New members of the tomato ERF family show specific expression pattern and diverse DNA-binding capacity to the GCC box element

Barthélémy Tournier[a,1], Maria Theresa Sanchez-Ballesta[a,1], Brian Jones[a], Edouard Pesquet[b], Farid Regad[a], Alain Latché[a], Jean-Claude Pech[a], Mondher Bouzayen[a,*]

[a] *UMR 990 INRA/INP-ENSAT, Pole de Biotechnologie Végétale, 24 Chemin de Borde Rouge, P.O. Box 107, 31326 Castanet Tolosan Cedex, France*
[b] *UMR 5546 CNRS/UPS, Pole de Biotechnologie Végétale, 24 Chemin de Borde Rouge, Auzeville P.O. Box 17, 31326 Castanet Tolosan Cedex, France*

Abstract **Four new members of the ERF (ethylene-response factor) family of plant-specific DNA-binding (GCC box) factors were isolated from tomato fruit (*LeERF1–4*). Phylogenetic analysis indicated that LeERF2 belongs to a new ERF class, characterized by a conserved N-terminal signature sequence. Expression patterns and *cis*/*trans* binding affinities differed between the LeERFs. Combining experimental data and modeled three-dimensional analysis, it was shown that binding affinity of the LeERFs was affected by both the variation of nucleotides surrounding the DNA *cis*-element sequence and the nature of critical amino acid residues within the ERF domain.**

Key words: Ethylene response factor; 3D modeling; *Lycopersicon esculentum*

1. Introduction

Phytohormones mediate development and stress responses by modulating the expression of specific subsets of hormone response genes. Ethylene, for example, affects the expression of a group of pathogenesis-related (*PR*) genes upon pathogen attack, wounding, abnormal temperatures, and drought stress [1,2]. A *cis*-acting promoter element, the GCC box, is both necessary and sufficient to confer ethylene responsiveness to a number of these *PR* genes in several plant species [3,4]. Ethylene response factors (ERFs), *trans*-acting factors that specifically bind the GCC box, have been identified in several plant species [3,4]. ERFs contain a highly conserved, plant-specific DNA-binding domain consisting of 58–59 amino acids that bind to DNA as a monomer, with high affinity. This domain is related to the AP2 domain in the *Arabidopsis* gene *APETALA2* [5], although ERF proteins contain a single DNA-binding domain whereas the APETALA2-type proteins typically contain two [6]. The ERF domain is comprised of three β-sheets and an α-helix [7]. Importantly, critical amino acids involved in DNA binding are not conserved between the APETALA2- and ERF-type sequences [7,8].

As many as 124 ERF domain-containing proteins are predicted to be encoded in *Arabidopsis* [9]. Only a few members of this family have been characterized and most of these have been shown to participate in stress and/or hormonal responses [4,8,10,11]. In tomato, the ERF domain-containing *Pti4/5/6* gene products bind to the pathogenesis-related Pto protein kinase [12]. To date, they represent the only *ERF* genes characterized in this species. Over-expression of *Pti4* in *Arabidopsis* induces the expression of GCC box-containing genes [13] and confers enhanced resistance to pathogen attack [14].

In order to further examine the role of the *ERF* genes in the tomato, four new members were isolated and analyzed with regard to (i) their primary sequence, (ii) their expression, (iii) their binding capacity to DNA and (iv) their DNA-binding domain conformational structure.

2. Materials and methods

2.1. Cloning cDNAs of four LeERFs

Degenerate oligonucleotides designed from ERF domains were used in a polymerase chain reaction (PCR) with a tomato fruit cDNA template (AP2-5′: 5′-CCRTGGGGRAAATKKGCGGCK-3′; AP2-3′: 5′-CATAAGCVAVAKBGCRGCTTCYTC-3′). cDNA fragments *LeERF1–4* (accession numbers: AY192367–AY192370) were chosen for further characterization. Full-length cDNAs were isolated by a PCR strategy using a tomato fruit cDNA library. The four full-length coding regions were cloned into the pJG4.5 vector (Clontech).

