Molecular Plant
Letter to the Editor

CellPress
PARTNER JOURNAL

Crystal Structure of the *Arabidopsis thaliana* L1L/NF-YC3 Histone-fold Dimer Reveals Specificities of the LEC1 Family of NF-Y Subunits in Plants

Dear Editor,

The NF-Y transcription factor is a heterotrimer formed by evolutionarily conserved subunits: NF-YA, NF-YB, and NF-YC. NF-YB and NF-YC harbor a histone fold domain (HFD), structurally similar to that of nucleosome core histones, and form a tight dimer (Romier et al., 2003). NF-YA binds to the NF-YB/NF-YC dimer and provides exquisite sequence specificity for recognizing and binding the CCAAT box (Huber et al., 2012; Nardini et al., 2013), an important DNA regulatory element of all eukaryotes (Dolfini et al., 2009). In plants, each NF-Y subunit is expanded to 8–14 genes (Laloum et al., 2013). Within the NF-YB genes, *LEAFY COTYLEDON1* (*LEC1*) and *LEC1-LIKE* (*L1L*) form a conserved subfamily (Braybrook and Harada, 2008), originally identified in genetic experiments in *Arabidopsis thaliana* (At-*LEC1*, corresponding to At-*NF-YB9*, and At-*L1L*, corresponding to At-*NF-YB6*, respectively) for their key roles in embryo maturation (Kwong et al., 2003; Lee et al., 2003). Considerations based on available structures and sequence alignments of At-NF-YBs raise questions as to the capacity of LEC1/L1L to bind DNA. In addition, activation of seed-specific genes by L1L, such as *CRC*, *SUS2*, and the lipidogenic *FAD3*, involves indirect promoter tethering by interactions with bZIP67, binding to abscisic acid response elements (ABREs) (Yamamoto et al., 2009; Mendes et al., 2013).

To understand the specificities of plant LEC1-like NF-YBs, it is mandatory to unveil their structural features. To this aim, we crystallized and solved the structure of the core domains of At-L1L and At-NF-YC3 subunits at 2.3 Å resolution (see Supplemental Materials and Methods, Supplemental Table 1, and Supplemental Figures 1A and 2). As expected, both subunits adopt an HFD structure (helix α1–loop L1–helix α2–loop L2–helix α3) and interact in a head-to-tail fashion, forming a classical histone-like pair (Figure 1A). Both At-L1L and At-NF-YC3 show the typical structural features specific for NF-YB and NF-YC subunits, as highlighted by sequence/structure comparison with the mammalian NF-YB/NF-YC heterodimer (sequence identity of 63.4% for NF-YB and 73.8% for NF-YC; RMSD of 0.95 Å calculated over 170 Cα pairs) (Figure 1A) including, among others, the presence of an intra-chain Arg-Asp bidentate pair in both subunits (Arg114-Asp121 in At-L1L, and Arg121-Asp128 in At-NF-YC3), and the presence of a Trp residue, absolutely conserved in NF-YCs (Trp113 in At-NF-YC3), nestled in a hydrophobic pocket at the interface between the L1/L2 loops pair of NF-YB and NF-YC (see Supplemental Figures 2 and 3). The divergent regions that distinguish the protein backbone of At-L1L/NF-YC3 from that of mammalian NF-YB/NF-YC heterodimer are localized at the N and C termini (Figure 1A).

We then analyzed the surface of the At-L1L/NF-YC3 dimer to search for the presence of positively charged residues in the in α1, L1, and L2 regions, predicted to be involved in histone-like DNA contacts (see Supplemental Figure 2). Figure 1B (top view) shows that, bar minor lateral differences essentially related to At-L1L residues Asp84 in helix α2 and His79 in loop L1 (discussed below), the overall surface of the At-L1L/NF-YC3 dimer is very similar to mammalian NF-YB/NF-YC, thus perfectly suited for binding and bending the DNA (Nardini et al., 2013). We noticed also the presence of a wide, acidic surface involving the At-L1L helix α2 and the At-NF-YC3 helix αC, which matches perfectly the location of the NF-YA binding groove in NF-YB/NF-YC (Figure 1B, side view). Thus, *Arabidopsis* NF-YAs are expected to bind the HFD heterodimer as mammalian NF-YA does. Indeed, almost the totality of the mammalian HFD residues interacting with NF-YA and with DNA are conserved in At-L1L/NF-YC3 (see Supplemental Figure 2).

