JMB

Functional Dissection of the Plant-specific SBP-Domain: Overlap of the DNA-binding and Nuclear Localization Domains

Rainer P. Birkenbihl, Guido Jach, Heinz Saedler and Peter Huijser*

Max Planck Institute for Plant Breeding Research, Carl-von-Linné Weg 10, 50829 Cologne Germany

SBP-domain proteins are plant-specific putative transcription factors. They all contain the highly conserved 76 amino acid residue SBP-domain, shown to bind specifically to related motifs in the *Antirrhinum majus SQUA* promoter and the orthologous *Arabidopsis thaliana AP1* promoter. The structural basis for this sequence-specific binding of DNA are two Zn-finger like structures formed by the coordination of two zinc ions by conserved cysteine and histidine residues. Amino acid exchanges of the cysteine residues involved revealed that each of the Zn^{2+}-coordinating structures is essential for DNA binding. By random target-site selection studies, it is shown that the palindromic GTAC core motif is essential for efficient DNA binding with additional nucleotides preferred by different SBP-domain proteins. Despite their different functions and origin from plants at different evolutionary distances, the mode of DNA binding is conserved from the single-cell algae *Chlamydomonas reinhardtii* to the moss *Physcomitrella patens* and higher plants. At the C-terminal end of the SBP-domain, a putative bipartite nuclear localization signal is located, which overlaps with the DNA-binding domain, in particular with the second Zn^{2+}-binding structure. By immunolocalization of SPL3 and transient expression of SBP-green fluorescent protein fusion proteins in plant cells, it is shown that this nuclear localization signal is functional. Exchange of a highly conserved serine next to the nuclear localization signal by aspartate, which may mimic phosphorylation, resulted in a decreased nuclear import (SPL8), while DNA binding *in vitro* was abolished completely. In contrast, exchange by alanine increased nuclear import and left DNA binding intact. This suggests that the function of SBP-domain proteins is also regulated by post-translational modification on the levels of nuclear import and DNA binding.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: plant transcription factor; SBP-domain protein; DNA binding; nuclear import; Zn-binding structure

*Corresponding author

Introduction

SBP-domain proteins form a diverse family of putative transcription factors, which are defined by harboring a highly conserved domain of 76 amino acid residues, the SBP-domain. They are found in all green plants, from single-cell *Chlamydomonas* to

multicellular higher plants. Since they are not present in prokaryotes, fungi or animals, they seem to be specific for plant development. SBP-domain proteins were first isolated from *Antirrhinum majus* in an *in vitro* approach to identify regulators of the MADS-box gene *SQUAMOSA* (*SQUA*) by their ability to bind to its promoter. Therefore, they were called *SQUA* promoter-binding proteins (SBP).[1]

In *Arabidopsis thaliana*, the SBP proteins are encoded by a heterogeneous family of 16 genes (*SPL*-genes), which can be divided in subfamilies on the basis of sequence, size and gene structure. The largest genes (*SPL1, SPL7, SPL12, SPL14* and *SPL16*)

are expressed constitutively, while the mid-sized and small genes are upregulated mainly in flower development.[2–4] More hints that the smaller SPL proteins are involved in the process of flowering come from the findings that over-expression of *SPL3* results in early flowering,[2] and *SPL8* insertion mutants result in aberrant pollen sack development in *Arabidopsis*.[5] SPL14, which with a length of 1035 amino acid residues belongs to the large proteins, was reported to play a role in plant pathogen response (sensitivity to fumonisin B1) and in development of the plant architecture, especially in leaf morphology.[6] Beside these *Arabidopsis SPL* genes, only one other SBP-box gene, *LIGULELESS1* (*LG1*) from *Zea mays*, has been reported to exhibit a mutant phenotype, i.e. lacking the ligule at the boundary between blade and sheath of the leaf.[7]

Recently, the *SPL* genes attracted attention because the smaller ten genes, except *SPL8*, contain the only miRNA target sites for miRNA156 and miRNA157,[8] which provides a means to down-regulate them simultaneously, as shown recently in constitutive MIR156 over-expressor transgenic plants.[9]

To fulfill their role as transcription factors, the SBP proteins have to be imported into the nucleus and eventually to bind in a sequence-specific mode to DNA. Both functions seem to be covered by the SBP-domain. For the *A. majus* SBP1 and SBP2 proteins and the putative *Arabidopsis* orthologues, it was shown that they bind, besides a defined motif in the *SQUA*-promoter, to a similar sequence in the *AP1* promoter, the presumed *Arabidopsis SQUA* orthologue. Using truncated versions of the respective proteins in *in vitro* DNA-binding studies led to the finding that the SBP-domain is necessary and sufficient for sequence-specific DNA binding.[1,2]

The structures of the SBP-domains of SPL4 and SPL7 in solution were resolved recently by NMR.[10] The domain contains two, non-interleaved Zn^{2+}-coordinating structures, which are formed by eight conserved cysteine and histidine residues. Modeling of the bound *SQUA*-promoter motif to the structure of the SBP-domain suggested that the DNA-binding domain covers almost the whole SBP-domain and that mainly the basic amino acid residues, arranged to form a positively charged face, are involved in general and specific DNA binding. The stoichiometry of DNA binding was shown to be 1:1.

Regulation of subcellular compartmentalization generally represents an important level of controlling the function of a protein. In the case of a transcription factor, this concerns in particular its nuclear import with the help of a nuclear localiz-ation signal (NLS). Here, the most common regulation mechanism is masking the NLS by post-translational modification close to the NLS. By this, efficient binding of importin α to the NLS is prevented and nuclear import is disturbed.[11–13] The nuclear import of the SBP-domain proteins seems to be promoted by the SBP-domain. Its C-terminal end harbors two clusters of basic amino acid residues and fits perfectly the consensus sequence for an NLS of the bipartite type.[14,15]

Here, we have analyzed the structural base of the DNA-binding activity and the functionality of the putative NLS by making use of SBP-proteins carrying amino acid exchanges in the SBP-domain. DNA binding was monitored by electrophoretic mobility-shift assays (EMSAs) and subcellular localization investigated by transient expression of SPL proteins in plant cells. The DNA sequence specificity was analyzed in more detail by random oligonucleotide selection assays. By this, we could confirm the palindromic tetranucleotide GTAC to be essential for DNA binding by SPL1, SPL3 and SPL8. By employing the *AP1* promoter motif and an SBP target sequence from *Chlamydomonas*, together with SBP-domains from *Arabidopsis*, *Physcomitrella* and *Chlamydomonas* produced in *Escherichia coli*, we demonstrate that DNA binding and sequence specificity of SBP-domains is conserved from single-cell algae to higher plants.

The SBP-domain structure has been discussed only with respect to DNA binding, despite the fact that the SBP-domain harbors a putative NLS.[10] Here, we show that this NLS is functional and that some of the amino acid residues, which were discussed on the basis of the domain structure to be important for general and sequence-specific DNA binding, are essential also for nuclear import. This structural and functional overlap of the DNA binding and nuclear localization domains is under-lined by the finding that both functions may be regulated by modification of the same conserved serine residue close to the NLS. Besides regulating transcription of the genes by their promoters and control of the transcripts by miRNAs, here another, post-translational level of regulating SBP protein activity seems to exist.

