HBP-1a and HBP-1b with various competitors. A ^{32}P-labeled 90 bp probe containing the H3 hexamer was incubated with bacterially expressed HBP-1a or HBP-1b in the presence of 2.5 ng, 5 ng, 10 ng or 20 ng of various competitors shown in Table I. Radioactivities of the retarded complexes and free probe were measured with Imaging analyzer (Fuji film).

interact with the hexamer homologs, seem to be involved in the expression of various genes: a cell cycle-regulated gene (histone H3), a catabolic-related gene (NITR), some light-controlled genes (CHS G box, CAB and RBCA), and at least one developmentally regulated gene (*LegA*). If these hexamer homologs function as *cis*-acting elements, they would not necessarily be functionally equivalent. Therefore, it seems unlikely that HBP-1a and -1b participate in the regulation of all of these genes. Rather, it is highly probable that different transcription factors, which exhibit similar binding specificity to that of HBP-1a or -1b, would be selectively associated with the hexamer homologs present in various promoters. In fact, we have isolated three other distinct cDNAs encoding DNA binding proteins of a leucine zipper-type; their binding specificity is somewhat different from that of either HBP-1a (1a-17) or HBP-1b (1b-c38) (A.Sakamoto, T.Tabata and M.Iwabuchi, unpublished results). Now, a group of DNA binding proteins, including

Table I. DNA binding specificity of HBP-1a and HBP-1b

Genes[a]	Sequences[b]	Distance from TATA[c]	Relative binding affinity[d] (%)	
			HBP-1a	HBP-1b
H3	T C G G C C A C G T C A C C A A T C C G	126	100	100
H3m1	T C G G C C A A G T A A C C A A T C C G		<1	<1
H3m2	T C G G A A A C G T C A C C A A T C C G		<1	8
H3m3	T C G G C C A C G T C A A A A A T C C G		<1	<1
H3m4	T C G G C C A A G T C A C C A A T C C G		2	<1
H3m5	T C G G C C A C G C C A C C A A T C C G		2	<1
H3m6	T C G G C C A C G T G A C C A A T C C G		3	<1
H3m7	T C G G C C A C G T C C C C A A T C C G		3	<1
DBP	T G A T G C C A C G T C A C C T C T G T (REV)	177	20	16
NITR	T T T T G C C A C G T C A C A T G C T G	21	13	16
LEGA	G T T G G A C A C G T C A T T G A A A G (REV)	19	6	10
CAB	A T T T G C C A C G T C T G A T T G T G (REV)	201	3	<1
CHS G	T T A T T C C A C G T G G C C A T C C G	119	8	<1
RBCA	T G G C G T A A C G T C A T T T G T T C (REV)	104	<1	3
CaMV	A T C C C T T A C G T C A G T G G A G A (REV)	39	<1	32
CaMV2	G G A T T G T G C G T C A T C C C T T A (REV)	27	<1	10
NOS	A G C A C A T A C G T C A G A A A C C A	91	<1	13
NOS2	T T T A A T G A G C T A A G C A C A T A	79	<1	3
OCS	A C C A A A A A C G T A A G C G C T T A	149	<1	4
MAS	G A T A C T T A C G T C A C G T C T T G	24	<1	57

[a]Sources for sequences: H3, wheat histone H3 (Tabata *et al.*, 1983); DBP, *dbp* of *Arabidopsis thaliana* (Alliotte *et al.*, 1989); NITR, rice nitrate reductase (Choi *et al.*, 1989); LEGA, pea legumin A (Lycett *et al.*, 1985); CAB, chlorophyll *a/b* binding protein of *Nicotiana plumbaginifolia* (Castresana *et al.*, 1988); CHS G, parsley chalcone synthase (Herrmann *et al.*, 1988); RBCA, spinach ribulose bisphosphate carboxy/oxygenase activase (Werneke *et al.*, 1989); CaMV, cauliflower mosaic virus (Franck *et al.*, 1980); NOS, nopaline synthase (Depicker *et al.*, 1982); OCS, octopine synthase (De Greve *et al.*, 1982); MAS, mannopine synthase (Velten *et al.*, 1984).

