Biochimica et Biophysica Acta 1799 (2010) 454–462

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

DNA binding by the Arabidopsis CBF1 transcription factor requires the PKKP/RAGRxKFxETRHP signature sequence

Donatella Canella [a,b,1], Sarah J. Gilmour [a], Leslie A. Kuhn [b,c,d], Michael F. Thomashow [a,*]

[a] MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824, USA
[b] Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA
[c] Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824, USA
[d] Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:
Received 16 October 2009
Accepted 23 November 2009
Available online 27 November 2009

Keywords:
AP2/ERF domain
Arabidopsis
CBF/DREB1 transcription factor
DNA binding
Signature sequence

ABSTRACT

The CBF/DREB1 transcriptional activators are key regulators of plant freezing tolerance. They are members of the AP2/ERF multi-gene family, which in *Arabidopsis* comprises about 145 members. Common to these proteins is the AP2/ERF DNA-binding domain, a 60-amino-acid fold composed of a three-stranded β-sheet followed by a C-terminal α-helix. A feature that distinguishes the CBF proteins from the other AP2/ERF proteins is the presence of "signature sequences," PKKP/RAGRxKFxETRHP (abbreviated PKKPAGR) and DSAWR, which are located immediately upstream and downstream, respectively, of the AP2/ERF DNA-binding domain. The signature sequences are highly conserved in CBF proteins from diverse plant species suggesting that they have an important functional role. Here we show that the PKKPAGR sequence of AtCBF1 is essential for its transcriptional activity. Deletion of the sequence or mutations within it greatly impaired the ability of CBF1 to induce expression of its target genes. This impairment was not due to the mutations eliminating CBF1 localization to the nucleus or preventing protein accumulation. Rather, we show that this loss of function was due to the mutations greatly impairing the ability of the CBF1 protein to bind to its DNA recognition sequence, the CRT/DRE element. These results establish that the ability of the CBF proteins to bind to the CRT/DRE element requires amino acids that extend beyond the AP2/ERF DNA-binding domain and raise the possibility that the PKKPAGR sequence contributes to determining the DNA-binding specificity of the CBF proteins.

1. Introduction

The *Arabidopsis* CBF cold response pathway has a prominent role in cold acclimation, the process whereby certain plants increase in freezing tolerance in response to low non-freezing temperatures [1,2]. The pathway includes rapid cold induction of three genes encoding transcription factors, *CBF1*, *-2* and *-3* [3,4]—also known as *DREB1b*, *-c*, *-a*, respectively [5]—that bind to the CRT/DRE DNA regulatory element present in the promoters of CBF-target genes [6,7]. Induction of the CBF regulon of genes, which comprises about 100 members [8,9], leads to an increase in freezing and drought tolerance [5,10].

The CBF proteins are members of the AP2/ERF family of transcription factors [11]. The AP2/ERF protein family, comprising 145 members in *Arabidopsis*, is defined by the conserved 60-amino-acid AP2/ERF DNA-binding domain [12,13]. The AP2/ERF domain is composed of a three-strand β-sheet structure followed by an α-helix (14). The 3D solution structure of the AtERF1 AP2/ERF domain has shown that arginine and tryptophan residues within the β-sheet contact nucleotides of the binding site within the major groove of the DNA [14]. Those key residues are well conserved among members of the AP2/EREBP family, and yet different AP2 proteins display different DNA-binding preferences. Therefore, binding specificity within the family is imparted by additional residues within or outside the canonical DNA-binding domain. It has been proposed that the specificity determinants in these two subfamilies lie within the AP2/EREBP DNA-binding domain [15,16]. Sakuma et al. [16] have shown that specific amino acids at two conserved positions within the AP2/ERF domains of DREB and ERF proteins affect DNA-binding affinity. At the present, however, it is still unclear whether additional residues within the two families can impart specific DNA preferences.

The primary feature that distinguishes the CBF transcription factors from the other 145 AP2/ERF family members in *Arabidopsis* is the "signature sequences" that flank the AP2/ERF domain [17]. These sequences, PKK/RPAGRxKFxETRHP, which we will refer to as PKKPAGR, and DSAWR, are located immediately up- and downstream, respectively, from the AP2/ERF domain in the CBF proteins. The signature sequences are highly conserved in CBF proteins from

* Corresponding author. Tel.: +1 517 355 2299; fax: +1 517 353 9168.
E-mail address: thomash6@msu.edu (M.F. Thomashow).
[1] Current address: Université de Lausanne, Centre Intégratif de Génomique (CIG), CH-1015 Lausanne-Dorigny, Lausanne, Switzerland.

1874-9399/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbagrm.2009.11.017

diverse plant species suggesting that they have an important function. Here we test this hypothesis for the PKKPAGR sequence.

2. Materials and methods

2.1. Plant material and growth conditions

Arabidopsis thaliana plants ecotype Colombia-0 (Col-0) or Wassilewskija-2 (Ws-2) were grown either in soil as described previously [18] or on solidified Gamborg's B5 medium (Caisson Laboratories, North Logan, UT, USA) for 2 weeks at 22°C under 100 μmol m^{-2} s^{-1} constant light, as previously described [9, 18]. Transgenic plants overexpressing CBF1 (lines G6 and G26) or carrying vector pGA643 (line B6) in Ws-2 background have been described previously [18]. *Arabidopsis* transformation was performed using the floral dip method [19].

2.2. RNA isolation and analysis

Plant material was harvested in liquid nitrogen and total RNA extracted using RNeasy Plant Mini kits (Qiagen, Valencia, CA, USA) with modifications as described [20]. Total RNA (5–10 μg) was fractionated in 1% formaldehyde gels and transferred onto nylon membranes as described [21]. Membranes were hybridized in Church buffer (1% BSA, 1 mM EDTA, 0.5 M $NaPO_4$, pH 7.2, 7% SDS) [22] at 65 °C overnight. Blots were hybridized with ^{32}P-labeled fragments prepared using the Random Primers DNA Labeling System (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions and washed under high stringency [22]. Following the washes, the membranes were exposed to a phosphorimager screen (BioRad, Hercules, CA, USA), which was scanned and then quantified using QuantityOne software (BioRad, Hercules, CA, USA). RNA levels were normalized by comparison to 18S rRNA determined from the same blots.

2.3. Protein isolation from plants

Total protein extracts were prepared from 2-week-old *Arabidopsis* seedlings by grinding frozen tissue in protein extraction buffer (20 mM Tris–HCl, pH 8.0, 50 mM NaCl, 5 mM EDTA, 0.05% SDS) supplemented with protease inhibitor cocktail (Roche Applied Science, Indianapolis, IN, USA) followed by centrifugation for 15 min at 4 °C at maximum speed in a microcentrifuge. Protein concentration was determined using Bradford reagent (BioRad, Hercules, CA, USA) with BSA as the standard.