2.2. Electromobility shift assay (EMSA)

Full-length *LeERF*s were generated by in vitro translation using the TNT-T7 Quick for PCR DNA coupled transcription/translation system and labeled with Transcend™ Biotin-Lysyl-tRNA as described by the manufacturer (Promega). A T7 RNA polymerase consensus binding site and a polyA tail were introduced into polynucleotide primer sequences, TNT-PJG4.5-5′ (5′-CAACGGCTAATACGACTCACTATAGGGTATCCCACCCCTCCTACCCTTATGATGTGCC-3′) and TNT-PJG4.5-3′ (5′-$(T)_{30}$ACCCGACAACCTTGATTGGAGACTTG-3′). In vitro translated proteins were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (14%), transferred, hybridized with streptavidin-horseradish peroxidase and revealed by enhanced chemiluminescence as described by the manufacturer. Probes were prepared by Klenow fragment filling with [α-^{32}P]dCTP of the hybridized oligonucleotides and gel purified. Sequences corresponded to Fig. 5B, with an overhang made by an *Spe*I site on the 5′ end of the forward oligonucleotide, and an *Xba*I site on the 5′ end of the reverse oligonucleotide. Binding reaction in 20 μl was performed with 2 μl of in vitro translated proteins as described in [8]. Binding reactions were resolved on a 6% polyacrylamide gel, run in 0.5×TBE buffer at 10 V/cm for 2 h, then gels were dried and exposed to X-ray film at −80°C.

*Corresponding author. Fax: (33)-5-62 19 35 73.
E-mail address: bouzayen@ensat.fr (M. Bouzayen).

[1] These authors contributed equally to this work.

2.3. Plant material

Tomato plants (cvs. *microtom* and *Ailsa Craig*) were grown in soil under standard greenhouse conditions. Samples were collected from four independent plants. Four-week-old *microtom* plants were used for all treatments. The ethylene treatment was performed as described previously [15]. For the wounding experiment, plants were wounded in situ with a razor blade and the plants were left under growth conditions. For the 1-methylcyclopropene (1-MCP) experiment, plants were pre-treated with 5 μl/l of 1-MCP in sealed chambers for 16 h and then subjected to the same wounding treatment.

2.4. RNA extraction, Northern blots and RT-PCR

Total RNA was extracted by the cetyltrimethyl-ammonium bromide method and reverse transcription (RT) PCR analysis were carried out as previously described [16]. RNA samples were fractionated, blotted and hybridized as described in [17]. Probes for *LeERF*s and the *E8* 3′ untranscribed region (UTR) were prepared by random primer labeling with [α-^{32}P]dCTP.

2.5. Molecular modelling and multiple 3D alignment

Homology modeling was performed using program facilities at the EXpert Protein Analysis SYstem proteomics server (EXPASY; http://expasy.ch) of the Swiss Institute of Bioinformatics. SWISS-MODEL Protein Modeling Server Version 36.0003 was used for sequence homology-based 3D modelling [18,19]. All LeERF models were based on AtERF1 (PDB accession number: 1GCC/2GCC/3GCC). 3D multiple alignment of LeERFs was done using Swiss-Pdb Viewer 3.7 where all LeERF layers of the calculated PDB model were superimposed together in order to analyze divergent amino acid side chains.

3. Results

3.1. Isolation and sequence analysis of four ERFs from tomato

Four partial cDNA clones, *LeERF1–4*, were isolated from tomato fruit cDNAs using degenerate primers targeted to the highly conserved ERF domain. All of the traits commonly associated with ERFs were identified in the full-length predicted LeERF sequences (Fig. 1). In addition to the ERF domain, these include putative nuclear localization signals and acidic domains that have been shown in other species to act as activation domains [20]. Phylogenetic analysis (Fig. 1) indicated that the LeERFs could be assigned to previously described ERF classes [8]. LeERF1 belongs to the highly conserved ERF class I. LeERF3 belongs to the class II putative repressor ERFs containing a conserved EAR repressor motif (ERF-associated amphiphilic repression) [11]. LeERF4 is related to ERF class III proteins, although it lacks an acidic domain in its shorter C-terminal moiety. LeERF2 and putative homologues from different plant species are well conserved, even outside the ERF domain. They display a novel and highly conserved N-terminal motif of unknown function (MCGGAII/L) (see complementary data) and define a new ERF class named class IV (Fig. 1).