[b]The hexamer motifs are boxed. Mutated nucleotides are underlined. For comparative purposes the sequences indicated by (REV) are written in the opposite orientation.

[c]The distances in bp of these sequence elements from their respective TATA boxes.

[d]Expressed as the ratio of H3 competitor to heterologous competitors needed to achieve 50% reduction in binding of HBP-1a or HBP-1b to the ^{32}P-labeled H3 probe in gel mobility shift assays (Figure 5). Note that the values do not represent the absolute binding affinities of HBP-1a and HBP-1b.

HBP-1a and -1b, may be regarded as a HBP-1 (Histone gene Binding Protein-1) family.

Recently, leucine zipper-type transcription factors have been identified in other plants as well. The *O2* locus, which is known by classical genetic studies to activate the maize zein gene *in trans*, encodes a protein which belongs to the leucine zipper class of transcription factors (Hartings *et al.*, 1989; Schmidt *et al.*, 1990). The basic region of the O2 protein is more similar to that of HBP-1a than to that of HBP-1b, prompting us to speculate that the target site of O2 within the zein promoter may resemble the HBP-1a binding site. Chua and his co-workers have isolated two cDNAs, TGA1a and TGA1b, probably encoding ASF-1 and HSBF, respectively, from tobacco on the basis of their ability to bind to the H3 hexamer (Katagiri *et al.*, 1989; Lam *et al.*, 1990). Judging from their DNA binding properties, ASF-1 and HSBF seem to correspond to HBP-1b and -1a, respectively. However, there was no significant homology between HBP-1a and HSBF except for the basic and leucine zipper regions, although the amino acid sequence of HBP-1b was 51% homologous to that of ASF-1.

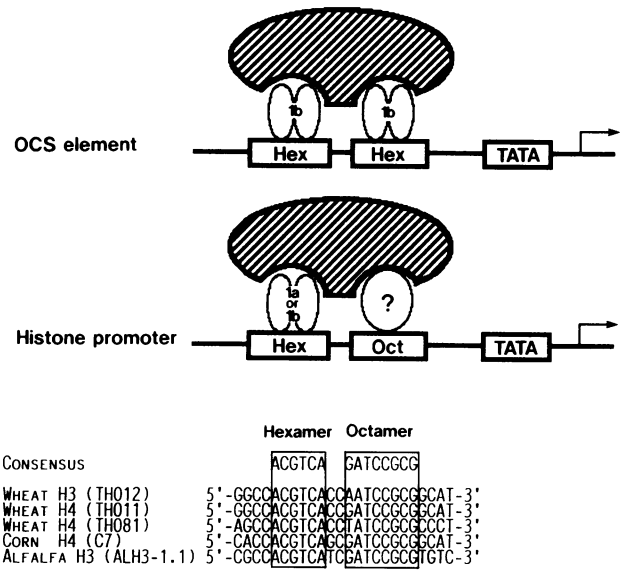
Most recently, cDNAs encoding other sequence-specific DNA binding proteins of a leucine zipper-type have been isolated in two laboratories: one is EmBP-1, which interacts specifically with ABRE (abscisic acid response element) in wheat (Guiltinan *et al.*, 1990) and the other is OCSBF-1 (OCS enhancer binding factor) in maize (Singh *et al.*, 1990). The DNA binding motif, CACGTGGC, within ABRE completely matches the consensus G box motif. Since HBP-1a can bind to the G box motif of the CHS gene, it would be important to know the relationship between HBP-1a and EmBP-1 to understand the function of the HBP-1 family.