2.4. Mutagenesis of the PKKPAGR signature sequence

Site-directed mutagenesis [23] was performed to convert tripeptides within the wild-type PKKPAGR signature sequence of CBF1 to stretches of three alanines. Primers were designed (M1–M5 F and R, see Supplementary Table 1) to introduce the desired mutations using the QuikChange mutagenesis kit (Stratagene, Cedar Creek, TX, USA) according to the manufacturer's instructions. A NotI restriction site was included in these primers for the screening of plasmids containing the mutated sequences. The template DNA used was full-length *CBF1* cDNA in pBS/SK$^-$. The ΔPKK mutant, which lacks the entire PKKPAGR region, was made using a modified protocol based on the QuikChange method [24]. To overcome the tendency of the perfectly complementary mutagenic primers to anneal to each other rather than to the target sequence, a two-stage PCR was performed, running two separate single-primer reactions before the final PCR amplification. Primers ΔPKKm F and R were used for the PKKPAGR deletion mutant (Supplementary Table 1). Mutated versions of *CBF1* were amplified by PCR using primers mPKK F and R, which introduced a BglII site (see Supplementary Table 2). These fragments were cloned into the BglII site of the binary vector, pGA643 [25], downstream of the CaMV 35S promoter.

Point mutations in the RKKFRET motif were designed using environment-specific substitution tables [26], which allow one to choose substitutions compatible with its predicted helical structure. Point mutations were introduced into CBF1 using the QuikChange kit (Stratagene, Cedar Creek, TX, USA) using the primers shown in Supplementary Table 3 and following the manufacturer's protocol. For screening the clones harboring the desired mutation, the primers were designed to disrupt a pre-existing Sau96I restriction site in all mutant sequences, except for the Phe40→Ala and Phe40→Pro mutations, where a new XhoI restriction site was inserted. The mutant versions of *CBF1* were amplified by PCR using the primers mPKK F and R shown in Supplemental Table 2 and cloned into pGA643 as described above.

2.5. Construction of protein fusions

Wild-type and mutated versions of CBF1 were tagged with 6xMyc by subcloning *CBF1* sequences into the binary vector pKVB24 [27], which contains a 6xMyc tag under the control of the CaMV 35*S* promoter. The translational fusions resulted in 6xMyc at the N-terminus of the CBF1 proteins. Primers for PCR amplification of the *CBF1* sequence were mycCBF1 F and R (see Supplementary Table 2), which added SmaI/SacI ends.

Wild-type and mutated versions of CBF1 were tagged with GFP:GUS by cloning into the binary vector pEZT-CL(GUS). This vector was engineered by inserting a GUS fragment with BamHI ends into the pEZT-CL plant expression vector [28]. The primers used to amplify the GUS sequence were GUS F and R (see Supplementary Table 2). This resulted in an in-frame fusion of GUS to GFP under the control of the CaMV 35S promoter. The eGFP gene in pEZT is based on mGFP4 [29] and contains additional mutations (S65T, Y66H) to increase intrinsic GFP fluorescence [30]. Full-length and 5′ deletions of *CBF1* were amplified by PCR using the primers shown in Supplementary Table 2, which added XhoI ends. These *CBF1* deletions were subcloned into the XhoI site of pEZT-CL(GUS), which generated in-frame fusions to GFP:GUS. The XhoI insert for NIa was PCR-amplified from the yeast plasmid pAVA367 [28], containing the in-frame fusion NIa-GFP, using the primers NIa F and R shown in Supplementary Table 2.

Wild-type and mutated versions of a 258-bp fragment of *CBF1* encoding amino acids 27–112 were fused to the Maltose Binding Protein (MBP) by cloning into the pMAL-c2X expression vector (New England BioLabs, Beverly, MA, USA) downstream of the *malE* gene. This resulted in a translational fusion of $CBF1_{27\text{-}112}$ to the C-terminus of MBP. The primers used, MBPCBF1 F and R, which include XbaI and XmnI sites, are shown in Supplementary Table 2.

2.6. Expression and purification of MBP:CBF1 proteins

Constructs with an MBP tag in the expression vector pMAL-c2X (New England BioLabs, Beverly, MA, USA) were transformed into *Escherichia coli* strain BL21-CodonPlus(DE3)-RIL (Stratagene, La Jolla, CA, USA) and protein expression induced by addition of 1 mM IPTG to the bacterial suspension. Cells were lysed by sonication, and the soluble protein fraction was separated by centrifugation. The supernatant containing the fusion proteins was loaded onto an amylase column pre-equilibrated in column buffer (20 mM Tris–HCl, 200 mM NaCl, 1 mM EDTA). MBP-CBF1 proteins were eluted in column buffer containing 10 mM maltose.

2.7. Expression and purification of 6xHis:CBF1 proteins

Constructs with a 6xHis tag were introduced into the pET28a$^+$ expression vector (Novagen/EMD, San Diego, CA, USA) and the plasmids were transformed into *E. coli* strain BL21-CodonPlus(DE3)-RIL (Stratagene, La Jolla, CA, USA) for optimal expression of the recombinant proteins [31]. Protein expression was induced with 1

mM IPTG, and the cells were harvested by centrifugation and resuspended in 10 ml of cell lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0). Cells were lysed by sonication in the presence of protease inhibitor cocktail (Roche Applied Science, Indianapolis, IN, USA). The soluble protein fraction was separated by centrifugation and the supernatant containing 6xHis-T7-CBF1 proteins was loaded over a nickel column equilibrated with wash buffer (as lysis buffer but with 20 mM imidazole). The proteins were eluted by increasing the imidazole to 250 mM.

2.8. Western analysis

Proteins were separated by SDS–polyacrylamide gel electrophoresis (PAGE) on 4–20% gradient gels (ISC BioExpress, Kaysville, UT, USA). After electrophoresis, the proteins were transferred to polyvinylidene difluoride (PVDF) membranes by electrotransfer. Membranes were blocked in 5% non-fat milk powder in Tris-buffered saline (TBS)–0.1% Tween-20 for 1 h at room temperature, and then incubated overnight at 4 °C with a suitable dilution of the primary antibody. The primary antibodies used were α-MBP (New England Biolabs, Beverly, MA, USA), α-Myc monoclonal antibody (Roche Applied Science, Indianapolis, IN, USA), and α-GFP monoclonal antibody (Pierce Biotechnology, Rockford, IL, USA). The secondary antibody was horseradish peroxidase-coupled anti-mouse IgG antibody (Pierce Biotechnology, Rockford, IL). Immunoreactive bands were visualized using the Amersham enhanced chemiluminesence assay (GE Healthcare, Piscataway, NJ, USA) following the manufacturer's instructions.