Results and Discussion

The SBP-domain has been highly conserved during plant evolution

SBP-genes exist only in green plants, from single-cell algae to the most highly developed plants. There is sequence information for more than 160 full-size SBP-domains in electronic databases. In all cases, the SBP-domains are very similar, with a high or absolute conservation at certain positions, as demonstrated in Figure 1. Here, 166 sequences from 52 species were aligned. Maize *Liguleless1* and its presumed orthologue in rice were left out, because glycine inserted at position 40 had resulted in a gap in the alignment. The best-conserved amino acid residues are cysteine (at positions 4, 9, 26, 45, 48 and 64) and histidine residues (at positions 29 and 52) used to coordinate two zinc ions. Similarly well conserved are the basic amino acid residues supposed to be involved in general and specific DNA binding,[10] and in nuclear translocation of the
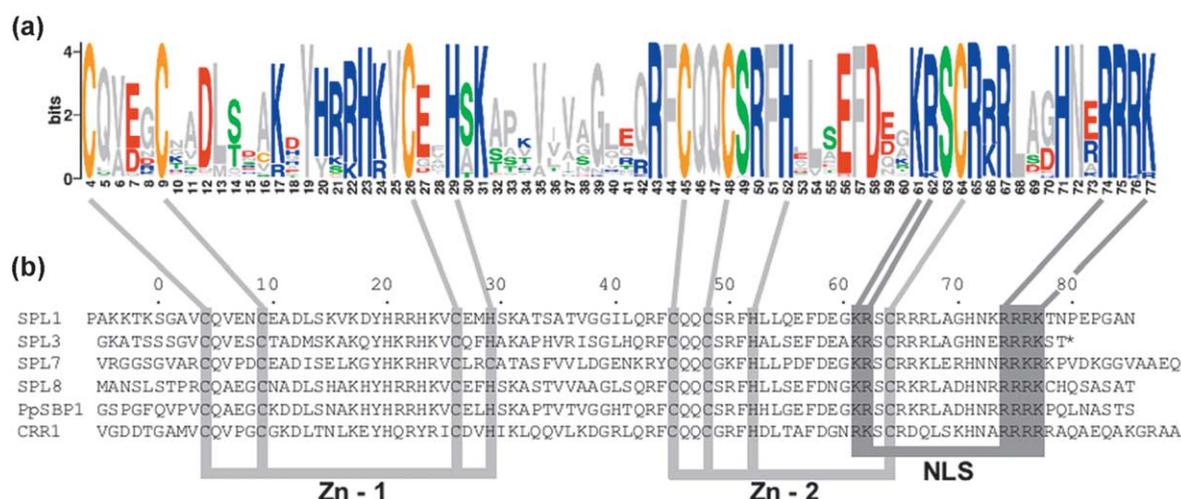
Figure 1. Sequence conservation in the SBP-domain. (a) The 166 SBP-domain sequences from 52 species were aligned and the sequence conservation at a particular position expressed as a stack of letters using the program WebLogo (WebLogo.berkeley.edu).[32] The overall height of each stack indicates the sequence conservation at that position (measured in bits), whereas the height of symbols within the stack reflects the relative frequency of the corresponding amino acid. Cysteine residues are in yellow, basic amino acid residues involved in DNA binding and nuclear localization are in blue, acidic amino acid residues are in red, and threonine and serine are in green. The numbering is according to the numbers in the NMR structure.[10] (b) Shown are the amino acid sequences of the SBP-domains used in DNA-binding assays: SPL1, SPL3, SPL7, SPL8 from *Arabidopsis*, PpSBP1 from *Physcomitrella* and CRR1 from *Chlamydomonas*. The asterisk (*) at the end of the SPL3 sequence indicates its natural stop. The conserved basic amino acids of the NLS are shaded dark grey, while the two Zn-coordinating structures (Zn-1 and Zn-2) with the involved cysteine and histidine residues are in light grey.

proteins (at positions 61, 62, 74, 75, 76 and 77). Among other well-conserved residues is serine 63, which is identified by the prediction program NetPhos 2.0 as a very likely phosphorylation site.

SBP-domain proteins recognize and bind specifically to the GTAC core motif

Knowledge of their optimal DNA-binding sequence is very helpful to identify candidate target genes of transcription factors. From previous work, it is known that SBP1 and SBP2 from *A. majus*,[1] *Arabidopsis* SPL3,[2] SPL4 and SPL7,[10] and birch BpSPL1[16] bind *in vitro* sequences encompassing a ten nucleotide motif found in common with the *A. majus SQUA* and the orthologous *Arabidopsis AP1* promoters. To investigate which nucleotides are really needed for binding, and if there are different sequence specificities among different SBP proteins, we performed random oligonucleotide selection based on EMSAs. Since we were not able to keep whole proteins under physiological conditions in solution, and since it was known that the SBP-domain alone is sufficient for sequence-specific binding, we used recombinant SBP-domains for the DNA-binding assays. The SBP-domains of SPL1, SPL3 and SPL8, representing three major subclasses of the *SPL* genes, were expressed in *E. coli* with extensions of about ten residues on both sides, except for SPL3, which naturally terminates two residues behind the SBP-domain (Figure 1(b)). The peptides were purified by affinity chromatography making use of a His$_6$ tag.

First, the minimal DNA motif size still bound with full efficiency by the SBP-domains was determined. For this we used the *AP1* promoter-derived binding site. Without any loss of binding strength by the SPL1 SBP-domain, the DNA could be shortened to a 15 bp fragment (AP1-15, sequence in Figure 3(a)), which was used as positive control in all EMSAs unless stated otherwise. Removing one more base-pair from both sides still had no effect, but further shortening from either side resulted in consecutive loss of binding strength until at 10 bp DNA-binding was lost completely (not shown). As judged from the gels, the unbound DNA substrate was still double-stranded, but it was not clear if essential nucleotides of the binding motif were lost or the DNA molecule was just too short to be bound by the SBP-domain.

On the basis of this finding, a random primer selection experiment was carried out using 62 nt long oligonucleotides containing a random central string of 16 nucleotides flanked by defined ends for PCR amplification and cloning. Three consecutive rounds of DNA binding, recovering the protein–DNA complexes from EMSA gels and amplification of bound DNA by PCR were done before cloning for sequencing. At that stage, more than 50% of the DNA was bound by the proteins. In all, 96 clones from SPL1 and SPL3 binding were sequenced, both resulting in a high preference for a GTAC core motif, which is contained also in the bound *AP1* promoter motif. Unfortunately, one nucleotide of this motif (G or C) was provided in more than 50% of the sequences by the fixed sequences flanking the
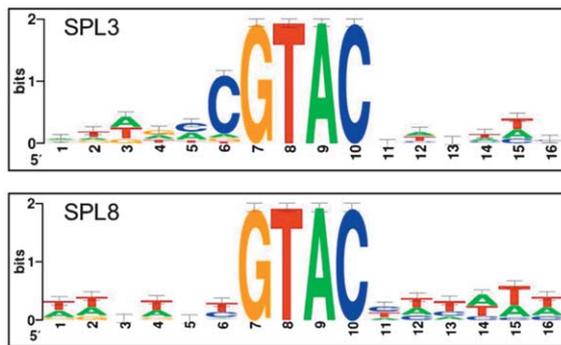
Figure 2. SBP-domains bind to the GTAC core sequence. A random primer selection experiment was carried out with oligonucleotides (see Materials and Methods), containing a random central string of 16 nucleotides with GT at positions 7 and 8. Three consecutive rounds of DNA binding, recovering the protein-DNA complexes from EMSA gels and amplification of bound DNA by PCR were done before cloning and sequencing. The 67 sequences from SPL3 and 73 sequences from SPL8 binding were aligned, and the sequence conservation at a particular position expressed as a stack of letters using the program WebLogo (WebLogo.berkeley.edu).[32] The overall height of each stack indicates the sequence conservation at that position (measured as content of information in bits), whereas the height of symbols within the stack reflects the relative frequency of the corresponding nucleotide.

random part of the oligonucleotides, thus preventing us from learning more about the distribution frequency of adjacent nucleotides.

In order to locate the binding motif in the middle of the random stretch of nucleotides, and with the knowledge that the GTAC seemed to be essential for binding, we performed a second experiment with oligonucleotides containing GT at positions 7 and 8 of the random part and three consecutive A or T nucleotides as first nucleotides in the fixed sequence (Figure 2). This time, the GT was again completed to the palindromic GTAC core motif for SPL3 (67 sequences analyzed) and SPL8 (73 sequences) with a high preference for at least one more C at the 5'-end of the motif for SPL3 (83% *versus* 32% for SPL8) (Figure 2). No more conclusions concerning a preference for specific nucleotides at other positions in the binding site could be drawn, except a slight preference for A or T nucleotides flanking the core motif. Binding assays with selected sequences from this experiment confirmed that SPL3 was indeed more selective. While the SBP-domains of SPL1 and SPL8 bound to the same sequences with similar efficiencies, SPL3 bound only to a subset (not shown).