The HBP-1 family can be compared with the ATF/CREB family in mammals: first, both contain multiple members of leucine zipper-type transcription factors; second, the core sequence of their target sites is the hexamer motif; and lastly, the hexamer motif exists within promoters of various plant and mammalian genes. It has been shown that ATF/CREB members are functionally distinguishable, although all ATF proteins can bind to ATF sites (hexamer motifs) *in vitro*. For instance, ATF-2 supports an E1a-inducible transcription of the adenovirus early genes, but unlike CREB, does not support a cAMP-inducible transcription response (Liu and Green, 1990).

Thus, it is proposed that in plant and animal cells, similar sets of *cis*-elements and *trans*-acting factors may commonly be utilized for regulation of gene transcription.

Possible role of HBP-1a and HBP-1b

The results of the *in vitro* DNA binding experiments have led us to conclude that the target genes of HBP-1a and HBP-1b are mainly histone genes. However, given that the OCS element, which can bind HBP-1b, confers tissue-specific gene expression in plants, we must consider the function of HBP-1b. Using a reporter gene construct, the OCS element in the CaMV 35S or the OCS promoter is preferentially expressed in proliferating tissues, for instance, root tips and shoot apex in transgenic tobacco (Benfey *et al.*, 1989; Fromm *et al.*, 1989). This evidence prompts us to speculate that HBP-1b would be a transactivator in the cell




```
                                  Hexamer  Octamer
CONSENSUS                         ACGTCA   GATCCGCG
WHEAT H3 (TH012)        5'-GGCCACGTCACCAATCCGCGGCAT-3'
WHEAT H4 (TH011)        5'-GGCCACGTCACCGATCCGCGGCAT-3'
WHEAT H4 (TH081)        5'-AGCCACGTCACCTATCCGCGCCCT-3'
CORN  H4 (C7)           5'-CACCACGTCAGCGATCCGCGGCAT-3'
ALFALFA H3 (ALH3-1.1) 5'-CGCCACGTCATCGATCCGCGTGTC-3'
```

Fig. 6. Schematic representation of a model in which homologous or heterologous factors simultaneously contact a common target protein, thereby regulating cell cycle-dependent transcription. The diagram shows the interaction of HBP-1a, HBP-1b and an as yet unidentified octamer binding protein with their target sites within the OCS element and the wheat histone promoter. Sequence comparisons of the promoter regions of several plant H3 and H4 histone genes are depicted at the bottom. The hexamer and octamer motifs are boxed. Sources for sequences: wheat H3 (TH012), Tabata *et al.* (1984); wheat H4 (TH011), Tabata *et al.* (1983); wheat H4 (TH081), Y.Fujimoto and M.Iwabuchi (unpublished results); corn H4, Philipps *et al.* (1986); alfalfa H3, Wu *et al.* (1988).

cycle-dependent transcription of wheat histone genes. If this is indeed the case, what is the function of HBP-1a? We would prefer to think that HBP-1a also participates in the regulation of cell cycle-dependent transcription of wheat histone genes because HBP-1a and -1b mRNAs, as well as histone mRNA, are abundant in proliferating tissues of wheat seedlings (M.Minami, T.Nakayama and M.Iwabuchi, unpublished results). In addition, the relative expression of the HBP-1a and HBP-1b mRNAs, in the early germination stage of wheat seeds, resembles that of the histone mRNA: the amounts of three mRNAs reached the first peak 24–28 h after imbibition (M.Minami, T.Nakayama and M.Iwabuchi, unpublished results). Although we do not know anything about functional differences between HBP-1a and HBP-1b, it seems likely that distinct target protein(s), which would be component(s) of the transcriptional apparatus, may contact HBP-1a or HBP-1b since the putative activation domains of HBP-1a and HBP-1b, which would interact with target proteins, are different from each other.