2.9. Electrophoretic mobility shift assays

Fragments 23 bp in length containing the CRT/DRE element from the *COR15a* and *COR78* promoters were prepared by synthesizing and annealing both strands using the oligomers shown in Supplementary Table 4. A mutant version of the *COR15a* CRT/DRE element in which the core CCGAC had been altered was used for the competition assays. The resulting double stranded oligonucleotides were end-labeled

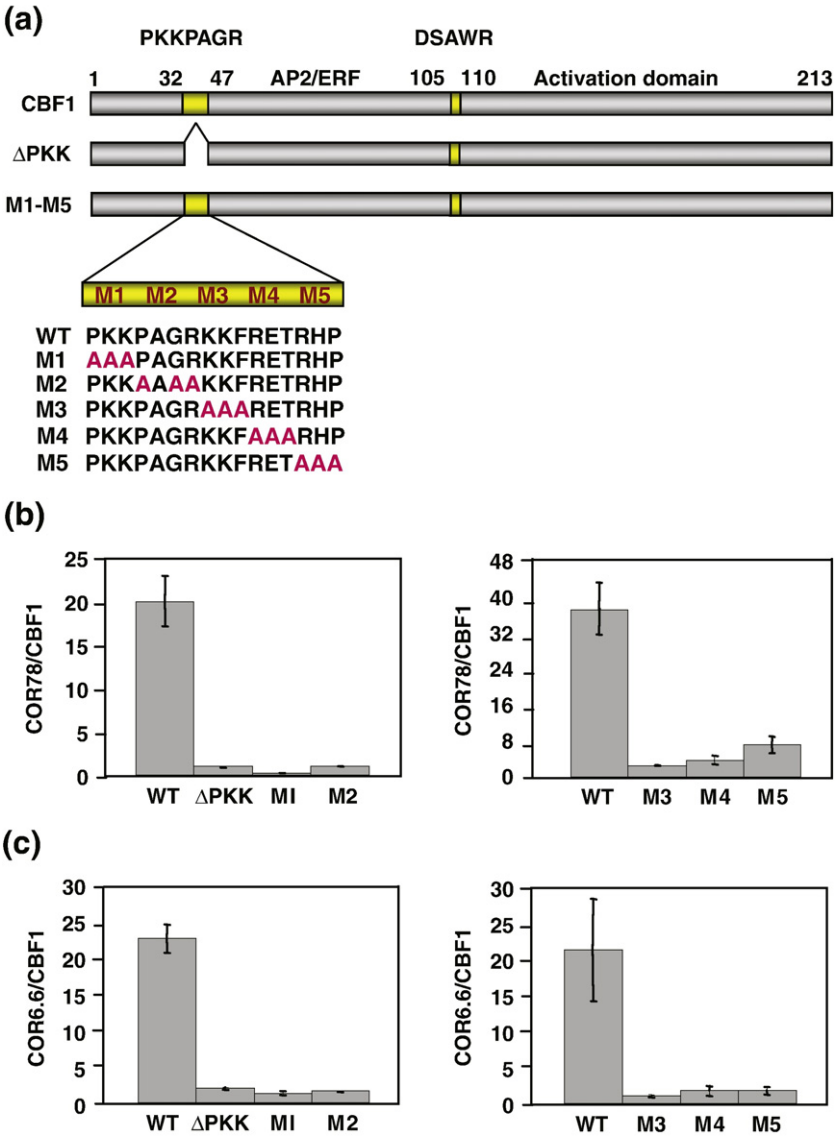

Fig. 1. Mutations within the PKKPAGR signature sequence impair CBF1 function. (a) Schematic of full-length CBF1, CBF1 lacking the PKKPAGR motif (ΔPKK), and mutations within the PKKPAGR motif (M1–M5). Amino acids shown in red indicate where the wild-type sequence has been substituted with alanines. (b) *COR78*/*CBF1* transcript ratios in transgenic plants overexpressing wild-type or mutated *CBF1* transgenes. Ratios were determined by quantification of northern blots as described in Materials and methods. Five to six independent transgenic lines were used for each mutation. The error bars indicate standard error. WT was significantly different from ΔPKK and M1–M5 (P value<0.0001 in all cases). (c) *COR6.6*/*CBF1* transcript ratios in transgenic plants overexpressing wild-type or mutated *CBF1* transgenes. Ratios were determined as in (b). WT was significantly different from ΔPKK and M1–M5 (P value<0.0001 in all cases).

with ^{32}P and purified through a Sephadex G-50 column. The binding of recombinant protein to the CRT/DRE-containing DNA probes was tested by incubating 0.5 ng of ^{32}P-probe labeled by end-filling with a gradient concentration of each recombinant protein in the presence/absence of 100 ng unlabeled competitor DNA for 20 min at room temperature. After incubation, samples were separated on non-denaturing polyacrylamide gels. The gels were dried and exposed to a phosphorimager screen. The screen was scanned and quantified using QuantityOne software (BioRad, Hercules, CA, USA).

2.10. Fluorescence imaging of Arabidopsis root tips overexpressing CBF1:GFP:GUS

Arabidopsis seedlings were grown on vertically oriented plates and allowed to grow for 6–7 days before being imaged. Laser confocal images were collected using an upright LSM Zeiss 510 META microscope (Carl Zeiss, Thornwood, NY, USA) equipped with a 40× oil immersion objective. To visualize the nuclei, *Arabidopsis* seedlings were incubated with a solution of 50 μg/ml propidium iodide (PI) and 5 μg/ml RNase in the dark for 15–30 min. Samples were rinsed with distilled water and mounted in tap water. GFP fluorescence images were obtained using Argon ion laser excitation of 488 nm with a 505/530 nm bandpass filter. PI fluorescence images were collected using an excitation line of 543 nm with a 560-nm longpass filter. Postacquisition image processing was done with the LSM 5 Image Browser and Adobe Photoshop 5.0 software (Adobe Systems, San Jose, CA, USA).

2.11. ANOVA

Statistical difference in COR/CBF1 ratios of transgenic plants overexpressing a wild-type *CBF1* transgene or *CBF1* transgenes mutated in the signature sequences was assessed by analysis of variance (ANOVA) using SAS Proc Mixed procedures (SAS Institute, Cary, NC, USA), version 9.1. To test for significant difference between COR/CBF1 ratios between transgenic plants overexpressing wild-type *CBF1* or the mutated versions, we estimated least-square means for each genotype and compared them to the least-square means of the control plants overexpressing wild-type *CBF1*. These estimates were used to calculate the *t*-values and the statistical significance at $P<0.0001$ for all the transgenic lines.

3. Results

3.1. The PKKPAGR signature sequence is required for CBF1 to induce expression of COR genes

The importance of the PKKPAGR motif in CBF1 function was assessed by determining whether changes in the sequence affected the ability of the transcription factor to induce *COR* gene expression. Two types of modification were made; a complete deletion of the PKKPAGR motif, ΔPKK, and short alanine substitutions, M1–M5, within the motif (Fig. 1a). The mutant versions of CBF1 were placed under control of the strong constitutive CaMV 35*S* promoter, transformed into *Arabidopsis*, and the transcript levels for two CBF1 target genes, *COR78* and *COR6.6*, were determined (Fig. 1b and c; Supplementary Figs. 1 and 2). The results indicated that each of the mutations greatly impaired CBF function; i.e., the ratios of the *COR* to *CBF1* transcripts were much higher in the lines transformed with the wild-type (WT) CBF1 construct than they were in the transgenic lines expressing the CBF1 ΔPKK and M1–M5 mutations. In all cases, the difference in COR78/CBF1 and COR6.6/CBF1 transcript ratios between plants overexpressing WT CBF1 and any of the mCBF1 was highly significant ($P<0.0001$) according to the ANOVA analysis.

3.2. The PKKPAGR sequence is not required for protein stability

The reduced level *COR* gene expression observed in the PKKPAGR mutant lines could have been due to the mutations resulting in protein instability; i.e., the mutant versions of the CBF1 protein could have been much less stable than the WT protein and thus accumulate to much lower levels than that of the WT protein given a similar transcript level. We were unable to test this possibility in the existing transgenic lines as we were unsuccessful in developing an antibody against full-length CBF1, or specific peptides of it, that could clearly detect the CBF1 protein in cold-acclimated WT *Arabidopsis* plants or in transgenic plants overexpressing CBF1 under control of the CaMV 35*S* promoter. Therefore, we tagged the WT CBF1 protein and the M1–M5 mutant CBF1 proteins with c-Myc (see Materials and methods), transformed the constructs into *Arabidopsis*, and tested several transgenic lines for *CBF1* and *COR* gene transcript levels (Fig. 2a) and the accumulation of WT and mutant CBF1 proteins (Fig. 2b).