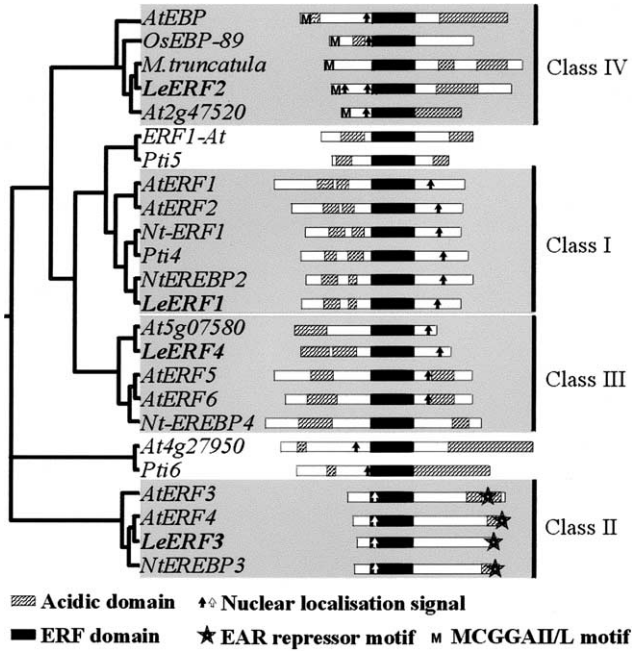

Fig. 1. Phylogenetic tree of close homologues of the *LeERF*s. Classes I to III are from [8]. *Medicago truncatula* (TC32494) is from a TIGR EST database. Other gene names are accessible through GenBank.

3.2. Organ- and ripening-associated expression of LeERFs in tomato plants

LeERF transcript accumulation indicated a specific pattern of expression for each *LeERF* (Fig. 2A). *LeERF1* transcripts were detected in stems and those for *LeERF3* in stems and leaves, and to a lesser extent in open flowers. *LeERF4* transcripts accumulated strongly in leaves. Interestingly, *LeERF2* transcripts were most abundant in ripe fruit. To further address the role of the *LeERF*s during tomato fruit development and ripening, gene expression analysis was performed on a tomato variety (*Ailsa Craig*) more suited to the study of fruit ripening and where tomato ripening mutants such as *Nr* (Never-ripe), *nor* (non-ripening) and *rin* (ripening inhibitor) are available. Fig. 2B indicates that *LeERF2* is the only one to show ripening-associated expression and that *LeERF2* transcripts did not accumulate in the ripening mutants, thus mimicking the expression of *E8*, a tomato ripening-associated gene [21].

3.3. Ethylene- and wound-associated expression of LeERF genes

As ERFs are thought to mediate ethylene and stress responses, wound- and ethylene-induced *LeERF* transcript accumulation was examined in tomato leaves. The expression of wound- and ethylene-induced *PR1b1*, a basic *PR* gene was analyzed in parallel to confirm the efficacy of the treatments [22]. Fig. 3 shows that *LeERF1/4* and *PR1b1* were strongly induced by ethylene, while *LeERF2/3* transcripts were unaffected. *LeERF1/2/4* and *PR1b1* transcripts accumulated in response to wounding, while *LeERF3* transcript levels diminished. Treatment with 1-MCP, an inhibitor of ethylene perception, prior to wounding strongly reduced transcript accumulation of *LeERF1/4* and completely inhibited *PR1b1*. By contrast it stimulated the accumulation of *LeERF2/3* transcripts (Fig. 3).

3.4. LeERFs are GCC box-binding proteins

To assess the capacity of the four LeERFs to bind a GCC box, DNA-binding experiments were performed using the same amount of in vitro translated proteins (Fig. 4A). All four LeERF proteins were able to specifically bind the tomato osmotin promoter GCC box (Fig. 4C). Binding activity was dramatically reduced by competition with an unlabeled probe and completely abolished when both G residues within the GCCGCC motif were replaced by T residues. Fig. 4C also shows strong differences in LeERF-binding activities. Several GCC box-containing sequences from a variety of plant species