With the help of the program Patmatch† the *Arabidopsis* genomic sequence ("GenBank whole

† http://www.arabidopsis.org/cgi-bin/patmatch/nph-patmatch.pl

genome" and "Locus Upstream Sequences-1000 bp") was searched for the GTAC and the CGTAC motif, and statistics performed in order to find significance (frequency and distribution relative to translation start) for this binding motif. It is remarkable that in the genome both motifs are under-represented by a factor of about 2, but as a dimer with a spacing of up to 30 bp the palindromic tetramer appears roughly 25 times more frequent in the promoter regions than expected. For the pentamer CGTAC, the situation is more complicated, because now the motif has an orientation, and anyway the tetranucleotide is included.

For the motifs, we established long candidate gene lists, which will be compared to gene lists derived from other experiments with the aim to identify target genes, like crosslinked-chromatin immunoprecipitation (X-CHIP)[17] or transcript profiling in response to *SPL* gene activation. Thus, in the promoters of putative target genes, derived from microarray hybridization with RNA isolated after SPL8 induction, the GTAC motif was significantly over-represented compared to a random gene list (S. Schwarz & P.H., unplublished results). In the case of other DNA-binding proteins known to interact with SBP-domain proteins, identified for example by yeast two-hybrid-analysis, promoter regions of putative target genes can be searched for the combined presence of both binding sequences.

DNA-binding by the SBP-domain proteins has been conserved during evolution

SBP-domain proteins, some of which are suggested to be transcription factors involved in plant development,[5,6] are found in all green plants, in single-cell *Chlamydomonas* with relative little differentiation during development, as well as in highly complex organisms like trees. Because the SBP-domain is extremely well conserved in all of these otherwise heterogeneous proteins, the question arises of whether the biological function of this domain in all plants is related and to what extent the biochemical mechanisms are the same.

To compare the DNA-binding specificity of SBP proteins from evolutionarily very distant plant species, we chose, beside SPL1, SPL3, SPL7 and SPL8, representing different subfamilies of the SPL proteins from *Arabidopsis*, an SBP-domain protein from *Chlamydomonas* CRR1,[18] (this information was generously provided before publication by Janette Kropat and Sabeeha Merchant, UCLA, USA) and one from moss *Physcomitrella* PpSBP1 (our unpublished results). The SBP-domains of these proteins, to the extent indicated in the alignment in Figure 1, were expressed in *E. coli*, purified and subjected to DNA-binding assays. Increasing amounts of the peptides were incubated either with the *Arabidopsis*-derived 15 bp *AP1* promoter fragment or with a copper response element (CuRE) from *Chlamydomonas* that, except for the GTAC tetranucleotide, exhibits no homology to the *AP1* substrate (Figure 3). In *in vivo* expression studies of this

(a) *A.t.* AP1-15:
TGGTCC**GTAC**AATGT
ACCAGG**CATG**TTACA

(b) *C.r.* Cyc6 CuRE:
GAGCGCACCT**GTAC**CTGCCAGTAT
CTCGCGTGGA**CATG**GACGGTCATA

(c)

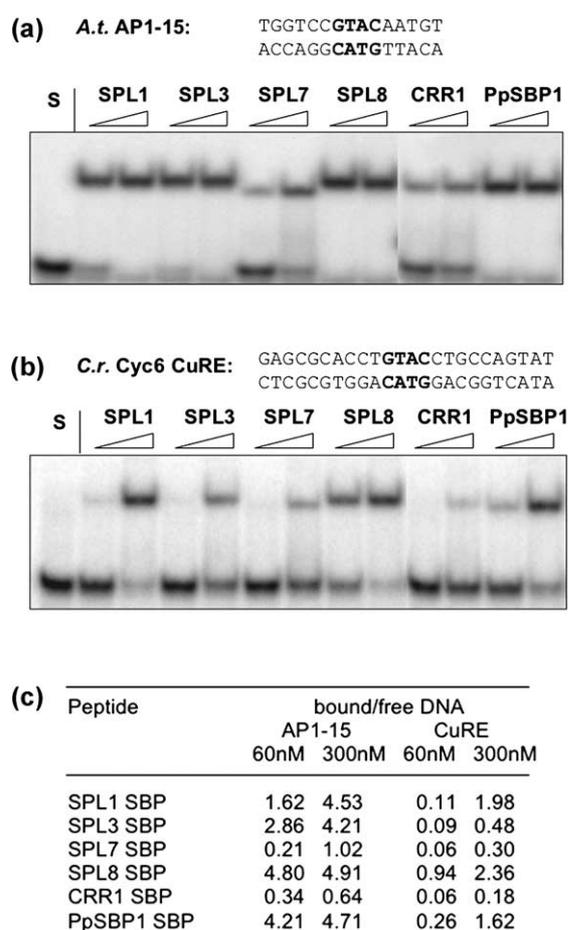| Peptide	bound/free DNA			
	AP1-15		CuRE	
	60nM	300nM	60nM	300nM
SPL1 SBP	1.62	4.53	0.11	1.98
SPL3 SBP	2.86	4.21	0.09	0.48
SPL7 SBP	0.21	1.02	0.06	0.30
SPL8 SBP	4.80	4.91	0.94	2.36
CRR1 SBP	0.34	0.64	0.06	0.18
PpSBP1 SBP	4.21	4.71	0.26	1.62

Figure 3. Conservation of sequence-specific DNA binding during evolution. Recombinant SBP-domains from different species (SPL1, SPL3, SPL7 and SPL8 from *Arabidopsis*, PpSBP1 from *Physcomitrella* and CRR1 from *Chlamydomonas*) were analyzed for DNA binding using radio-labeled DNA substrates (a) AP1-15 from *Arabidopsis*, and (b) Cyc6 CuRE from *Chlamydomonas*, the respective sequences containing the GTAC core motif (bold) are indicated: 10 ng (60 nM) and 50 ng (300 nM) of each peptide was used for EMSA. The first lane (S) contains DNA substrate without protein. (c) The binding efficiencies from (a) and (b) are quantified by the ratio of bound to free DNA for both substrates (AP1-15 and CuRE) at peptide concentrations of 60 nM and 300 nM.

CuRE had been shown to be essential for a graduated response, mediated by CRR1, to the concentration of Cu in the growth medium.[19]

As shown in Figure 3, all SBP-domains bound to both DNA substrates, but with different efficiencies. In general, it can be stated that for all peptides, binding to the AP1-derived motif was stronger than to the CuRE. Additionally, the peptides that bound strongest to the *AP1* DNA also bound best to the CuRE. An exception from this was SPL3, which was, together with SPL1, SPL3, SPL8 and PpSBP1, among the best binders to AP1-15, but apart from them among the weak binders of the CuRE. This reflects the higher level of sequence selectiveness observed in the random primer selection experiment.

Interestingly, SPL7 and CRR1, the two SBP-domains with the least similarity to the other peptides, but the greatest similarity to each other, exhibited a significantly lower affinity to both DNAs.

The relatively weak binding of CRR1 to its own putative target might be due to the fact that, here, interaction has to function in a graduated mode. In this mechanism, a second CuRE in the same promoter is probably involved,[19] which was bound also in EMSAs, but with an even lower level of efficiency (not shown).

SBP-domain binding to DNA is Zn^{2+}-dependent and established by two Zn^{2+}-binding structures

The alignment of SBP-domains shows that a number of conserved Cys and His residues are spaced in a way known from Zn^{2+}-coordinating structures, which, among other possible functions, may bind DNA.[20] The NMR structure of the SBP-domain revealed that indeed two Zn-finger like structures are formed: Zn-structure 1 is formed by the amino acid residues C4, C9, C26 and H29 (numbers according to Figure 1 and numbering in the published NMR structure);[10] and Zn-structure 2 is formed by C45, C48, H52 and C64. However, it has not been shown whether Zn^{2+} is really needed for DNA binding, or if both Zn^{2+}-binding structures are involved.