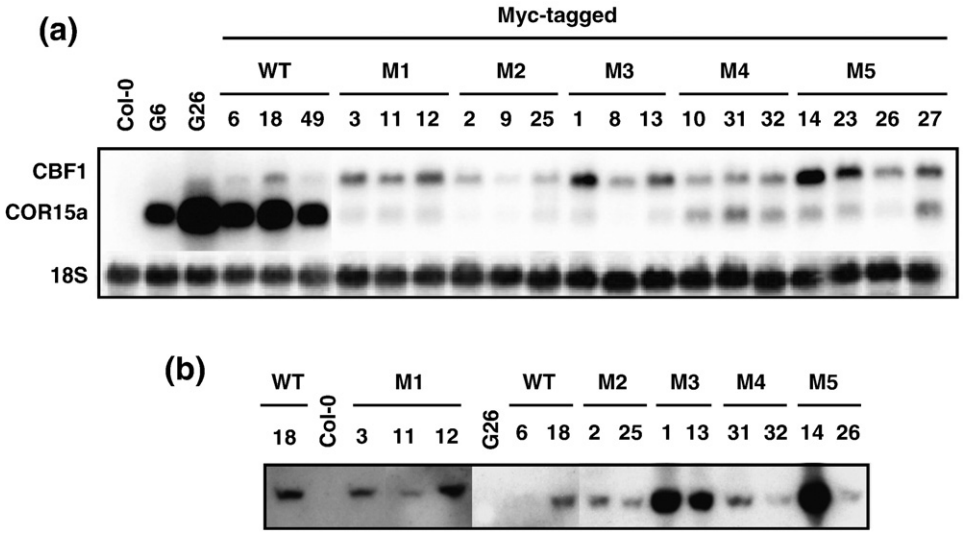

Fig. 2. Reduced function of CBF1 carrying mutations within the PKKPAGR signature sequence is not due to reduced protein levels. (a) Northern blot of transgenic *Arabidopsis* lines overexpressing 6xMyc-tagged CBF1 (WT) or 6xMyc-tagged CBF1 with mutations in the PKKPAGR motif (M1–M5 as in Fig. 1). The numbers indicate the different transgenic lines. G6 and G26 are two independent lines overexpressing wild-type CBF1 without a Myc-tag. Col-0 is a non-transgenic control. 18S ribosomal RNA was used as a loading control. (b) Western blot of transgenic lines overexpressing 6xMyc-tagged CBF1 without (WT) or with (M1–M5) mutations in the PKKPAGR region as detailed in (a). 6xMyc-tagged CBF1 was detected using a monoclonal anti-Myc antibody as described in Materials and methods.

Transgenic plants overexpressing the c-Myc:WT CBF1 protein (Fig. 2a, WT transgenic lines 6, 18, 49), like those expressing the non-tagged WT CBF1 protein (Fig. 2a, lines G6 and G26), were found to accumulate high levels of *COR15a* transcripts. In contrast, transgenic plants overexpressing the c-Myc:M1–M5 mutant CBF1 proteins accumulated low levels of *COR15a* transcripts (Fig. 2a, M1–M5 transgenic lines). This was despite the fact that the *CBF1* transcript levels were generally greater in the plants expressing the c-Myc:M1–M5 CBF1 transgenes than they were in the plants expressing the c-Myc:WT CBF1 transgene. Thus, as observed in the experiments using the non-tagged CBF1 proteins (Fig. 1b and c), CBF1 function was severely impaired by mutations within the PKKPAGR. Further, the results indicated that this decrease in activity was not due to protein instability as the c-Myc:M1–M5 CBF1 proteins accumulated to levels comparable to those in the transgenic lines expressing the c-Myc:WT CBF1 protein (Fig. 2b). For instance, the two lines (2 and 25) expressing the c-Myc:M2 CBF1 transgene had *CBF1* transcript (Fig. 2a) and protein (Fig. 2b) levels that were very similar to those in c-Myc:WT CBF1 line 18. Moreover, the c-Myc:M3 CBF1 lines 1 and 13 as well as the c-Myc:M5 CBF1 line 14 accumulated considerably greater levels