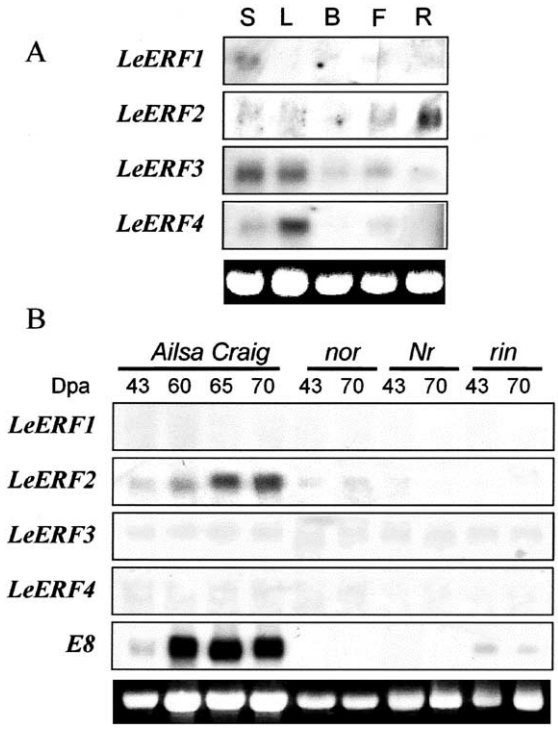

Fig. 2. Northern blots hybridized with *LeERF*s and *E8* 3′-UTR probes on plant organs (A) and during fruit ripening (B). Equivalence of lane loading is demonstrated by 18S rRNA ethidium bromide staining. Stems (S), leaves (L), buds (B), open flowers (F) and red fruits (R) are from *microtom*. Fruit are from tomato after 43 (immature), 60 (mature green), 65 (breaker) and 70 (red) days post anthesis (dpa) and from *Ailsa Craig* mutants *Nr* (Never-ripe), *rin* (ripening inhibitor) and *nor* (non-ripening) after 43 and 70 dpa.

including tomato, tobacco and *Arabidopsis thaliana* were used for mobility gel shift assays (Fig. 4B). The strongest binding activity was found with tobacco chitinase and tomato osmotin GCC boxes, though the latter showed slightly weaker binding (Fig. 4D). LeERF1/3/4 proteins showed weak binding to the tomato chitinase GCC box, while no binding was detected

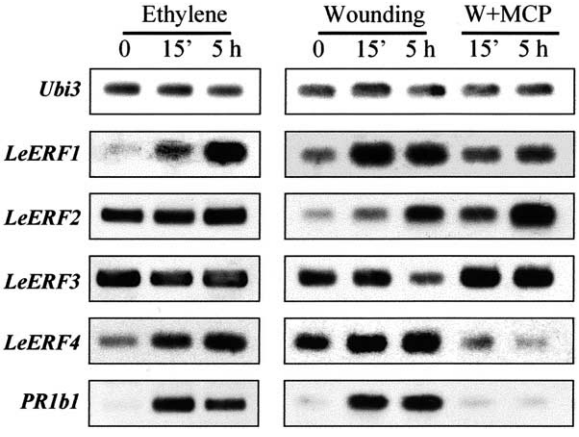

Fig. 3. RT-PCR analysis of *LeERF* transcript accumulation in fully expended leaves treated with ethylene, or mechanically wounded with or without 1-MCP pre-treatment. In each experiment the internal reference *ubi3* was co-amplified with *LeERF* or *PR1b1*. Results are representative of three separate experiments.

with LeERF2. None of the four LeERFs were able to bind to the GCC-like elements from the tomato *cel5* glucanase and *Arabidopsis* DRE/CRT elements. A perfect conservation of the GCCGCC core motif was strictly required for binding and sequences surrounding the core GCC box strongly influenced binding affinity (Fig. 4B). Regardless of the origin of the GCC box, the strongest binding was exhibited by LeERF4 and LeERF3, and the weakest by LeERF1/2.