We wished to ascertain whether trimers composed of plant NF-Y subunits indeed bind DNA. For this purpose, we performed electrophoretic mobility shift assay (EMSA) experiments, using the At-L1L/NF-YC3 HFD dimer and the At-NF-YA6 subunit, widely expressed in seeds. These experiments extend the information previously published (Calvenzani et al., 2012), where At-L1L trimerization and DNA binding were tested in EMSA by using recombinant mouse NF-YA and NF-YC.

In the absence of genetic information on a CCAAT-containing promoter expressed in seeds, hence potentially targeted by our At-NF-Y trimer, we used a fluorescently labeled high affinity CCAAT box probe derived from the human HSP70 promoter (see Supplemental Materials and Methods). EMSA analysis with increasing amount of the HFD heterodimer or At-NF-YA6, in the presence of the respective counterpart, shows robust DNA binding only in the presence of At-L1L/NF-YC3/NF-YA6 trimers and not of the separated subunits (Figure 1C). Furthermore, the bands corresponding to the At-NF-Y/DNA complex are specific for CCAAT, since it is competed away by an unlabeled CCAAT oligo but not by the one with an unrelated sequence (Figure 1D). The Asn67→Asp mutation in helix α1 of At-L1L, a key contact site between NF-YB and the DNA phosphate backbone in the mammalian NF-Y/DNA complex (Sinha et al., 1996; Zemzoumi et al., 1999; Nardini et al., 2013), abrogates the ability of At-L1L/NF-YC3 to bind DNA in EMSA experiments, without affecting heterodimerization. A similar result is obtained with the Glu96→Arg mutation in helix α2 of At-L1L, a site

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and IPPE, SIBS, CAS.

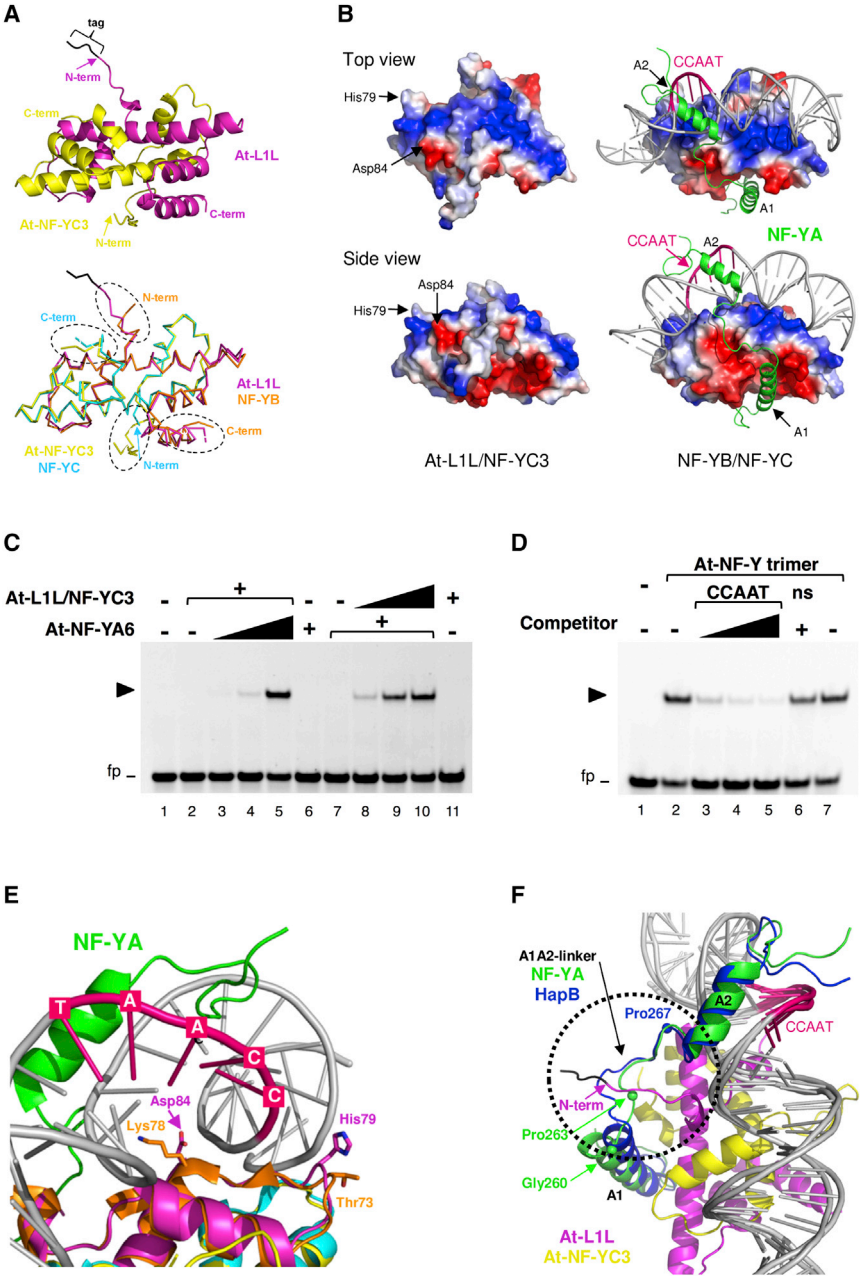

Figure 1. Structure and DNA Binding of At-L1L/NF-YC3.