The OCS element contains two hexamer homologs, which have been shown to be required for activating the function of this element (Ellis *et al.*, 1987; Benfey *et al.*, 1989). Thus, a dimer of HBP-1b, which binds to each of two adjacent hexamer homologs, could better contact another distinct target protein simultaneously (Figure 6). On the other hand, the hexamer motif and a highly conserved octamer motif, GATCCGCG, which is located 2 bp downstream of the hexamer reside in the wheat histone H3 and H4 promoters (Figure 6). We therefore speculate that in histone genes, an as yet unidentified octamer binding factor may exist. This octamer binding factor may participate with HBP-1a or -1b in transcription regulation, perhaps in a cell cycle-dependent manner (Figure 6). Such regulatory mechanisms with multiple transcription factors have been proposed in other transcription regulation systems, for instance, the yeast transcription factor GAL4 binds to multiple sites on DNA to activate transcription synergistically (Carey *et al.*, 1990), and GAL4 also works synergistically with a heterologous factor ATF (Lin *et al.*, 1990). It remains to be determined whether the hexamer or octamer motif, or both, is really involved in the cell cycle dependency of plant histone gene transcription.

Materials and methods

Isolation of the HBP-1b cDNA

A wheat λgt11 cDNA library was screened using a catenated oligomer containing the CaMV hexamer as a probe according to the methods of Vinson *et al.* (1988) with slight modifications (Tabata *et al.*, 1989). We isolated three clones from 3×10^6 plaques after three rounds of purification. The insert of one of them (λc8) was recloned into M13mp19 and sequenced by the dideoxy sequencing method using Sequenase (USB). A 29 bp oligonucleotide derived from the 5′ end portion of the λc8 insert was synthesized and used as a probe to screen the same library by plaque hybridization to isolate a full-size cDNA. Of 3×10^6 plaques screened, 26 clones remained positive after three rounds of plaque purification. Among them, we chose 1b-c38, which contained the longest insert, and determined the nucleotide sequence as above.

Preparation of bacterially expressed HBP-1a and HBP-1b

The T7 polymerase expression vector pAR2106 (Studier and Moffat, 1986) was used to express each of the HBP-1a and -1b cDNA clones in *E.coli* cells. The plasmid pAR1a-17 (HBP-1a expression plasmid) was constructed by subcloning a 1.3 kb *Acc*II–*Eco*RI fragment containing an entire HBP-1a coding region into pAR2106 using *Bam*HI linkers. Oligonucleotide-directed mutagenesis was used to create an *Nde*I site at the initiating ATG of the 1b-c38 (HBP-1b) coding sequence by using the oligonucleotide GGCCTCTGCCATATGAGATTATTCATA. The plasmid pAR1b-38 (HBP-1b expression plasmid) was constructed by subcloning the 2 kb *Nde*I–*Eco*RI fragmrent of 1b-c38, mutated as above, into pAR2106. For the truncated HBP-1a protein, a *Nhe*I–*Xho*I fragment containing the amino-terminal half of HBP-1a was removed from pAR1a-17 (Figure 4A) and religated to create pAR1a-17BL. For the truncated HBP-1b protein, a *Sph*I–*Taq*I fragment of pAR1b-c38, which contained the T7 promoter and the amino-terminal half of HBP-1b (Figure 4A) was subcloned into pBR322 to create pAR1b-c38BL.

Proteins were expressed in the bacterial strain BL21(DE3), which contains the T7 RNA polymerase gene under the control of the inducible *lacUV5* promoter (Studier and Moffatt, 1986). Cells transformed with the expression plasmids were grown at 37°C in super broth with 100 μg of ampicillin per ml. The cells were grown to an A_{600} of 0.5, IPTG (isopropyl-β-D-thiogalactopyranoside) was added to the culture to a final concentration of 1 mM, then the cells were grown for an additional 2 h.