To prove the requirement of Zn^{2+} for DNA binding, the SBP-domain of SPL1 was depleted of Zn^{2+} using EDTA. Simply adding EDTA to the protein did not impair DNA binding, a hint that, if required for DNA binding, the Zn^{2+} is bound tightly in the molecule. This was found also with the SBP proteins from *A. majus*[1] and SPL14, when its binding to genomic DNA was tested.[6] Addition of EDTA to the native protein was sufficient to disturb the formation of the Zn-structures in SPL4 and SPL7,[10] but the consequences for DNA binding have not been tested. In our hands, denaturation with 8 M urea in the presence of EDTA removed the Zn^{2+}. After renaturation by dialysis, the proteins treated with or mock-treated without EDTA, were submitted to EMSA. While the peptide treated with EDTA did not bind DNA, the peptide without EDTA treatment still bound the DNA with the original level of efficiency (Figure 4). Addition of Zn^{2+} to the depleted peptide restored DNA binding, while Mg^{2+} had no effect. This demonstrates clearly the Zn^{2+}-dependency of DNA binding by SBP proteins.

To get more information about the involvement of the single Zn-structures in DNA-binding, amino acid residues in the SPL1-SBP-domain participating in the coordination of the zinc ions were exchanged for alanine and the respective polypeptides tested by EMSA (Figure 5). For Zn-structure 1, the exchange of C9 or H29 resulted in almost complete loss of DNA binding. Only at higher concentrations of protein could some DNA binding be detected. In Zn-structure 2, exchange of the outer C45 or C64 impaired DNA binding completely, as judged from
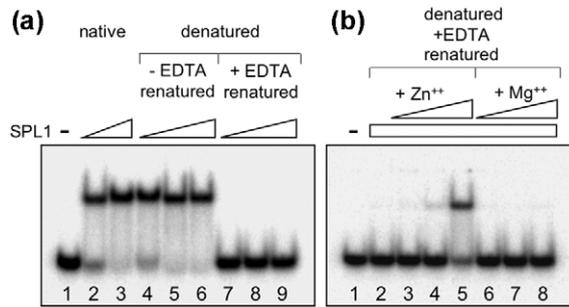
Figure 4. Zn-dependence of DNA binding by SBP-domain proteins. (a) Increasing amounts (2 ng, 10 ng, and 50 ng) of native peptide SPL1-SBP was denatured with 8 M urea, mock-treated (lanes 4–6) or treated with 50 mM EDTA (lanes 7–9), renatured by dialysis and submitted to EMSA. Also shown is radio-labeled AP1-15 substrate without peptide ($-$) and DNA binding of native untreated peptide (2 ng and 10 ng, lanes 2 and 3). (b) Zn^{2+}-depleted peptide (50 ng) was preincubated with increasing concentrations of Zn^{2+} (10 μM, 100 μM and 1000 μM; lanes 3–5) or Mg^{2+} (lanes 6–8) and analyzed by EMSA.

the gel, but exchange of H52, also participating in Zn^{2+} coordination according to the NMR structure, led to a reduction in binding efficiency to only about half. This could mean that the structure is still stabilized to some extent by the three remaining cysteine residues and other intramolecular forces. This result underlines the uncommon character of this second Zn-structure, which contains a very short middle and an unusually extended C-terminal knuckle. In summary, the data presented demonstrate that the integrity of both Zn-structures together is needed for efficient DNA binding.
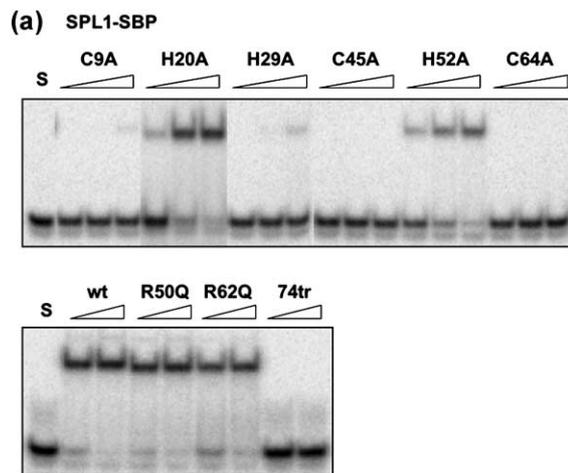
Other amino acid residues involved in DNA binding

The NMR structure of the SBP-domain shows that all conserved basic amino acids form a positively charged surface involved in binding the negatively charged DNA. Modeling the *AP1* recognition sequence to the structure provided further insight into the role of single amino acids in general and sequence-specific DNA binding.[10] To get more experimentally based information about the interaction with DNA, we analyzed additional SBP-domain mutant peptides. Truncation of the SPL1 SBP-domain at position 74, which removes the basic amino acid residues R75, R76 and K77 close to the C-terminal end of the domain, led to a complete loss of DNA binding. Exchanging single conserved basic amino acid residues, like H20A (in SPL1) within Zn-structure 1, R50Q (in SPL1), which is encoded by the codon generated upon splicing of the conserved intron in the SBP-box, and R62Q (in SPL1 and SPL3), which belongs to the first basic cluster of the bipartite NLS, had only little effect on

the apparent binding strength in EMSAs (Figures 5 and 6). Whether the mutations relaxed the sequence specificity could not be deduced from those experiments.

Sequence conservation in proteins is often a sign for functional importance of the respective region. The finding that destroying Zn-structure 1 at the N-terminal end of the SBP-domain as well as removing basic amino acids from the C-terminal end of the SBP-domain both abolished DNA binding, indicates that the DNA-binding domain encompasses the whole SBP-domain. Thus, the boundaries of the conserved SBP-domain seem to have been determined by the evolutionarily conserved mechanism of DNA binding to an also conserved DNA-binding site.

With respect to DNA binding, we also tested mutations in serine S63, an evolutionarily highly

Figure 5. DNA binding of SPL1 SBP-domain carrying selected amino acid exchanges. (a) Recombinant peptides of the SPL1 SBP-domain carrying substitutions in the indicated amino acids (in the one-letter code and position in the SBP-domain) were analyzed by EMSA using AP1-15 as DNA substrate. Beside the amino acid exchanges, a peptide truncated behind R74 (74tr) was tested. Increasing amounts, 2 ng (12 nM), 10 ng (60 nM) and 50 ng (300 nM) upper panel, and 10 ng (60 nM) and 50 ng (300 nM) lower panel, of the respective peptides were used in the assays. Lanes labeled (S) contained DNA substrate without protein. (b) The binding efficiencies from (a) are quantified by the ratio of bound to free DNA for peptide concentrations of 60 nM and 300 nM.

(b)

Peptide	bound/free DNA	
	60nM	300nM
SPL1 SBP wt	3.55	7.51
SPL1 SBP C9A	0.09	0.12
SPL1 SBP H20A	1.61	2.58
SPL1 SBP H29A	0.11	0.19
SPL1 SBP C45A	0.08	0.09
SPL1 SBP H52A	0.97	1.72
SPL1 SBP C64A	0.08	0.08
SPL1 SBP R50Q	3.73	5.02
SPL1 SBP R62Q	1.66	3.15
SPL1 SBP 74tr	0.02	0.02

conserved residue and a possible target for post-translational modification (see below in the context of nuclear import). In SPL3 and SPL8 (which were also tested for nuclear import), S63 was changed either to alanine, thus destroying the putative modification site, or to aspartic acid, which may mimic a phosphorylated serine. The resulting peptides were analyzed by EMSA using again AP1-15 as substrate. For both SPL3 and SPL8, S63A was still proficient in binding DNA, indicating that the serine as such is not needed for DNA binding. Conversely, the S63D exchange resulted in complete loss of DNA binding by both SBP-domains (Figure 6). From our current knowledge, we cannot say whether the introduction of a negative charge at this position impaired DNA binding merely by repulsion of the also negatively

Figure 6. DNA-binding analysis of SPL3 and SPL8 SBP-domains carrying selected amino acid exchanges. Increasing amounts, 2 ng (12 nM), 10 ng (60 nM) and 50 ng (300 nM) of the (a) SPL3 and (b) SPL8 SBP-domains containing amino acid exchanges (indicated in the one-letter code and position in the SBP-domain) were assayed by EMSA using AP1-15 as DNA substrate. Lanes labeled (S) contain DNA substrate without protein. In (c) the binding efficiencies from (a) and (b) are quantified by the ratio of bound to free DNA for peptide concentrations of 60 nM and 300 nM.

charged DNA, or if the mutation has structural consequences leading to loss of DNA-binding activity. In any case, this result suggests strongly that phosphorylation of S63 will also lead to a loss of DNA binding.