(A) Ribbon diagram showing the histone fold dimer, and structural comparison of At-L1L/NF-YC3 with NF-YB/NF-YC (PDB: 4CSR), where the most divergent regions are highlighted by dashed circles.

(B) Electrostatic surface of the At-L1L/NF-YC3 dimer (top and side views) compared with that of NF-YB/NF-YC (PDB: 4AWL). Blue and red colors indicate positively and negatively charged regions, respectively. At-L1L His79 and Asp84 are indicated. NF-YA and the bound DNA are represented in ribbon and stick models, with the CCAAT box highlighted.

(C) Trimerization and DNA binding of At-L1L/NF-YC3 dimer with At-NF-YA6 was assessed by EMSA with Hsp70 CCAAT box DNA probe by addition of increasing doses of the respective NF-Y subunit(s) counterpart, as indicated. DNA-binding mixes containing the HFD dimer (60 nM; lanes 2–5) or At-NF-YA6 (60 nM; lanes 7–10) were added with protein dilution buffer (DB; lanes 2, 7) or increasing amounts (20, 60 or 180 nM) of At-NF-YA6–6His or At-L1L/NF-YC3 (lanes 3–5; 8–10, respectively). As negative controls, At-NF-YA6–6His (lane 6) or At-L1L/NF-YC3 (lane 11) were incubated alone with the probe, at the highest concentration of the dose curve (180 nM). Lane 1, probe alone DNA-binding mix without NF-Y subunits added. The At-NF-Y/DNA complex is indicated by an arrowhead. fp, free probe.

(D) CCAAT specificity of the At-L1L/NF-YC3/NF-YA6 NF-Y trimer/Hsp70 DNA complex was assessed in EMSA by addition of increasing doses of unlabeled CCAAT DNA or unrelated oligo (ns, nonspecific) competitors. DNA-binding mixes containing 20 nM labeled Hsp70 CCAAT probe alone (lanes 1, 2, 7) or with increasing amounts of unlabeled oligo competitors (lanes 3–5, Hsp70 CCAAT competitor, 100, 500, or 2000 nM; lane 6, ns competitor, 2000 nM) were added with At-NF-Y trimer subunits (lanes 2–7, 60 nM At-L1L/NF-YC3 and 180 nM At-NF-YA6-6His). Lane 1, probe alone DNA-binding mix without NF-Y subunits added. The At-NF-Y/DNA complex is indicated by an arrowhead. fp, free probe.

(E) Position of the Asp84 and His79 residues. The superposition of the At-L1L/NF-YC3 dimer on the NF-Y/DNA complex (PDB: 4AWL) shows the position of At-L1L Asp84 and His79 relative to DNA and to the corresponding NF-YB residues Lys78 and Thr73.

(F) Structure of the At-L1L N terminus. The superposition of the At-L1L/NF-YC3 dimer on the mammalian and *A. nidulans* NF-Y trimers (PDB: 4AWL and 4G92, respectively) shows the overlay of the At-L1L N terminus relative to the NF-YA and HapB A1A2-linkers. The dashed circle highlights the structural elements described in the text.

demonstrated to be diagnostic for NF-YA association and trimerization (Sinha et al., 1996) (Supplemental Figures 1C and 4A). Altogether, the data indicate that At-L1L and At-NF-YC3 are *bona fide* NF-Y subunits able to form a canonical HFD heterodimer, with a trimerization and a DNA-binding mode conforming to the mammalian NF-Y.