Induced cells were harvested by centrifugation and resuspended in 1/20 volume of Z0 buffer [25 mM HEPES, pH 7.9, 1 mM EDTA, 1 mM phenylmethyl sulfonyl fluoride (PMSF), 1 mM DTT and 0.15% Nonidet P-40]. Cells were frozen in liquid nitrogen and then thawed on ice. Lysozyme was added to a final concentration of 0.5 mg/ml, and the cell suspension was incubated on ice for 15 min. After this incubation, KCl was added to a final concentration of 1.3 M. The lysates were gently rotated at 4°C for 15 min, then centrifuged at 350 000 *g* for 45 min. The supernatant was dialyzed against the buffer Z0 with 100 mM KCl and 20% glycerol. Dialyzed samples were incubated with Sepharose 4B coupled with calf thymus DNA. The resin was packed in an Econocolumn (BioRad). Proteins were eluted with Z0 with 600 mM KCl and used for DNA binding assays.

Gel mobility shift assay

A *Hin*dIII–*Eco*RI fragment containing the 5′ flanking region, −185 to −130, including the hexamer motif of the wheat histone H3 gene (Tabata *et al.*, 1984) and flanking polylinker sites was isolated, 3′-end-labeled with Klenow enzyme and [α-^{32}P]dCTP, and used as a probe. The binding reaction mixture contains 1–2 μl of cell extracts containing bacterially expressed HBP-1a and -1b, or nuclear extracts prepared from wheat germ (Mikami *et al.*, 1987), 0.25 ng of the probe , 1 μg of poly d(I-C), 0–20 ng of 20 bp competitors, 12 mM HEPES (pH 7.9), 12% glycerol, 60 mM KCl, 0.2 mM EDTA, and 0.5 mM DTT. All 20 bp competitors were

chemically synthesized using Cyclone Plus DNA synthesizer (Milligen/Biosearch). The reaction mixture was loaded on 5% native polyacrylamide gels containing 0.5 × TBE. Radioactivities of the migrated complex and the free probe were measured by Image analyzer (Fuji film).

Methylation interference analysis

The same DNA fragment as that used for gel mobility shift assays was 5′-end labeled on either strand with T4 polynucleotide kinase and [γ-^{32}P]ATP, partially methylated on G residues with dimethylsulfate, and used as a probe. This probe was subjected to the gel mobility shift assay described above, then bound and unbound DNAs were eluted from the gel, cleaved with 1 M piperidine, and analyzed on a 6% sequencing gel.

Cross-linking of the proteins

The plasmids pAR1a-17 and pAR1b-c38 were linearized by cutting with *Eco*RI and used as templates for RNA synthesis. RNA was synthesized with T7 RNA polymerase using the transcription kit mCAP (Stratagene). *In vitro* translation of the purified RNA was carried out using wheat germ extract (Amersham) in the presence of [^{35}S]methionine.

After *in vitro* translation, 5 μl of products were incubated with 0.01% of glutaraldehyde in 75 mM phosphate buffer (pH 7.0) at 20°C for 30 min. The products were directly loaded on a 10% SDS–polyacrylamide gel. Fluorography was performed after treatment of the gel with Amplify (Amersham).

Acknowledgements

This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan, by grants from the Naito Foundation, Ciba–Geigy Foundation (Japan), Takeda Science Foundation, and the Research Council, Ministry of Agriculture, Forestry, and Fisheries, Japan, original and creative research project on biotechnology.