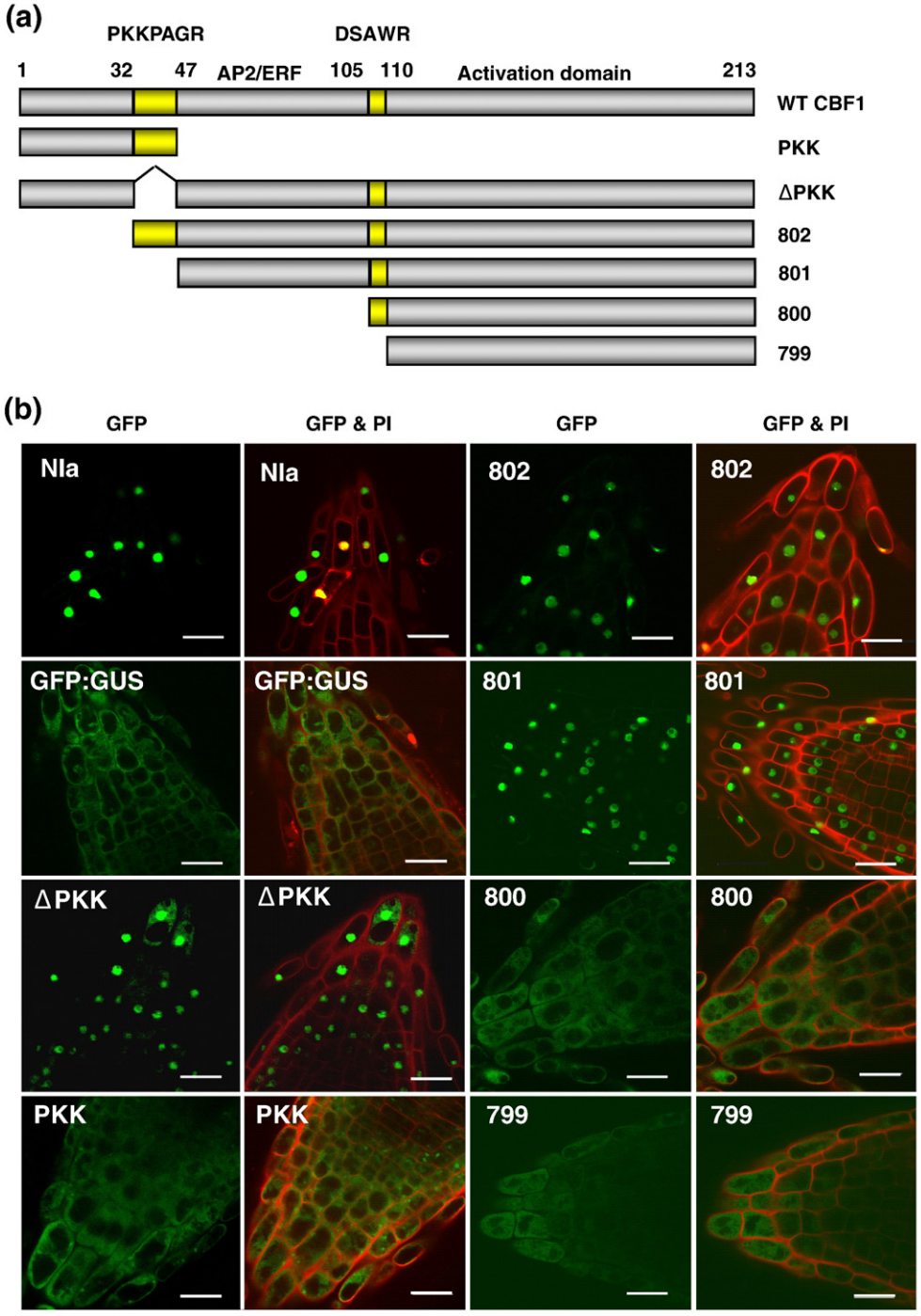

Fig. 3. The PKKPAGR motif is not required for targeting of CBF1 to the nucleus. (a) Schematic of CBF1:GFP:GUS constructs. AP2/ERF is the AP2/ERF DNA-binding domain. (b) Translational fusions to GFP:GUS of each of the CBF1 constructs shown in (a) were generated and transformed into *Arabidopsis* plants. Root tips were examined by confocal microscopy as described in Materials and methods. The panels show images obtained from plants expressing the GFP:GUS constructs indicated. NIa is nuclear inclusion protein a fused to GFP:GUS, and GFP:GUS is GFP:GUS alone. GFP fluorescence images were obtained (GFP), and propidium iodide was used to visualize the nuclei; these images were overlaid (GFP and PI).

of CBF1 protein than did the WT c-Myc:WT CBF1 lines 6 and 18 (Fig. 2b), yet the *COR15a* transcript levels in the M3 and M5 lines were much lower than those in the WT lines (Fig. 2a).

3.3. The PKKPAGR motif is not required for targeting CBF1 to the nucleus

When the *Arabidopsis* CBF1 protein was first described [32], it was suggested that the PKKPAGR sequence might be a nuclear localization signal (NLS). This was based on sequence similarities between the PKKPAGR motif and known NLS sequences for other plant proteins [33–35]. If the PKKPAGR sequence indeed has this function, then it is possible that the inability of mutant PKKPAGR CBF1 proteins to induce *COR* gene expression might be due to the proteins not being imported into the nucleus. To test this, *Arabidopsis* was transformed with constructs that had different parts of the CBF1 protein fused to the GFP:GUS reporter protein (Fig. 3a), and protein localization was determined using laser confocal microscopy (Fig. 3b). Western analysis was also performed to verify that fluorescence detected in plants corresponded to the expression of full-length proteins (Supplementary Fig. 3). The results indicated that the GFP:GUS reporter protein alone did not accumulate in the nucleus (Fig. 3b, GFPGUS). This was expected as the GFP:GUS fusion protein has a mass of approximately 100 kDa, which is well above the limit for free diffusion into the nucleus [36,37]. However, fusion of the NIa nuclear localization signal (NIa) [38] to the GFP:GUS reporter (Fig. 3b; NIa) or fusion of the CBF1 protein without the PKKPAGR region to the GFP:GUS reporter (Fig. 3b; ΔPKK) resulted in nuclear localization of the fusion proteins. Thus, the PKKPAGR motif was not required for nuclear localization of CBF1, indicating that the greatly reduced ability of the ΔPKK and M1–M5 proteins to activate *COR* gene expression was not due to inactivation of an NLS required for the proteins to be imported into the nucleus.

Additional protein fusions were examined in an attempt to localize the CBF1 NLS sequence(s). Fusion of the 47 N-terminal residues of CBF1 to GFP:GUS did not result in nuclear localization of the protein fusion (Fig. 3b, PKK), indicating that the PKKPAGR motif was not sufficient to promote nuclear localization. Deletion of CBF1 amino acids 1–32 (construct 802) or 1–47 (construct 801) did not significantly affect nuclear localization of the reporter fusion, but further deletions removing either the AP2/ERF domain (construct 800) or the AP2/ERF domain and the DSAWR sequence (construct 799) did (Fig. 3b). These results were consistent with the PKKPAGR motif not being essential for localizing CBF1 to the nucleus and indicated that an NLS was located within the AP2/ERF DNA-binding domain.

3.4. Mutations within the PKKPAGR signature sequence greatly impair CBF1 binding to the CRT/DRE DNA regulatory element

We hypothesized that the PKKPAGR motif might play a role in DNA binding based on the proximity of this motif to the AP2/ERF domain and the presence of positively charged residues that could potentially provide favorable electrostatic interactions with the negatively charged DNA backbone. This hypothesis was tested using EMSA to assess the DNA-binding activity of the WT and PKKPAGR mutant versions of CBF1 described in Fig. 1a. The ΔPKK and M1–M5 proteins were fused to a histidine tag, expressed in *E. coli*, purified and assayed for binding to the CRT/DRE DNA regulatory element (see Materials and methods). A band shift was observed with the WT CBF1 protein (Fig. 4). The binding was abolished by addition of unlabeled wild-type competitor CRT/DRE DNA (Fig. 4, W), but not by addition of a version of the sequence in which the core CRT/DRE binding sequence, CCGAC,

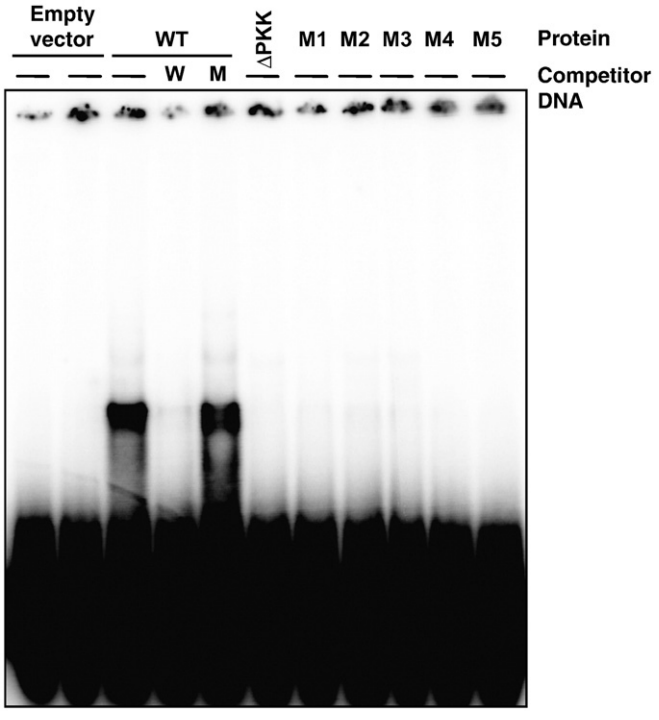

Fig. 4. Mutations in the PKKPAGR motif impair DNA binding. Electrophoretic mobility shift assays using the CRT/DRE element from the *COR15a* promoter as a probe. ΔPKK and M1–M5 mutations of CBF1 are as shown in Fig. 1. Recombinant protein (200 ng) was used in a 12-μl binding reaction with 0.5 ng of radiolabeled probe as described in Materials and methods. Where indicated 100 ng of unlabeled competitor DNA (the *COR15a* CRT/DRE element) was added. W, wild-type competitor DNA; M, mutated competitor DNA in which the core DNA-binding sequence, CCGAC, was altered as described [32].