Nevertheless, significant differences also emerge. Among the HFD residues involved in DNA interactions, Asp84 of At-L1L (Asp55 in LEC1) stands out, since it is typically Lys/Arg in other plant, yeast, and mammalian NF-YBs (Supplemental Figure 2A). In mammalian NF-YB, the corresponding Lys78, located in helix α2, contacts the DNA phosphate backbone, 2 bp upstream of CCAAT (Nardini et al., 2013). This residue was a major focus of our attention, since genetic experiments pinpointed Asp55 as crucial for LEC1 function *in vivo*: an Asp→Lys mutation led to a dramatic loss of LEC1 activity, while a Lys→Asp substitution was sufficient to confer partial LEC1 features to the more "canonical" At-NF-YB3 (Lee et al., 2003). EMSA experiments on At-L1L wt and Asp84→Lys mutant show that both proteins

bind DNA with similar affinities (Supplemental Figures 1C and 4A). The current dimer structure, superimposed on the quaternary NF-Y/CCAAT complex (Figure 1E), suggests that the electrostatic repulsion between the negatively charged Asp84 and the DNA phosphate backbone would favor a slightly shifted DNA trajectory, stabilized on the opposite side of the double helix by His79, which could compensate for the missing DNA contact resulting from the Lys→Asp substitution at position 84 in At-L1L (or 55 in LEC1) (Figure 1E). Notably, this His residue is absolutely conserved in plants and diagnostic for LEC1 family members (see Supplemental Figure 2A). Note that the corresponding mammalian Thr73 is not in contact with DNA (Figure 1E), and essentially all "canonical" NF-YBs in plants have an Asn residue at this position (see Supplemental Figure 2A). In summary, the At-L1L model posits that Asp84 would "shift" DNA toward His79, making contacts with the complementary strand, also upstream of CCAAT. This structural variability in the protein–DNA interactions at the At-L1L Asp84 site (Lys in standard NF-YB subunits) supports the idea that this protein region in different NF-YB subunits might provide selectivity for different composition and flexibility of GC versus AT sequences flanking the CCAAT nucleotides, thus accounting for the functional differences in gene activation that distinguish members of the L1L/LEC1-type subfamily from other NF-YBs.

An additional difference is observed at the N-terminal parts of At-L1L and At-NF-YC3, both more extended relative to their mammalian counterparts (Figure 1A). In the crystal unit cell, the At-L1L N terminus inserts between the core domain and the N-terminal region of an adjacent At-NF-YC3 molecule (see Supplemental Figure 5A). Formally, it is possible that crystal growth/packing select for the extended conformation of the N termini of both subunits, and that the physiological conformations would be different. The following considerations, however, suggest otherwise. The N terminus of At-NF-YC3 extends differently from the short stretch of NF-YC (Figure 1A), but superimposition with NF-Y from *A. nidulans* (termed Hap; Huber et al., 2012) reveals a similar backbone path (see Supplemental Figure 5B): in HapE (corresponding to NF-YC), this stretch leads to the long αN helix that runs antiparallel to and contacts the α2 helix. Plant NF-YCs show conservation in this area (Huber et al., 2012; see caption of Supplemental Figure 2B), hinting at the presence of an αN, although not to the full extension of HapE. On the At-L1L side, superimposition of the extended N terminus reveals that it is located in the region occupied by the A1A2-linker of NF-YA (and HapB) within the NF-Y/DNA complex (Figure 1F). For this to be physiologically relevant, one would have to assume that the At-NF-YA A1A2-linker follows a different pathway due to specific amino acid composition: indeed, this is an area of relative divergence across kingdoms and paralogs. In mammalian NF-YA, two residues, Gly260 and Pro263, guide the structure of the A1A2 linker by providing the flexibility and directionality, respectively, needed to orient the A2 helix toward the CCAAT box site (Figure 1F). These two residues are absent in plant NF-YAs, including At-NF-YA6. Furthermore, in the At-NF-YA2 homolog, the A1A2 linker has four additional residues, suggesting a specific structuring of this region (see Supplemental Figure 2C). Independently on the structure of the A1A2 linker, the plant helix A2 should be brought in register with that found in the mammalian counterpart to allow correct recognition/binding of the CCAAT box; a Pro residue (Pro205 in At-NF-YA6) is completely conserved in plants, but not in mammalian NF-YA. Importantly, a similar Pro residue is also present in fungi, notably *S. cerevisiae* and *A. nidulans*, where Pro267 indeed contributes to orient the A2 helix of the HapB subunit, despite a different trajectory of the A1A2-linker (Figure 1F). However, EMSA experiments on a Pro205→Gly mutant (Supplemental Figure 1C) indicate that the mutant retains a DNA-binding affinity comparable to that of the wt protein (Supplemental Figure 4B). Thus, our analysis points out that this Pro residue, conserved in plants, may be responsible for favoring the correct functional orientation of the A2 helix toward the DNA in plant NF-YAs, but it is not essential for efficient DNA association and to stabilize the DNA bound complex, indicating a certain degree of structural adaptability of the A1A2-linker region.