References

Alliotte,T., Tire,C., Engler,G., Peleman,J., Caplan,A., Van Montagu,M. and Inze,D. (1989) *Plant Physiol.*, **89**, 743–752.
Ares,M., Jr, Chung,J.-S., Giglio,L. and Weiner,A.M. (1987) *Genes Dev.*, **1**, 808–817.
Benfey,P.N., Ren,L. and Chua,N.-H. (1989) *EMBO J.*, **8**, 2195–2202.
Berkowitz,L.A. and Gilman,M.Z. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 5258–5262.
Block,A., Dangl,J.L., Hahlbrock,K. and Schulze-Lefert,P. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 5387–5391.
Bouchez,D., Tokuhisa,J.G., Llewellyn,D.J., Dennis,E.S. and Ellis,J.G. (1989) *EMBO J.*, **8**, 4197–4204.
Busch,S.J. and Sassone-Corsi,P. (1990) *Trends Genet.*, **6**, 36–40.
Carey,M., Lin,Y.-S., Green,M.R. and Ptashne,M. (1990) *Nature*, **345**, 361–364.
Castresana,C., Garcia-Luque,I., Alonso,E., Malik,V.S. and Cashmore,A.R. (1988) *EMBO J.*, **7**, 1929–1936.
Choi,H.K., Kleinhofs,A. and An,G. (1989) *Plant Mol. Biol.*, **13**, 731–733.
Clerc,R.G., Corcoran,L.M., LeBowitz,J.H., Baltimore,D. and Sharp,P.A. (1988) *Genes Dev.*, **2**, 1570–1581.
Courey,A.J. and Tjian,R. (1988) *Cell*, **55**, 887–898.
De Greve,H., Dhaese,P., Seurinck,J., Lemmers,M., Van Montagu,M. and Schell,J. (1982) *J. Mol. Appl. Genet.*, **1**, 499–511.
Depicker,A., Stachel,S., Dhaese,P., Zambryski,P. and Goodman,H.M. (1982) *J. Mol. Appl. Genet.*, **1**, 561–573.
Ellis,J.G., Llewellyn,D.J., Walker,J.C., Dennis,E.S. and Peacock,W.J. (1987) *EMBO J.*, **6**, 3203–3208.
Franck,A., Guilley,H., Jonard,G., Richards,K. and Hirth,L. (1980) *Cell*, **21**, 285–294.
Fromm,H., Katagiri,F. and Chua,N.-H. (1989) *Plant Cell*, **1**, 977–984.
Gonzalez,G.A., Yamamoto,K.K., Fischer,W.H., Karr,D., Menzel,P., Biggs,W., III, Vale,W.W. and Montminy,M.R. (1989) *Nature*, **337**, 749–752.
Guiltinan,M.J., Marcotte,W.R., Jr and Quatrano,R.S. (1990) *Science*, **250**, 267–271.
Hai,T., Liu,F., Coukos,W.J. and Green,M.R. (1989) *Genes Dev.*, **3**, 2083–2090.
Hartings,H., Maddaloni,M., Lazzaroni,N., Di Fonzo,N., Motto,M., Salamini,F. and Thompson,R. (1989) *EMBO J.*, **8**, 2795–2801.
Harvey,R.P., Robins,A.J. and Wells,J.R.E. (1982) *Nucleic Acids Res.*, **10**, 7851–7863.
Herrmann,A., Schulz,W. and Hahlbrock,K. (1988) *Mol. Gen. Genet.*, **212**, 93–98.
Hoeffler,J.P., Meyer,T.E., Yun,Y., Jameson,J.L. and Habener,J.F. (1988) *Science*, **242**, 1430–1433.
Katagiri,F., Lam,E. and Chua,N.-H. (1989) *Nature*, **340**, 727–730.
Kawata,T., Nakatsuka,A., Tabata,T. and Iwabuchi,M. (1989) *Biochem. Biophys. Res. Commun.*, **164**, 387–393.
LaBella,F., Sive,H.L., Roeder,R.G. and Heintz,N. (1988) *Genes Dev.*, **2**, 32–39.
Lam,E., Benfey,P.N., Gilmartin,P., Fang,R.-X. and Chua,N.-H. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 7890–7894.
Lam,E., Katagiri,F. and Chua,N.-H. (1990) *J. Biol. Chem.*, **265**, 9909–9913.