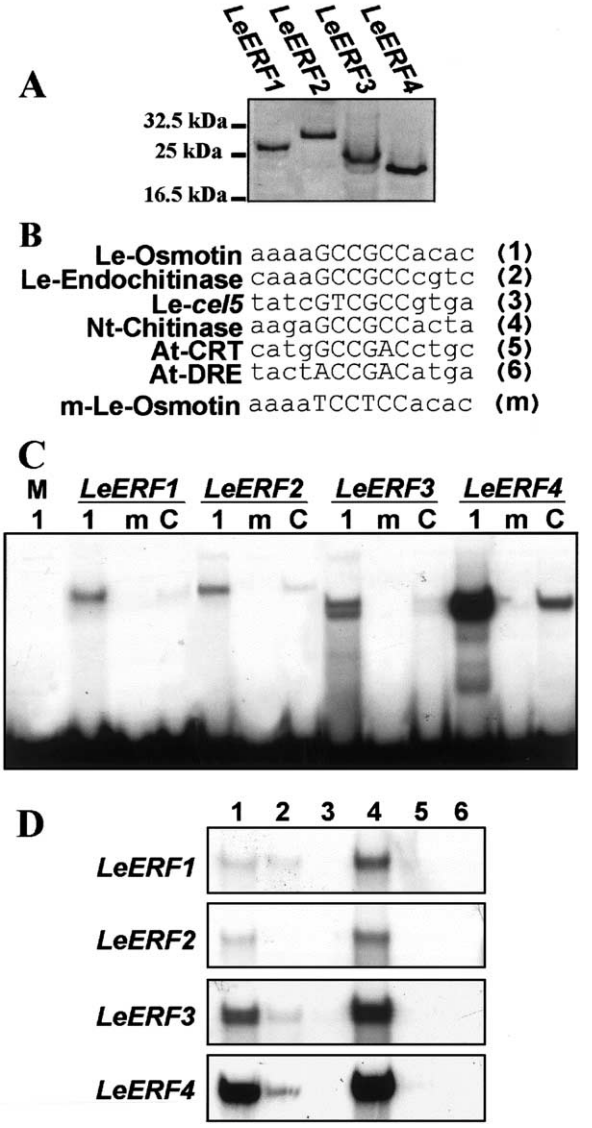

Fig. 4. A: Western blots with 2 μl of in vitro translated LeERF. B: *Cis*-elements used in EMSA were the following: (1) Le-osmotin (AF093743), (2) Le-endochitinase (A32906), (3) tomato β-1,4-glucanase cel5 (AF077340), (4) Nt-chitinase (X16938), (5) At-CRT from Cor15a promoter (U01377) and (6) At-DRE element from rd29A promoter (D13044) and (m) mutated Le-osmotin. C: EMSAs were performed with 2 μl of in vitro translated LeERF1/2/3/4 proteins and 1 ng of [α-^{32}P]dCTP-labeled either wild-type (1) or mutated (m) Le-osmotin probes. The competition assay (c) was performed as in (1) except that 1 μg of Le-osmotin-unlabeled probe was added in the incubation medium. The control mock experiment (M) consisted of 2 μl of the in vitro translation reaction mixture minus template DNA. These in vitro 'empty' translation products were then incubated with Le-osmotin-labeled probe. D: Each in vitro translated LeERF was challenged with the probes listed in B.

3.5. 3D modeling of LeERF DNA-binding domain

A 3D model alignment (Fig. 5), based on the core homology of AtERF1, was performed in order to address the relationship between the conformational structure of the DNA-binding domains and the binding activity of the four LeERFs [7]. Homology with AtERF1 within this domain totaled 84.1%, 76.35%, 71.3% and 73.8% for LeERF1–4, respectively. Models were highly reliable based on the analysis of their Ramachandran diagram compared with the crystal structure of AtERF1 (see complementary data). This showed 80.7–79.1% in the core region, 14.5–12.9% in the allowed region and 8–6.4% found in the disallowed region for LeERF1–4. Gaps and strong amino acid variations in the linear alignment (Fig. 5A) located in the α-helix and the β-turn 2 should not account for the differences in binding activity as these regions are not directly involved in interaction with DNA [7]. Com-