The SBP-domain contains a functional nuclear localization signal

To fulfill their role as transcription factors, SBP-domain proteins have to be translocated into the nucleus. The nuclear import is usually initiated by binding of importin α to a certain amino acid sequence, the NLS.[11] Close to its C terminus, the SBP-domain contains a bipartite NLS, which, together with or even as part of the DNA-binding domain, has been conserved during evolution. Two basic amino acid residues (K61 and R62 for all SPL proteins) are followed by an 11 residue spacer and a second block of four basic amino acid residues (RRRK; RRKR in SPL7), thus perfectly fitting the consensus sequence described.[14] At least the short SPL proteins from *Arabidopsis* do not contain a second candidate NLS, as judged by sequence analysis. The others, especially the large SPL proteins with more than 1000 residues, contain some additional basic regions, which also might promote nuclear import, functioning as monopartite NLS. It has been reported for the maize SBP-domain protein *LIGULELESS* (LG1) that it still is imported into the nucleus when the SBP-domain was removed.[7] The nuclear localization has been demonstrated also for full-length SPL14, which belongs to the large proteins, but the functional NLS was not mapped.[6]

To test the functionality of the SBP-NLS, the subcellular distribution of SPL3 was investigated by immunolocalization using purified SPL3-directed antibodies. We compared floral stigmatic tissues from wild-type *Arabidopsis* plants with transgenic plants over-expressing SPL3 ectopically.[2] While no SPL3 protein was detectable in wild-type cells, clear nuclear signals were obtained for the plant over-expressing SPL3. Here, in contrast to the wild-type cells, the high SPL3 content in the nuclei together with the bound antibody also resulted in a strong quenching of the nuclear 4',6-diamidino-2-phenylindole (DAPI) staining (Figure 7).

To further investigate the functionality of the SBP NLS, we performed nuclear import experiments in tobacco BY-2 protoplasts with transiently expressed SPL proteins. For these experiments we chose SPL3, SPL4 and SPL5 because of their small size. Especially SPL3, with only 131 residues, consists mainly of the SBP-domain and virtually no extension at its C terminus. Additionally, we analyzed with respect to nuclear import, SPL9, SPL15 and SPL8, the only *SPL* gene with a described mutant phenotype[5] at the time we performed the experiments. For these studies, we destroyed the putative NLS by C-terminal truncations, which removed the C-terminal basic cluster or the complete NLS. Both types of mutations were
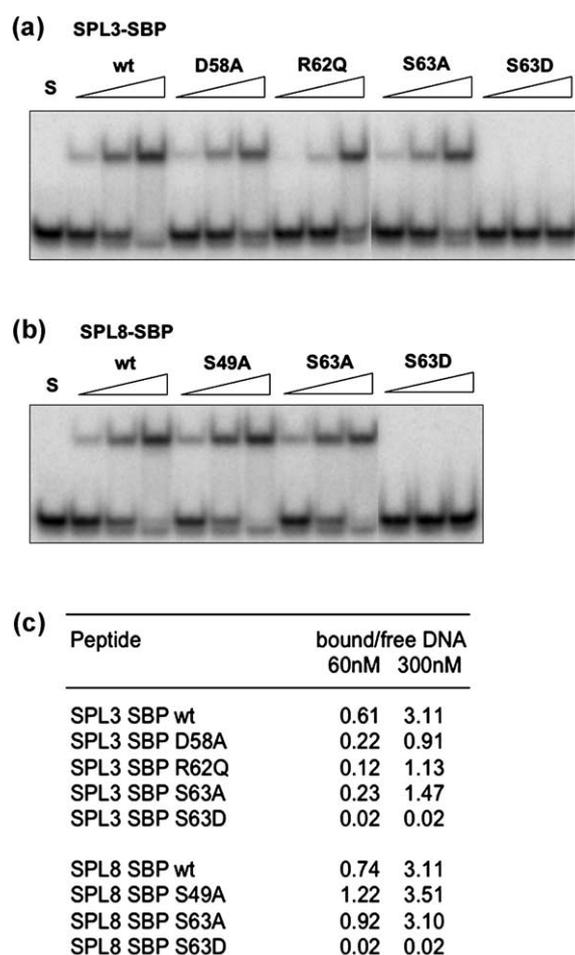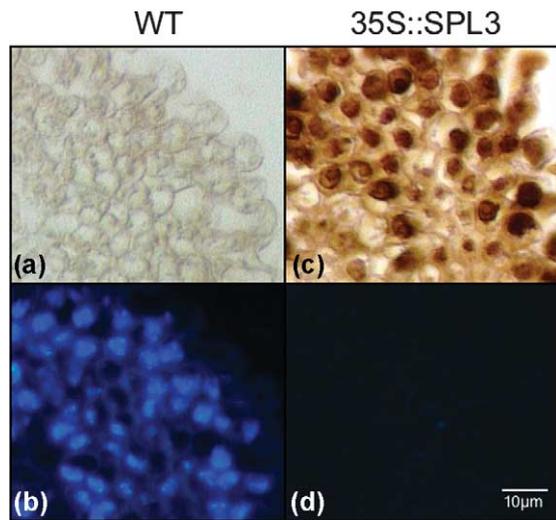
WT   35S::SPL3

Figure 7. Immunolocalization of SPL3 in *Arabidopsis*. Transverse sections of stigmatic tissue of *Arabidopsis* Col 0 (A and B) and plants ectopically over-expressing SPL3 (C and D) were probed with an affinity-purified antibody raised against SPL3. Sections were mounted with DAPI to localize nuclei and analyzed by bright field (A and C) and fluorescent light microscopy (B and D). Notice that in D the DAPI signal is quenched.

expected to result in less efficient nuclear import compared to wild-type.

The proteins were expressed as N-terminal translational fusions to two copies of green fluorescent protein (2×GFP) driven by the constitutive CaMV S35-promoter. Unfortunately 2×GFP, despite not being a nuclear protein itself and with a molecular mass of about 54 kDa, tends to enter the nucleus to a certain extent, thus resulting in a background nuclear signal, even if no functional NLS is contained in the protein to be tested. Because of this, efficiency of nuclear import had to be judged by comparing the intensity of fluorescence in the nucleus with that retained in the cytoplasm.

Nuclear import was monitored 20 h after transformation of tobacco protoplasts. In the transient expression assays, the small SPL proteins SPL3, SPL4 and SPL5, and the midsized SPL9 fused to GFP were imported efficiently into the nucleus with little fluorescence left in the cytoplasm. SPL15 was imported with the highest level of efficiency (Figure 8); SPL8 behaved differently. While having been expressed well, its nuclear import was clearly less efficient. When C-terminally truncated versions of SPL8 were analyzed, truncation after position 65 in the SBP-domain (SPL8 65tr), which removes the second basic cluster of the NLS, and after position 14 (SPL8 14tr), which removes almost the whole SBP-domain, import efficiency dropped to that of GFP alone. Truncations of SPL3 omitting the complete second part of the NLS (SPL3 66tr) or the three C-terminal residues of the second part (SPL3 74tr, not shown), respectively, also resulted in a clear decrease in nuclear translocation compared

to that of the wild-type, but with some import left (Figure 8, Table 1). Together with the immunolocalization of over-expressed SPL3, these results clearly demonstrate the functional relevance of the SBP-domain NLS.