Lee,K.A.W. and Green,M.R. (1987) *EMBO J.*, **6**, 1345–1353.
Lin,Y.-S. and Green,M.R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 3396–3400.
Lin,Y.-S., Carey,M., Ptashne,M. and Green,M.R. (1990) *Nature*, **345**, 359–361.
Liu,F. and Green,M.R. (1990) *Cell*, **61**, 1217–1224.
Lycett,G.W., Croy,R.R.D., Shirsat,A.H., Richards,D.M. and Boulter,D. (1985) *Nucleic Acids Res.*, **13**, 6733–6743.
Maekawa,T., Sakura,H., Kanei-Ishii,C., Sudo,T., Yoshimura,T., Fujisawa,J., Yoshida,M. and Ishii,S. (1989) *EMBO J.*, **8**, 2023–2028.
Mermod,N., O'Neill,E.A., Kelly,T.J. and Tjian,R. (1989) *Cell*, **58**, 741–753.
Mikami,K., Tabata,T., Kawata,T., Nakayama,T. and Iwabuchi,M. (1987) *FEBS Lett.*, **223**, 273–278.
Mikami,K., Takase,H., Tabata,T. and Iwabuchi,M. (1989a) *FEBS Lett.*, **256**, 67–70.
Mikami,K., Sakamoto,A., Takase,H., Tabata,T. and Iwabuchi,M. (1989b) *Nucleic Acids Res.*, **17**, 9707–9717.
Nakayama,T., Ohtsubo,N., Mikami,K., Kawata,T., Tabata,T., Kanazawa,H. and Iwabuchi,M. (1989) *Plant Cell Physiol.*, **30**, 825–832.
O'Neill,E.A., Fletcher,C., Burrow,C.R., Heintz,N., Roeder,R.G. and Kelly,T.J. (1988) *Science*, **241**, 1210–1213.
Philipps,G., Chaubet,N., Chaboute,M.-E., Ehling,M. and Gigot,C. (1986) *Gene*, **42**, 225–229.
Roesler,W.J., Vandenbark,G.R. and Hanson,R.W. (1988) *J. Biol. Chem.*, **263**, 9063–9066.
Schmidt,R.J., Burr,F.A., Aukerman,M.J. and Burr,B. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 46–50.
Schulze-Lefert,P., Dangl,J.L., Becker-Andre,M., Hahlbrock,K. and Schulz,W. (1989) *EMBO J.*, **8**, 651–656.
Singh,K., Dennins,E.S., Ellis,J.G., Llewellyn,D.J., Tokuhisa,J.G., Wahleithner,J.A. and Peacock,W.J. (1990) *Plant Cell*, **2**, 891–903.
Studier,F.W. and Moffatt,B.A. (1986) *J. Mol. Biol.*, **189**, 113–130.
Tabata,T., Sasaki,K. and Iwabuchi,M. (1983) *Nucleic Acids Res.*, **11**, 5865–5875.
Tabata,T., Fukasawa,M. and Iwabuchi,M. (1984) *Mol. Gen. Genet.*, **196**, 397–400.
Tabata,T., Takase,H., Takayama,S., Mikami,K., Nakatsuka,A., Kawata,T., Nakayama,T. and Iwabuchi,M. (1989) *Science*, **245**, 965–967.
Tanaka,M., Grossniklaus,U., Herr,W. and Hernandez,N. (1988) *Genes Dev.*, **2**, 1764–1778.
Tokuhisa,J.G., Singh,K., Dennis,E.S. and Peacock,W.J. (1990) *Plant Cell*, **2**, 215–224.
Velten,J., Hain,R. and Schell,J. (1984) *EMBO J.*, **3**, 2723–2730.
Vinson,C.R., LaMarco,K.L., Johnson,P.F., Landschulz,W.H. and McKnight,S.L. (1988) *Genes Dev.*, **2**, 801–806.
Werneke,J.M., Chatfield,J.M. and Ogren,W.L. (1989) *Plant Cell*, **1**, 815–825.
Wu,S.-C., Bogre,L., Vincze,E., Kiss,G.B. and Dudits,D. (1988) *Plant Mol. Biol.*, **11**, 641–649.

Received on January 14, 1991; revised on February 14, 1991

Note added in proof

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X56781 (for HBP-1a) and X56782 (for HBP-1b).