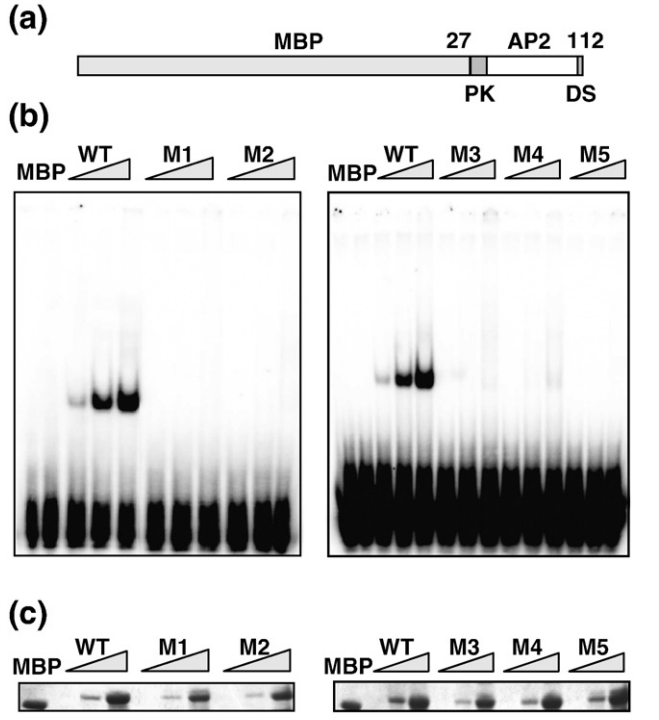

Fig. 5. The PKKPAGR motif is required for CBF1 to bind to the CRT/DRE element. (a) Schematic of maltose binding protein-CBF1 construct. Amino acids 27 to 112, containing the PKKPAGR motif (PK), the AP2/ERF DNA-binding domain (AP2), and the DSAWR motif (DS) were translationally fused to maltose binding protein (MPB). (b) Electrophoretic mobility shift assays using the CRT/DRE element from the *COR15a* promoter as a probe. Increasing amounts (0.2, 2.0, 15 μg) of each recombinant protein were used in a 25-μl binding reaction in the presence of 0.5 ng of radiolabeled probe as described in Materials and methods. WT, wild-type fusion protein as indicated in (a); M1–M5, fusion proteins with mutations in the PKKPAGR region as detailed in Fig. 1; MBP, MBP alone. (c) Western blot of the same proteins in the same amounts as in (b) probed with anti-MBP antibody.

was mutated (Fig. 4, M). These results indicated that the WT CBF1 protein bound specifically to the CRT/DRE sequence. In sharp contrast, little if any binding could be detected with the ΔPKK and M1–M5 mutant CBF1 variants, indicating that PKKPAGR sequence had a major role in DNA binding.

The requirement for the PKKPAGR motif in binding to the CRT/DRE motif was tested further using a construct that contained the CBF1 AP2/ERF domain with surrounding signature sequences fused to the maltose binding protein (Fig. 5a). As with the full-length CBF1 protein, the CBF1 AP2/ERF domain with the surrounding signature sequences bound to the CRT/DRE element (Fig. 5b). This binding was nearly abolished by each of the M1–M5 mutations, again indicating an important role for the PKKPAGR signature sequence in binding to the CRT/DRE element.

3.5. Specific amino acid side chain chemistry within the PKKPAGR signature sequence is required for effective binding of CBF1 to the CRT/DRE regulatory element

Additional analysis was conducted to determine whether specific amino acid side chain chemistry within the PKKPAGR motif was important for DNA binding. We focused on the RKKFRET region within the PKKPAGR motif as it was predicted to form an α-helix (D. Canella and L. Kuhn, unpublished). Both conservative and non-conservative amino acid changes were made to preserve or alter, respectively, side chain chemistry within the RKKFRET sequence (Fig. 6a). Likewise, in all but one case, the changes made were chosen such that they would be compatible with maintaining the predicted α-helical structure of the RKKFRET region [26]. The one exception was the Phe40→Pro mutation, which was designed to test the effect of a helix-breaking residue on the stability of the protein–DNA complex while maintaining a hydrophobic side chain. The effects of each mutation were assessed by EMSA.

The results indicated that the Arg37 and Phe40 residues were critical for DNA binding; the Arg37→Lys, Arg37→Ser and Phe40→Ala substitutions all resulted in greatly impaired binding to the CRT/DRE regulatory element (Fig. 6b). The loss of binding caused by substitution of Arg37 with Lys, two amino acids with very similar side-chain length and charge, indicated that preserving the positive charge was insufficient for binding and that the specific side-chain chemistry was critical. The loss of DNA binding caused by the Phe40→Ala substitution suggested that the aromatic ring in this residue was essential for interaction with the CRT/DRE element. Consistent with this suggestion was the finding that the Phe40→Tyr substitution preserved, or even enhanced, DNA binding. Finally, the finding that the Phe40→Pro substitution resulted in a near complete loss of DNA binding. The Phe40→Pro substitution resulted in near-complete loss of DNA binding. Given that Pro disfavors helicity, this result is consistent with the importance of a helical backbone conformation at this position for DNA binding.