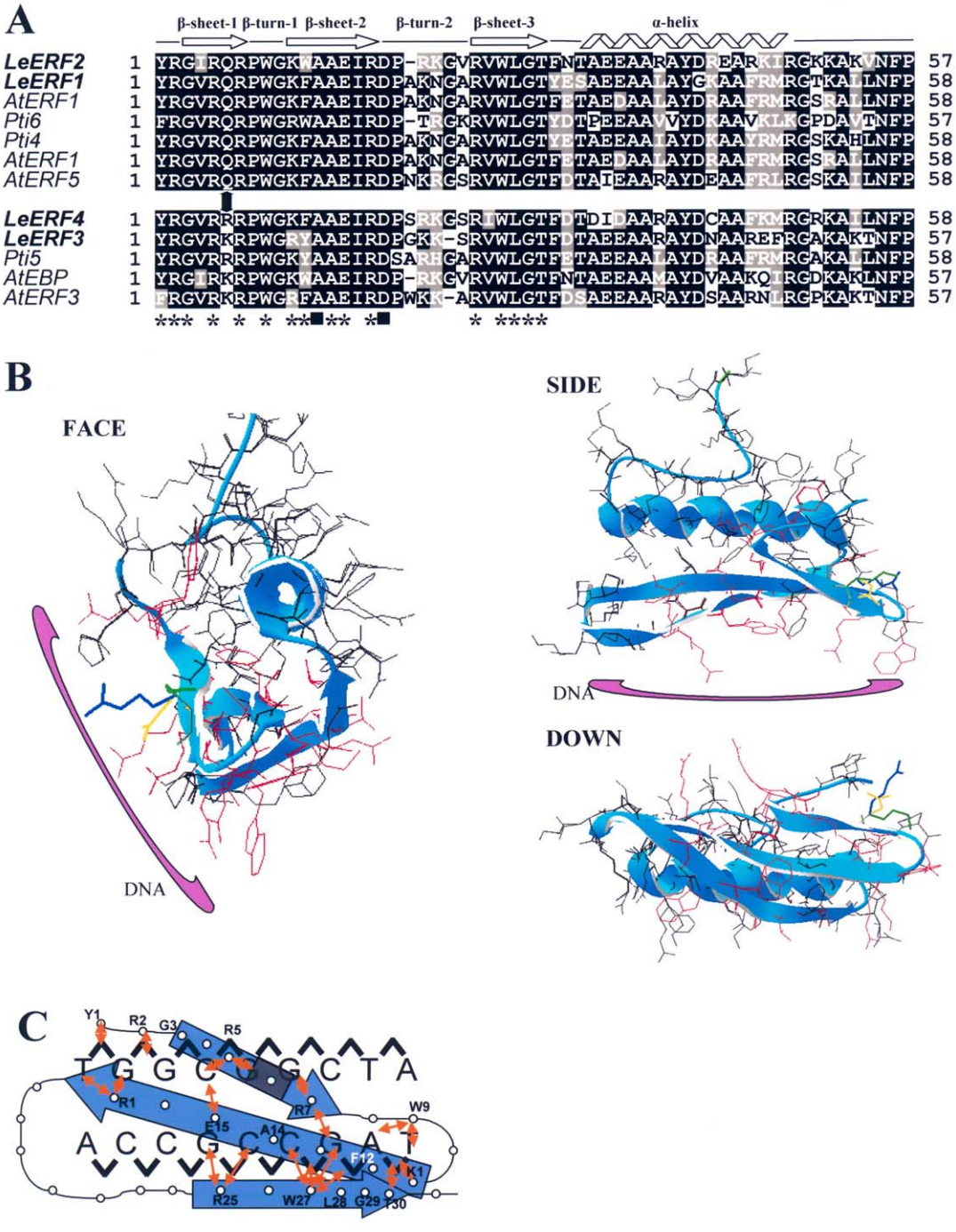

Fig. 5. A: Multiple alignment of LeERF DNA-binding domains. Stars indicate amino acids interacting with DNA in AtERF1 [7]. Black squares indicate amino acids different between ERF and DREB types [9]. B: 3D alignment of LeERF1/3/4 DNA-binding domain. Residue 6 is represented in yellow for LeERF1, green for LeERF3 and blue for LeERF4. Amino acids shown to interact with DNA, represented in red, are conserved. C: Schematic representation of the three β-sheets interacting with DNA based on AtERF1 structure [7]. β-sheets are represented by blue arrows. Residue 6 is shadowed within the first β-sheet. Red arrows represent interactions.