The function of SBP-domain proteins is also regulated at the level of nuclear import

For many cases, it has been described that covalent modification of nuclear proteins, especially close to the NLS, can prevent or promote their import into the nucleus.[12,13,21] Analyzing the
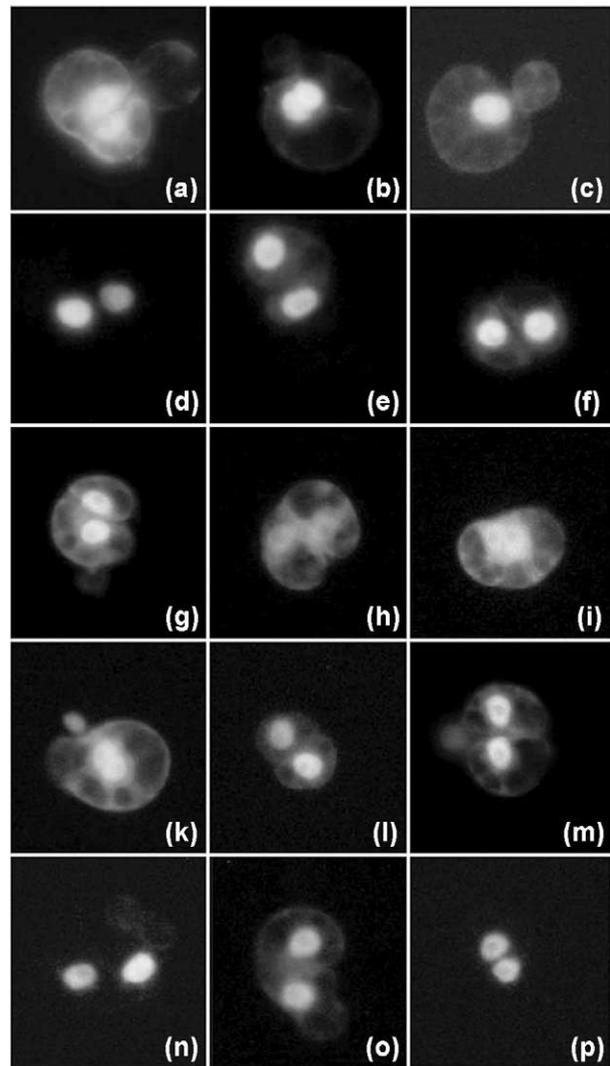

Figure 8. Nuclear import of SPL proteins fused to GFP. Constructs encoding SPL::2×GFP fusion proteins were transformed into tobacco protoplasts for transient expression. After 20 h, GFP distribution in the cells was examined by fluorescence microscopy. Proteins analyzed were: GFP (A), SPL3 wt (B), SPL3 66tr (C), SPL3 C64A (D), SPL4 wt (E), SPL5 wt (F), SPL8 wt (G), SPL8 65tr (H), SPL8 14tr (I), SPL8 S63D (K), SPL8 S63A (L), SPL8 C64A (M), SPL8 SBP (N), SPL9 wt (O), and SPL15 wt (P). Notice that the nuclear import of the SPL3 and SPL8 constructs is quantified in Table 1.

SBP-domain for possible post-translational modification sites identified S63 as a very likely target for phosphorylation (program used, NetPhos 2.0).[22]

To investigate the possible effect on nuclear import of a modification at serine S63, it was replaced in both SPL3 and SPL8 either by alanine (S63A) to prevent potential modification, or by aspartic acid (S63D) to mimic phosphorylation of the serine residue by the size of the side-chain and the charge. This approach has been applied successfully to study the regulation of nuclear import of other proteins.[21] The modified SPL3 and SPL8 proteins were again expressed transiently in tobacco BY-2 protoplasts as translational fusions to GFP. While SLP8 S63D, which mimics phosphorylation of S63, stayed preferentially in the cytoplasm, SPL8 wild-type, which might be phosphorylated at S63, was imported to a moderate level. SPL8 S63A with the putative phosphorylation site destroyed was imported best and was found predominantly in the nucleus (Table 1 and Figure 8). This result suggests, that phosphorylation of S63 (or another kind of modification) negatively influences nuclear import of SPL8.

For SPL3, this mechanism does not seem to apply. Here, SPL3 S63A and SPL3 S63D were imported with the same high level of efficiency as the wild-type protein (not shown). An even better import was achieved when just the SBP-domain of SPL8 (M179-T269 in the native protein) was fused to 2× GFP (Figure 8 and Table 1). Because the S63A and S63D mutations in this construct exhibited similar high-level import efficiencies (not shown), this negative regulation seems to depend on sequences outside the SBP-domain.

We also analyzed if formation or non-formation of Zn-structure 2, which overlaps with the NLS, has consequences for nuclear import. Therefore, we tested for SPL3 and SPL8 cysteine to alanine

Table 1. Constructs encoding SBP-domain proteins fused to 2×GFP under the control of the strong constitutive CaMV 35 S promoter were transformed into tobacco protoplasts for transient expression

Construct	Ratio N/C	n	Pos. in Figure 8
GFP control	1.4 ± 0.2	15	A
SPL3 wt::GFP	11.2 ± 3.4	16	B
SPL3 66tr::GFP	5.5 ± 1.7	15	C
SPL3 C64A::GFP	26.3 ± 14.3	10	D
SPL8 wt::GFP	7.6 ± 1.7	23	G
SPL8 65tr::GFP	1.8 ± 0.2	22	H
SPL8 14tr::GFP	1.6 ± 0.2	23	I
SPL8 S63D::GFP	4.0 ± 0.7	22	K
SPL8 S63A::GFP	10.6 ± 1.7	21	L
SPL8 C64A::GFP	8.7 ± 1.8	22	M
SPL8 SBP::GFP	21.6 ± 10.3	21	N

After 20 h GFP distribution in the cells was monitored by confocal laser scanning microscopy as described in methods. In the Table are listed behind the examined constructs, the average ratio of GFP-intensity in the nuclei *versus* cytoplasm (N/C) with the standard deviation, the number of cells examined (*n*), and the position of a representative cell in Figure 8.

exchanges at position 64, which on one hand is essential for Zn-coordination and on the other hand is situated next to the critical S63. In both proteins, the C64A exchanges resulted in an increase of import efficiency (Figure 8 and Table 1). For SPL8, this increase was only moderate compared to wild-type but it was very strong for SPL3. To exclude a positional effect, the experiment was repeated with SPL3 C45A, which also impairs Zn-structure 2, leading to the same outcome (data not shown). This result might be explained by the finding that the NLS has to be in an extended conformation for efficient binding by importin α,[23] and formation of the Zn-structure disturbs this interaction. Another possibility is that C64 itself might be modified, for example by glutathionylation. Thus, NF-kappa B and AP-1 carry a DNA-binding site containing a well-conserved cysteine residue, flanked by basic amino acids. The DNA-binding activities of both are modulated by SH-modifying agents.[24–26] The difference between SPL3 and SPL8 might reflect the overall better import of SPL3 wild-type and its mutants with impaired NLS.

Overlap of the DNA-binding domain and the NLS

The boundaries of the SBP-domain are determined and conserved during evolution by the amino acid residues that are essential for binding DNA. Thus, eliminating C10 or H29 of the Zn-structure 1, which starts with C3 at the N terminus of the domain, as well as a three-residue truncation of the NLS (R75-K77) at the C terminus of the domain abolished DNA binding. The latter alteration also decreased nuclear import severely, indicating that the NLS is physically included in the DNA-binding domain, and that certain residues are needed for both functions.

Despite the fact that all basic amino acid residues participate in the formation of the positively charged surface directed towards the DNA,[10] single amino acid exchanges, with the exception of the Zn-coordinating histidine residues, do not seem to harm DNA binding *in vitro*. In contrast, single mutations in the Zn-coordinating cysteine and histidine residues destroy DNA binding, probably because they lead to larger structural changes. Interestingly, eliminating Zn-structure 2 by a C64 to alanine exchange affects nuclear import positively. This opposite effect on DNA binding and nuclear import indicates that these are mechanistically independent processes, ruling out the possibility that nuclear localization of SBP proteins results simply from passive diffusion into the nucleus and retention by binding to the DNA.