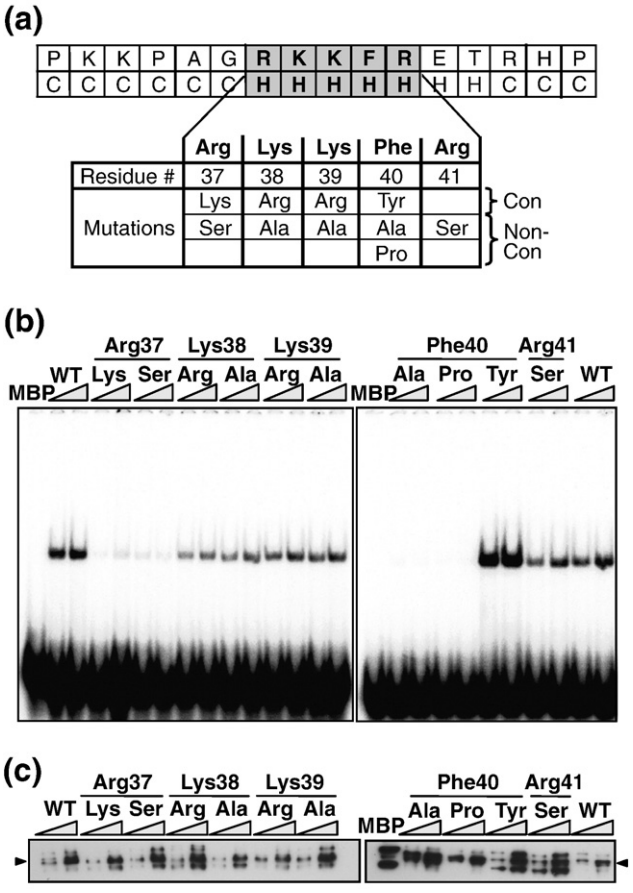

Fig. 6. Specific residues within the RKKFRET region of the PKKPAGR are crucial for CBF1 binding to the CRT/DRE promoter element. (a) The PKKPAGR sequence from CBF1 with the shaded boxes indicating amino acids that were mutated. Predicted α-helix (H) and random coil (C) are indicated. Conservative (Con) and non-conservative (Non-Con) amino acid changes are shown. (b) Electrophoretic mobility shift assays showing binding activity of maltose binding protein (MBP):CBF_{27-112} (see Fig. 5a) with mutations detailed in (a) using the CRT/DRE element from the *COR78* promoter as a probe. DNA-binding reactions were carried out in a 15-μl reaction containing increasing amounts (300 and 600 ng) of each recombinant protein in the presence of 0.5 ng of radiolabeled probe as described in Materials and methods. Non-mutated MBP:CBF_{27-112} (WT) and maltose binding protein alone (MBP) were used as controls. (c) Western blot of the same proteins in the same amounts as in (b) probed with anti-MBP antibody. The arrows indicate MBP:CBF_{27-112}.

4. Discussion

The CBF transcription factors of *Arabidopsis* have a key role in cold acclimation, controlling the expression of a regulon of more than 100 genes that contribute to freezing tolerance. Determining how the CBF proteins regulate the expression of the CBF regulon is basic to an overall understanding of the CBF cold response pathway. Here we further explore the structure–function relationships of the CBF proteins. The high degree of conservation of the CBF signature sequences among CBF transcription factors from diverse plant species suggested that they have an important functional role. Here we show that this is the case for the PKKPAGR sequence. Deletion of this sequence, or mutations within it, was found to greatly impair the ability of CBF1 to induce expression of target *COR* genes (Fig. 1, Supplementary Figs. 1 and 2). This functional impairment was not due to the mutations causing protein instability or loss of protein import into the nucleus. Rather, the PKKPAGR mutations were found to greatly impair the ability of CBF1 to bind to the CRT/DRE DNA recognition sequence.

A role for the PKKPAGR region in DNA binding was not anticipated as the AP2/ERF domain has typically been assigned the function of DNA binding for AP/ERF family proteins. Allen et al. [14] determined the 3D solution structure of the 60-amino-acid AP2/EREBP domain of the AtERF1 transcription factor and examined its interactions with its DNA-binding sequence, the GCC box. The AP2/ERF domain was found to be composed of a three-stranded anti-parallel structure followed by an α-helix that is packed parallel to the β-sheet. Arginine and tryptophan residues within the β-sheet were found to contact eight of nine consecutive nucleotides of the binding site within the major groove of the DNA. The high degree of sequence identity between CBF1 residues 46–107 and residues within the AtERF1 solution structure (Protein Data Bank entry 1gcc) implies the same arrangement of structural elements, with the full PKKP/RAGRxKFxETRHP sequence occurring immediately N-terminal to this domain. Thus, there was no prior reason to suspect an essential role for the PKKPAGR sequence in CBF1 binding to the CRT/DRE element.

A somewhat different picture was presented by Hao et al. [15]. These investigators found that the 10 amino acids immediately upstream of the AP2/ERF-binding domain of the EREBP2 transcription factor, a protein that also binds to the GCC box, were required for effective binding to the GCC box. However, unlike the CBF signature sequences, these 10 amino acids did not match the amino acids immediately flanking the AP2/ERF domains of other members of the EREBP protein family. Thus, the authors suggested that the 10 amino acids might not have a direct role in DNA binding, but instead, be required for the AP2/ERF domain to maintain an active conformation. The study did not include a functional analysis of the 10 amino acids, and thus, it was unclear whether specific amino acids were required, or that simply an N-terminal peptide extension was required to stabilize, in a non-specific manner, the AP2/ERF domain of EREBP2.

Our results show that the PKKPAGR sequence is required for CBF1 to effectively bind to the CRT/DRE element. Moreover, they show a need for specific amino acid side-chain chemistry at certain positions. In particular, the results show that the Arg37 and Phe40 residues are critical for CBF1 binding to the CRT/DRE regulatory element. Substituting Arg37 with Lys essentially abolished binding even though the two amino acids have very similar side-chain length and charge. In addition, the loss of DNA binding brought about by substituting Phe40 with Ala, and the preservation of binding (if not enhanced binding) observed with the Phe40 to Tyr substitution, indicates a requirement for a bulky hydrophobic side chain at this position. The challenge now is to determine the specific functions of the Arg37 and Phe40 side chains. One possibility is that they interact specifically with other CBF1 amino acids, stabilize the DNA-binding domain, and enable effective protein–DNA interaction. However, another possibility is that these residues interact directly with nucleotides surrounding the CRT/DRE core motif and contribute to binding site specificity. In this case, the residues would contribute to determining the composition of the CBF regulon. Future experiments will be directed at distinguishing between these possibilities.

Acknowledgments

This research was supported in part by grants to MFT from the NSF Plant Genome Project (DBI 0110124 and DBI 0701709), the Department of Energy (DE-FG02-91ER20021) and the Michigan Agricultural Experiment Station.

Appendix A. Supplementary data

Supplementary Table 1. Primers used for PKKPAGR mutagenesis.
Supplementary Table 2. Primers used in making constructs.
Supplementary Table 3. Primers used RKKFRET mutagenesis.
Supplementary Table 4. Oligomers used in EMSAs.

Supplementary Figure 1. Mutations in the PKKPAGR signature sequence impair CBF function. Northern blots of *Arabidopsis* overexpressing *CBF1* with mutations in the PKKPAGR region as shown in Fig. 1 were performed using *CBF1* and *COR78* as probes [(b) and (d)]. B6 is a transgenic line expressing the vector alone. 18S ribosomal RNA was used as a loading control. The blots were quantified [(a) and (c)] as described in Materials and methods.

Supplementary Figure 2. Mutations in the PKKPAGR signature sequence impair CBF function. Northern blots of *Arabidopsis* overexpressing *CBF1* with mutations in the PKKPAGR region as shown in Fig. 1 were performed using *CBF1* and *COR6.6* as probes [(b) and (d)]. B6 is a transgenic line expressing the vector alone. 18S ribosomal RNA was used as a loading control. The blots were quantified [(a) and (c)] as described in Materials and methods.

Supplementary Figure 3. Western blot analysis of protein extracts from *Arabidopsis* overexpressing CBF1-GFP:GUS constructs. Total protein extracts were made from 12-day-old seedlings, and Western blots were prepared as described in Materials and methods. Total protein (70 μg) was used per lane. Detection was with monoclonal antibody against GFP. Upper and lower panels represent short and long exposures of the same film, respectively.

Note: The supplementary material accompanying this article is available at (doi:10.1016/j.bbagrm.2009.11.017).