paratively, within the β-sheets involved in protein/DNA interactions all amino acids are either strictly conserved or replaced by similar residues, with the exception of residue 6 (Fig. 5B). Residue 6, located between the two R residues involved in the binding to G-1 and C-3 nucleotides of the GCC box, varies from an uncharged amino acid in LeERF1/2 to a basic amino acid in LeERF3/4 (Fig. 5C). Strikingly, the shift from basic highly charged residues (K and R in LeERF3 and 4, respectively) to uncharged residues (Q in LeERF1/2) correlated with a net decrease in binding activity. 3D multiple alignment suggests that spatial orientation of the side chain of this residue 6 influences binding affinity. That is, LeERF4, which showed the highest binding affinity, also harbors the R6 basic residue, deploying the most accessible side chain bearing a highly polar guanidinium group. In comparison, the side chain of the LeERF3 K6 basic residue appeared to be less deployed towards the DNA target and bears an NH_2 group with weaker polarity.

4. Discussion

Sequence analysis clearly indicated that *LeERF1–4* belong to the large ERF family of transcription factors unique to plants. Furthermore, peptide sequence analysis revealed that the two amino acids (A13, D18 in Fig. 5A) shown to define ERF-, but not DREB-type transcription factors are conserved in the four *LeERF*s isolated in this study [9]. Based on their structural organization and considering the transcriptional activities of the previously characterized ERFs, the four *LeERF*s fall into different classes. Class I (*LeERF1*) and class III (*LeERF4*) are activators, whereas class II are repressors (*LeERF3*). *LeERF2* and its homologues from monocot to dicot species are characterized by a unique N-terminal signature MCGGAII/L and cannot be assigned to any of the three previously defined ERF classes [8]. This N-terminal motif is only found in ERF genes, including *AtEBP* from *Arabidopsis* [23] and *OsEBP-89* from rice [20]. *LeERF2*, therefore, defines the new ERF class IV. Though the function of the MCGGAII/L motif has not been established, deletion studies indicated that it is required neither for nuclear localization nor for binding to the GCC box (data not shown).

All of the *LeERF*s showed specific transcript accumulation patterns. While most ripening-associated genes in climacteric fruit display ethylene responsiveness [24], none have been shown to contain a GCC motif in their promoter region [25–27]. The ripening-associated expression of *LeERF2*, a GCC box-binding ERF, described here for the first time, indicates that its target genes are likely to play a role in the ripening process.

Typically, ERFs mediate ethylene-regulated responses to both biotic and abiotic stresses [6,28]; consistently, in this study, *LeERF1/4* are up-regulated by ethylene and wounding. *LeERF2/3* were not found to be ethylene-responsive, but wounding induced *LeERF2* expression while it negatively regulated that of *LeERF3*. However, for all LeERFs, the wound response was at least partially mediated by ethylene.

All four LeERFs were capable of specific binding to GCC box-containing *cis*-elements. To date, only three GCC boxes have been characterized from tomato genes, basic endochitinase, *cel5* basic glucanase and osmotin. Surprisingly, LeERFs displayed weaker binding to these tomato GCC boxes than to the tobacco chitinase. Our data show not only that the binding required a perfectly conserved core GCCGCC box, but also that the sequences flanking this core box affected binding efficiency. Other tomato genes may contain GCC boxes with higher affinity to the isolated LeERFs.

Studies of the 3D alignment of modelled DNA-binding domains of the LeERFs indicated that variation in the binding affinity correlates with variation in the nature of residue 6, embedded between two amino acids shown to interact with DNA (Fig. 5). Changing residue 6 from basic charged to uncharged did not alter the specificity of interaction with the *cis*-element, but greatly decreased binding affinity. Residue 6 may interact with the DNA phosphopentose skeleton and thus, depending on its charge and accessibility, could affect the stability of the DNA–protein complex.

Our data revealed for the first time, beside amino acid residues directly engaged in the interaction with the target DNA, that the nature of critical amino acid 6 can greatly impact *cis*/*trans* binding affinity.

Acknowledgements: This work was supported in part by the EU (FAIR CT-95 0225), INRA (Action Transversale Structurante tomate, 2001–2003), and the Midi-Pyrénées Regional Council (Grants 99009080 and 01002710) to M.B. M.T.S.-B. was supported by a postdoctoral fellowship from INRA and the Ministerio de Educación y Ciencia, Spain.