Another single amino acid residue affecting both functions is S63, which is predicted to be a target for phosphorylation. Exchange to aspartate, mimicking phosphorylation, abolished DNA binding of SPL3 and SPL8 completely, and reduced nuclear import in the case of SPL8. In contrast, S63A left DNA binding intact, while nuclear import was enhanced. In the context that both functions overlap, this could

mean that the ultimate function of these transcription factors, to bind to a specific locus in the genome, is prevented by phosphorylation of S63 and, if phosphorylation occurred in the cytosol, nuclear import might be reduced by masking the NLS. Functional overlap of the DNA-binding domain with the NLS suggests also that as long as importin α is bound to the NLS, DNA binding is prohibited.

Materials and Methods

Production of recombinant SBP-domain proteins

For the expression of His$_6$-tagged peptides in *E. coli*, the DNA fragments encoding different SBP-domains were cloned into vector pET11a-his, which is a modified version of pET11a.[27,28] The fragments were generated by PCR using cDNAs as templates and proof-reading Pwo DNA polymerase (peQlab). The following primers with indicated restriction sites (underlined) were used for the amplification of the respective fragments encoding the peptides aligned in Figure 1(b):

SPL1 (accession number AJ011629)
sense, SacII: TTC TCC CCG CGG CCT GCC AAG AAG ACC AAA TCC GGA;
antisense, KpnI: TCC TCC GGT ACC CTA GTT AGC GCC AGG TTC GGG ATT TGT.

SPL3 (accession number Y09427)
sense, BglII: C TCC AGA TCT CTG GTT CCG CGG GGT AAA GCT ACA AGT AGT AGT GGA GT;
antisense, BamHI: TCC TCC GGA TCC TTA GTC AGT TGT GCT TTT CCG CCT.

SPL7 (accession number AJ011612)
sense, BglII: TTT TCC AGA TCT CTG GTT CCG CGG GTT CGC GGC GGG TCG GGC GTG GC;
antisense, BamHI: TAA TCC GGA TCC TCA TTG TTC TGC AGC AAC ACC TCC TTT ATC.

SPL8 (accession number AJ011642)
sense, BamHI: TCC TCC GGA TCC CTG GTT CCG CGG ATG GCG AAC TCG CTG AGC ACG;
antisense, BamHI: TCC TCC GGA TCC CTA GGT GGC GGA GGC TGA TTG GTG ACA.

PpSBP1 (accession number AJ968320)
sense, SacII: TCC TCC CCG CGG GCA GGT TCA CCT GGT TTC CAA GT,
antisense; KpnI: TCC TCC GGT ACC TCA TGA AGT AGA GGC ATT TAG TTG TGG.

CRR1 (accession number AY484394)
sense, SacII: TCC TCC CCG CGG GTG GGT GAC GAC ACC GGC GCC ATG,
antisense; KpnI: TCC TCC GGT ACC TCA CGC AGC GCG GCC CTT GGC CTG CTC.

Fragments were digested with the indicated restriction enzymes and ligated into the vector cut with BamHI or SacII/KpnI and transformed into *E. coli* BL21 Codon Plus (DE3) RIL (Stratagene).

For transient expression in tobacco protoplasts, *SPL* genes were cloned behind the 35 S promoter in-frame in front of the *GFP* gene. DNA of *SPL3*, *SPL4* (accession number AJ011631), *SPL5* (accession number AJ242960), *SPL8*, *SPL9* (accession number AJ011638), and *SPL15* (accession number NM_115654) were amplified by PCR using sense primers, which had a 25 nt homology to the 5′end of the open reading frames and contained the start codon ATG embedded in the cloning site NcoI. The

antisense primers consisted of 25 nt homologous to the 3′ends, omitting the stop codons. An extra C nucleotide had to be inserted in front of the NcoI cloning site to keep the frame, resulting in an extra alanine between the protein of interest and GFP. The fragments were digested with NcoI and ligated into expression vector pGJ619 also cut with NcoI. The constructs were amplified in *E. coli* DH5α.

Mutagenesis

Mutagenesis in all cases was carried out by inverse PCR, using back-to-back primers pointing outwards, one carrying the mutation and the other carrying the wild-type sequence. In this way, DNA fragments of the whole plasmid were produced, linearized at the point where the 5′-ends of the primers meet. Pwo polymerase was used to yield blunt-ended fragments and high-level sequence accuracy. The fragments were 5′phosphorylated with phage T4 polynucleotide kinase and recircularized by ligation before amplification in *E. coli* DH5α.

Expression and purification of SBP-domains

For expression in *E. coli*, the plasmids were transformed into strain BL21 Codon Plus (DE3) RIL (Stratagen), which contains a plasmid carrying extra genes for rare tRNAs in *E. coli*. The cells were grown in 500 ml of LB medium with 100 μg/ml of ampicillin and 87 μg/ml of chloramphenicol to an absorbance at 600 nm of 0.8. After induction of protein expression with IPTG (1 mM final concentration), the cells were grown at 37 °C for another 2 h. The cells were harvested by centrifugation and resuspended in 15 ml of lysis buffer (50 mM Tris (pH 7.5), 300 mM NaCl, 1 mM EDTA, 10 mM β-mercaptoethanol, 10% (v/v) glycerol). After one freeze–thaw cycle (−80 °C) the cells were disrupted by sonication and the lysates cleared by centrifugation at 30,000*g* for 30 min.

A sample (10 ml) of cleared lysates containing the His-tagged proteins were mixed with 1 ml of Ni-NTA agarose (Qiagen) and incubated for 1 h at 6 °C with slow shaking. The slurry was placed into a column and washed twice with 10 ml of lysis buffer containing 10 mM imidazole. The protein was eluted with lysis buffer containing 250 mM imidazole and stored at −20 °C.

Electrophoretic mobility-shift assays (EMSA)

EMSA was used to identify protein–DNA complexes. Two DNA substrates (AP1-15 and Cyc6 CuRE, indicated in Figure 3) were prepared by slowly hybridizing the (5′-^{32}P)-end labeled upper strand to a twofold excess of unlabeled lower strand in TE containing 10 mM MgCl$_2$. Various amounts of protein were incubated with 100 fmol of labeled DNA at room temperature in 10 μl of binding buffer (50 mM Tris–HCl (pH 8.0), 1 mM EDTA, 1 mM DTT). After 15 min, 2.5 μl of loading buffer (40 mM Tris (pH7.5), 1 mM EDTA, 25% glycerol, 0.05% (w/v) bromphenol blue, and 400 μg/ml of bovine serum albumin) was added to each sample. The mixtures were immediately loaded onto a native 10% (w/v) polyacrylamide gel and run at room temperature for 2.5 h at 7 V/cm in 0.5×TBE electrophoresis buffer. Bands on the gel were visualized by phosphorimaging.

Zn^{2+}-depleted protein

To deplete SPL1 SBP from Zn^{2+}, purified protein at a concentration of 1 mg/ml was denatured by adding urea powder to a final concentration of 8 M. A sample (100 μl) of denatured protein was than dialyzed on floating membranes (pore size 0.025 μm; Millipore Cat. no. VSWP02500) twice for 30 min, first against 20 ml of lysis buffer containing 20 mM EDTA, 2 M urea, and then the same buffer without urea. As a positive control, protein was treated in the same way but without EDTA. The proteins were than used in EMSAs.

Random oligonucleotide binding assay

The random primer selection experiment was carried out using 62 nt oligonucleotides containing a random central string of 16 nucleotides (or GT at positions 7 and 8) flanked by defined ends for PCR amplification and cloning: GGTCAGTTCAGCGGATCCTGTCG…16×N…GAGG CGAATTCAGTGCAACTGCG or GGTCAGTTCAGCGGATCCT GTTT…6×N.GT.8×N…AAAGCGAATTCAGTGCAACTGCG. To make double-stranded [^{32}P]DNA-substrate, 20 pmol of oligonucleotide was primed with an antisense oligonucleotide to the 3′ fixed sequence of the 62mer. This primer was elongated by Klenow enzyme using standard conditions with [^{32}P]dCTP included in the reaction. After precipitation in ethanol this substrate was used for the first of three consecutive rounds of DNA binding by SBP-domains, recovering the protein-DNA complexes from EMSA gels by elution and amplification of bound DNA by PCR, using primers to both defined terminal sequences of the 62mer under labeling conditions. After the third round, the amplified DNA was cloned for sequencing into TOPO-TA vector (Invitrogen) according to the supplier's protocol.