References

[1] V. Chinnusamy, J. Zhu, J.K. Zhu, Cold stress regulation of gene expression in plants, Trends Plant Sci. 12 (2007) 444–451.
[2] M.F. Thomashow, Plant cold acclimation: freezing tolerance genes and regulatory mechanisms, Annu. Rev. Plant Physiol. Plant Mol. Biol. 50 (1999) 571–599.
[3] S.J. Gilmour, D.G. Zarka, E.J. Stockinger, M.P. Salazar, J.M. Houghton, M.F. Thomashow, Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced *COR* gene expression, Plant J. 16 (1998) 433–442.
[4] J. Medina, M. Bargues, J. Terol, M. Perez-Alonso, J. Salinas, The *Arabidopsis* CBF gene family is composed of three genes encoding AP2 domain-containing proteins whose expression is regulated by low temperature but not by abscisic acid or dehydration, Plant Physiol. 119 (1999) 463–470.
[5] Q. Liu, M. Kasuga, Y. Sakuma, H. Abe, S. Miura, K. Yamaguchi-Shinozaki, K. Shinozaki, Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*, Plant Cell 10 (1998) 1391–1406.
[6] S.S. Baker, K.S. Wilhelm, M.F. Thomashow, The 5′-region of *Arabidopsis thaliana cor15a* has cis-acting elements that confer cold-, drought- and ABA-regulated gene expression, Plant Mol. Biol. 24 (1994) 701–713.
[7] K. Yamaguchi-Shinozaki, K. Shinozaki, A novel cis-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress, Plant Cell 6 (1994) 251–264.
[8] K. Maruyama, Y. Sakuma, M. Kasuga, Y. Ito, M. Seki, H. Goda, Y. Shimada, S. Yoshida, K. Shinozaki, K. Yamaguchi-Shinozaki, Identification of cold-inducible downstream genes of the *Arabidopsis* DREB1A/CBF3 transcriptional factor using two microarray systems, Plant J. 38 (2004) 982–993.
[9] J.T. Vogel, D.G. Zarka, H.A. Van Buskirk, S.G. Fowler, M.F. Thomashow, Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of *Arabidopsis*, Plant J. 41 (2005) 195–211.
[10] K.R. Jaglo-Ottosen, S.J. Gilmour, D.G. Zarka, O. Schabenberger, M.F. Thomashow, *Arabidopsis CBF1* overexpression induces *COR* genes and enhances freezing tolerance, Science 280 (1998) 104–106.
[11] J.L. Riechmann, E.M. Meyerowitz, The AP2/EREBP family of plant transcription factors, Biol. Chem. 379 (1998) 633–646.
[12] K.D. Jofuku, B.G. den Boer, M. Van Montagu, J.K. Okamuro, Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2, Plant Cell 6 (1994) 1211–1225.
[13] M. Ohme-Takagi, H. Shinshi, Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element, Plant Cell 7 (1995) 173–182.
[14] M.D. Allen, K. Yamasaki, M. Ohme-Takagi, M. Tateno, M. Suzuki, A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA, EMBO J. 17 (1998) 5484–5496.
[15] D. Hao, M. Ohme-Takagi, A. Sarai, Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in plant, J. Biol. Chem. 273 (1998) 26857–26861.
[16] Y. Sakuma, Q. Liu, J.G. Dubouzet, H. Abe, K. Shinozaki, K. Yamaguchi-Shinozaki, DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression, Biochem. Biophys. Res. Commun. 290 (2002) 998–1009.
[17] K.R. Jaglo, S. Kleff, K.L. Amundsen, X. Zhang, V. Haake, J.Z. Zhang, T. Deits, M.F. Thomashow, Components of the *Arabidopsis* C-repeat/dehydration-responsive element binding factor cold-response pathway are conserved in *Brassica napus* and other plant species, Plant Physiol. 127 (2001) 910–917.
[18] S.J. Gilmour, S.G. Fowler, M.F. Thomashow, *Arabidopsis* transcriptional activators CBF1, CBF2, and CBF3 have matching functional activities, Plant Mol. Biol. 54 (2004) 767–781.
[19] S.J. Clough, A.F. Bent, Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*, Plant J. 16 (1998) 735–743.
[20] D.G. Zarka, J.T. Vogel, D. Cook, M.F. Thomashow, Cold induction of *Arabidopsis* CBF genes involves multiple ICE (inducer of CBF expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature, Plant Physiol. 133 (2003) 910–918.
[21] J. Sambrook, D.W. Russell, Molecular Cloning, A laboratory manual, third ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2001.
[22] G. Church, G. Walter, Genomic sequencing, Proc. Natl. Acad. Sci. U. S. A. 81 (1984) 1991–1995.
[23] S. Li, M.F. Wilkinson, Site-directed mutagenesis: a two-step method using PCR and DpnI, Biotechniques 23 (1997) 588–590.
[24] W. Wang, B.A. Malcolm, Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis, Biotechniques 26 (1999) 680–682.

[25] G.E. An, P.R. Ebert, A. Mitra, S.B. Ha, Binary vectors, Plant Molecular Biology Manual, vol. A3, Kluwer Academic Publishing, Dordrecht, 1988, pp. 1–19.
[26] J. Overington, D. Donnelly, M.S. Johnson, A. Sali, T.L. Blundell, Environment-specific amino acid substitution tables: tertiary templates and prediction of protein folds, Protein Sci. 1 (1992) 216–226.
[27] K.E. Vlachonasios, M.F. Thomashow, S.J. Triezenberg, Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect *Arabidopsis* growth, development, and gene expression, Plant Cell 15 (2003) 626–638.
[28] M. Schena, D. Picard, K. Yamamoto, Vectors for constitutive and inducible gene expression in yeast, Methods Enzymol. 194 (1991) 389–398.
[29] J. Haseloff, K.R. Siemering, D.C. Prasher, S. Hodge, Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 2122–2127.
[30] B.P. Cormack, R.H. Valdivia, S. Falkow, FACS-optimized mutants of the green fluorescent protein (GFP), Gene 173 (1996) 33–38.
[31] T. Kleber-Janke, W.M. Becker, Use of modified BL21(DE3) *Escherichia coli* cells for high-level expression of recombinant peanut allergens affected by poor codon usage, Protein Expr. Purif. 19 (2000) 419–424.
[32] EJ. Stockinger, SJ. Gilmour, MF. Thomashow, *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 1035–1040.
[33] S. Abel, P. Oeller, A. Theologis, Early auxin-induced genes encode short-lived nuclear proteins, PNAS 91 (1994) 326–330.
[34] R. Lyck, U. Harmening, I. Hohfeld, E. Treuter, K.D. Scharf, L. Nover, Intracellular distribution and identification of the nuclear localization signals of two plant heat-stress transcription factors, Planta 202 (1997) 117–125.
[35] A.G. von Arnim, X.W. Deng, Light inactivation of *Arabidopsis* photomorphogenic repressor COP1 involves a cell-specific regulation of its nucleocytoplasmic partitioning, Cell 79 (1994) 1035–1045.
[36] R.J. Grebenok, E. Pierson, G.M. Lambert, F.C. Gong, C.L. Afonso, R. Haldeman-Cahill, J.C. Carrington, D.W. Galbraith, Green-fluorescent protein fusions for efficient characterization of nuclear targeting, Plant J. 11 (1997) 573–586.
[37] A. Kaffman, E.K. O'Shea, Regulation of nuclear localization: a key to a door, Annu. Rev. Cell Dev. Biol. 15 (1999) 291–339.
[38] J.C. Carrington, D.D. Freed, A.J. Leinicke, Bipartite signal sequence mediates nuclear translocation of the plant potyviral NIa protein, Plant Cell 3 (1991) 953–962.