References

[1] Abeles, F.B., Morgan, P.W. and Saltveit Jr., M.E. (1992) Ethylene in Plant Biology, 2nd edn., Academic Press, San Diego, CA.
[2] Leclercq, J., Adams-Phillips, L.C., Zegzouti, H., Jones, B., Latché, A., Giovannoni, J.J., Pech, J.C. and Bouzayen, M. (2002) Plant Physiol. 130, 1132–1142.
[3] Ohme-Takagi, M. and Shinshi, H. (1995) Plant Cell 7, 173–182.
[4] Solano, R., Stepanova, A., Chao, Q. and Ecker, J.R. (1998) Genes Dev. 12, 3703–3714.
[5] Jokofu, K.D., den Boer, B.G.W., Van Montagu, M. and Okamuro, J.K. (1994) Plant Cell 6, 1211–1225.
[6] Riechmann, J.L. and Meyerowitz, E.M. (1998) Biol. Chem. 379, 633–646.
[7] Allen, M.D., Yamasaki, K., Ohme-Takagi, M., Tateno, M. and Suzuki, M. (1998) EMBO J. 17, 5484–5496.
[8] Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H. and Ohme-Takagi, M. (2000) Plant Cell 12, 393–404.
[9] Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2002) Biochem. Biophys. Res. Commun. 290, 998–1009.
[10] Onate-Sanchez, L. and Singh, K.B. (2002) Plant Physiol. 128, 1313–1322.
[11] Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H. and Ohme-Takagi, M. (2001) Plant Cell 13, 1959–1968.
[12] Zhou, J., Tang, X. and Martin, G.B. (1997) EMBO J. 16, 3207–3218.
[13] Wu, K., Tian, L., Hollingworth, J., Brown, D.C. and Miki, B. (2002) Plant Physiol. 128, 30–37.
[14] Gu, Y.Q., Wildermuth, M.C., Chakravarthy, S., Loh, Y.T., Yang, C., He, X., Han, Y. and Martin, G.B. (2002) Plant Cell 14, 817–831.
[15] Jones, B., Frasse, P., Olmos, E., Zegzouti, H., Li, Z.G., Latché, A., Pech, J.C. and Bouzayen, M. (2002) Plant J. 32, 603–613.
[16] Zegzouti, H., Jones, B., Frasse, P., Marty, C., Maitre, B., Latché, A., Pech, J.C. and Bouzayen, M. (1999) Plant J. 18, 589–600.
[17] Sanchez-Ballesta, M.T., Zacarias, L., Granell, A. and Lafuente, M.T. (2000) J. Agric. Food Chem. 48, 2726–2731.
[18] Peitsch, M.C. (1996) Biochem. Soc. Trans. 24, 274–279.
[19] Guex, N. and Peitsch, M.C. (1997) Electrophoresis 18, 2714–2723.
[20] Yang, H.J., Shen, H., Chen, L., Xing, Y.Y., Wang, Z.Y., Zhang, J.L. and Hong, M.M. (2002) Plant Mol. Biol. 50, 379–391.

[21] Deikman, J. and Fischer, R.L. (1988) EMBO J. 11, 3315–3320.
[22] Tornero, P., Gadea, J., Conejero, V. and Vera, P. (1997) Mol. Plant Microbe Interact. 10, 624–634.
[23] Buttner, M. and Singh, K.B. (1997) Proc. Natl. Acad. Sci. USA 94, 5961–5966.
[24] Giovannoni, J. (2001) Annu. Rev. Plant Physiol. Plant Mol. Biol. 52, 725–749.
[25] Deikman, J., Xu, R., Kneissl, M.L., Ciardi, J.A., Kim, K.N. and Pelah, D. (1998) Plant Mol. Biol. 37, 1001–1011.
[26] Nicholass, F.J., Smith, C.J., Schuch, W., Bird, C.R. and Grierson, D. (1995) Plant Mol. Biol. 28, 423–435.
[27] Xu, R., Goldman, S., Coupe, S. and Deikman, J. (1996) Plant Mol. Biol. 31, 1117–1127.
[28] Singh, K.B., Foley, R. and Onate-Sanchez, L. (2002) Curr. Opin. Plant Biol. 5, 430–436.