Immunolocalization of SPL3

The immunolocalization experiment was carried out as described.[29]

Transient expression of SBP-domain proteins in plant cells

Protoplasts from tobacco BY2 suspension cultures were used for transient expression assays. Experiments were done as described,[30] but downscaled to allow for handling and cultivation of the cells in 2 ml Eppendorf tubes.

Quantification of nuclear import

SBP-domain protein GFP fusion constructs were expressed transiently in tobacco BY-2 protoplasts and imaged in a confocal microscope (Leica TCS SP2). The digital images were recorded such that all pixel values remained within the dynamic range. For relative quantification of nuclear *versus* cytoplasmic fluorescence signal, the average pixel values over the cytoplasm and the nucleus were determined using the image processing program ImageJ†.[31] At least ten different transformed protoplasts were evaluated for each construct.

---

† http://rsb.info.nih.gov/ij/

References

1. Klein, J., Saedler, H. & Huijser, P. (1996). A new family of DNA binding proteins includes putative transcriptional regulators of the *Antirrhinum majus* floral meristem identity gene *SQUAMOSA*. *Mol. Gen. Genet.* **250**, 7–16.
2. Cardon, G., Höhmann, S., Nettesheim, K., Saedler, H. & Huijser, P. (1997). Functional analysis of the *Arabidopsis thaliana* SBP-box gene *SPL3*: a novel gene involved in the floral transition. *Plant J.* **12**, 367–377.
3. Cardon, G., Höhmann, S., Klein, J., Nettesheim, K., Saedler, H. & Huijser, P. (1999). Molecular characterisation of the *Arabidopsis* SBP-box genes. *Gene*, **237**, 91–104.
4. Schmid, M., Uhlenhaut, N. H., Godard, F., Demar, M., Bressan, R., Weigel, D. & Lohmann, J. U. (2003). Dissection of floral induction pathways using global expression analysis. *Development*, **130**, 6001–6012.
5. Unte, U. S., Sorensen, A. M., Pesaresi, P., Gandikota, M., Leister, D., Saedler, H. & Huijser, P. (2003). *SPL8*, an SBP-box gene that affects pollen sac development in *Arabidopsis*. *Plant Cell*, **15**, 1009–1019.
6. Stone, J. M., Liang, X., Nekl, E. R. & Stiers, J. J. (2005). *Arabidopsis* AtSPL14, a plant-specific SBP-domain transcription factor, participates in plant development and sensitivity to fumonisin B1. *Plant J.* **41**, 744–754.
7. Moreno, M. A., Harper, L. C., Krueger, R. W., Dellaporta, S. L. & Freeling, M. (1997). Liguleless1 encodes a nuclear-localized protein required for induction of ligules and auricles during maize leaf organogenesis. *Genes Dev.* **11**, 616–628.
8. Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B. & Bartel, D. P. (2002). Prediction of plant microRNA targets. *Cell*, **110**, 513–520.
9. Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M. & Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Dev. Cell*, **8**, 517–527.
10. Yamasaki, K., Kigawa, T., Inoue, M., Tateno, M., Yamasaki, T., Yabuki, T. *et al*. (2004). A novel zinc-binding motif revealed by solution structures of DNA-binding domains of Arabidopsis SBP-family transcription factors. *J. Mol. Biol.* **337**, 49–63.
11. Görlich, D. & Kutay, U. (1999). Transport between the cell nucleus and the cytoplasm. *Annu. Rev. Cell Dev. Biol.* **15**, 607–660.
12. Kaffman, A. & O'Shea, E. K. (1999). Regulation of nuclear localization: a key to a door. *Annu. Rev. Cell Dev. Biol.* **15**, 291–339.
13. Jans, D. A., Xiao, C. Y. & Lam, M. H. (2000). Nuclear targeting signal recognition: a key control point in nuclear transport? *BioEssays*, **22**, 532–544.
14. Robbins, J., Dilworth, S. M., Laskey, R. A. & Dingwall, C. (1991). Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell*, **64**, 615–623.
15. Fontes, M. R., Teh, T., Jans, D., Brinkworth, R. I. &

Kobe, B. (2003). Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha. *J. Biol. Chem.* **278**, 27981–27987.

16. Lännenpää, M., Jänönen, I., Hölttä-Vuori, M., Gardemeister, M., Porali, I. & Sopanen, T. (2004). A new SBP-box gene BpSPL1 in silver birch (*Betula pendula*). *Physiol. Plant*, **120**, 491–500.

17. Orlando, V. (2000). Mapping chromosomal proteins *in vivo* by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem. Sci.* **25**, 99–104.

18. Quinn, J. M., Eriksson, M., Moseley, J. L. & Merchant, S. (2002). Oxygen deficiency responsive gene expression in *Chlamydomonas reinhardtii* through a copper-sensing signal transduction pathway. *Plant Physiol.* **128**, 463–471.

19. Quinn, J. M., Barraco, P., Eriksson, M. & Merchant, S. (2000). Coordinate copper- and oxygen-responsive Cyc6 and Cpx1 expression in *Chlamydomonas* is mediated by the same element. *J. Biol. Chem.* **275**, 6080–6089.

20. Berg, J. M. & Shi, Y. (1996). The galvanization of biology: a growing appreciation for the roles of zinc. *Science*, **271**, 1081–1085.

21. Harreman, M. T., Kline, T. M., Milford, H. G., Harben, M. B., Hodel, A. E. & Corbett, A. H. (2004). Regulation of nuclear import by phosphorylation adjacent to nuclear localization signals. *J. Biol. Chem.* **279**, 20613–20621.

22. Blom, N., Gammeltoft, S. & Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* **294**, 1351–1362.

23. Conti, E., Uy, M., Leighton, L., Blobel, G. & Kuriyan, J. (1998). Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell*, **94**, 193–204.

24. Pineda-Molina, E., Klatt, P., Vazquez, J., Marina, A., Garcia de Lacoba, M., Perez-Sala, D. & Lamas, S. (2001). Glutathionylation of the p50 subunit of NF-kappaB: a mechanism for redox-induced inhibition of DNA binding. *Biochemistry,* **40**, 14134–14142.

25. Klatt, P., Molina, E. P. & Lamas, S. (1999). Nitric oxide inhibits c-Jun DNA binding by specifically targeted S-glutathion. *J. Biol. Chem.* **274**, 15857–15864.

26. Matthews, J. R., Wakasugi, N., Virelizier, J. L., Yodoi, J. & Hay, R. T. (1992). Thioredoxin regulates the DNA binding activity of NF-kappa B by reduction of a disulphide bond involving cysteine 62. *Nucl. Acids Res.* **20**, 3821–3830.

27. Birkenbihl, R. P., Neef, K., Prangishvili, D. & Kemper, B. (2001). Holliday junction resolving enzymes of archaeal viruses SIRV1 and SIRV2. *J. Mol. Biol.* **309**, 1067–1076.

28. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.

29. Zachgo, S., Silva-Ede, A., Motte, P., Tröbner, W., Saedler, H. & Schwarz-Sommer, Z. (1995). Functional analysis of the *antirrhinum* floral homeotic *DEFICIENS* gene *in vivo* and *in vitro* by using a temperature-sensitive mutant. *Development*, **121**, 2861–2875.

30. Jach, G., Binot, E., Frings, S., Luxa, K. & Schell, J. (2001). Use of red fluorescent protein from *Discosoma* sp. (dsRED) as a reporter for plant gene expression. *Plant J.* **28**, 483–491.

31. Abramoff, M. D., Magelhaes, P. J. & Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics Int.* **11**, 36–42.

32. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190.

Edited by M. F. Summers