In conclusion, At-L1L is a *bona fide* NF-YB subunit that can trimerize and bind to the CCAAT box in a mammalian-like mode. The structural data reported here highlight differences that help explain the molecular mechanisms of LEC1/L1L activity and provide the basis for further genetic and structural analysis of plants trimer/DNA interactions.

ACCESSION NUMBERS

The coordinates of the structure have been deposited in the PDB with the accession code PDB: 5G49.

SUPPLEMENTAL INFORMATION

Supplemental Information is available at *Molecular Plant Online*.

FUNDING

This work was supported by UNIMI funding to N.G.

AUTHOR CONTRIBUTIONS

N.G., A.C.S., and D.S. performed the biochemical experiments; D.S., A.C.S., and M.N. crystallized the protein and solved the X-ray structure; N.G., M.N., and R.M. planned the experiments, analyzed the data, and wrote the manuscript.

ACKNOWLEDGMENTS

No conflict of interest declared.

Received: May 13, 2016
Revised: October 26, 2016
Accepted: November 14, 2016
Published: November 17, 2016

*Nerina Gnesutta, Dana Saad, Antonio Chaves-Sanjuan, Roberto Mantovani and Marco Nardini**

Dipartimento di Bioscienze, Università degli Studi di Milano, Via Celoria 26, 20133 Milan, Italy
***Correspondence: Marco Nardini (marco.nardini@unimi.it)**
http://dx.doi.org/10.1016/j.molp.2016.11.006

REFERENCES

Braybrook, S.A., and Harada, J.J. (2008). LECs go crazy in embryo development. Trends Plant Sci. **13**:624–630.

Calvenzani, V., Testoni, B., Gusmaroli, G., Lorenzo, M., Gnesutta, N., Petroni, K., Mantovani, R., and Tonelli, C. (2012). Interactions and

CCAAT-binding of *Arabidopsis thaliana* NF-Y subunits. PLoS One **7**:e42902.

Dolfini, D., Zambelli, F., Pavesi, G., and Mantovani, R. (2009). A perspective of promoter architecture from the CCAAT box. Cell Cycle **8**:4127–4137.

Huber, E.M., Scharf, D.H., Hortschansky, P., Groll, M., and Brakhage, A.A. (2012). DNA minor groove sensing and widening by the CCAAT-binding complex. Structure **20**:1757–1768.

Kwong, R.W., Bui, A.Q., Lee, H., Kwong, L.W., Fischer, R.L., Goldberg, R.B., and Harada, J.J. (2003). LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development. Plant Cell **15**:5–18.

Laloum, T., De Mita, S., Gamas, P., Baudin, M., and Niebel, A. (2013). CCAAT-box binding transcription factors in plants: Y so many? Trends Plant Sci. **18**:157–166.

Lee, H., Fischer, R.L., Goldberg, R.B., and Harada, J.J. (2003). *Arabidopsis* LEAFY COTYLEDON1 represents a functionally specialized subunit of the CCAAT binding transcription factor. Proc. Natl. Acad. Sci. USA **100**:2152–2156.

Mendes, A., Kelly, A.A., van Erp, H., Shaw, E., Powers, S.J., Kurup, S., and Eastmond, P.J. (2013). bZIP67 regulates the omega-3 fatty acid content of *Arabidopsis* seed oil by activating fatty acid desaturase3. Plant Cell **25**:3104–3116.

Nardini, M., Gnesutta, N., Donati, G., Gatta, R., Forni, C., Fossati, A., Vonrhein, C., Moras, D., Romier, C., Bolognesi, M., et al. (2013). Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. Cell **152**:132–143.

Romier, C., Cocchiarella, F., Mantovani, R., and Moras, D. (2003). The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. J. Biol. Chem. **278**:1336–1345.

Sinha, S., Kim, I.S., Sohn, K.Y., de Crombrugghe, B., and Maity, S.N. (1996). Three classes of mutations in the A subunit of the CCAAT-binding factor CBF delineate functional domains involved in the three-step assembly of the CBF-DNA complex. Mol. Cell. Biol. **16**:328–337.

Yamamoto, A., Kagaya, Y., Toyoshima, R., Kagaya, M., Takeda, S., and Hattori, T. (2009). *Arabidopsis* NF-YB subunits LEC1 and LEC1-LIKE activate transcription by interacting with seed-specific ABRE-binding factors. Plant J. **58**:843–856.

Zemzoumi, K., Frontini, M., Bellorini, M., and Mantovani, R. (1999). NF-Y histone fold alpha1 helices help impart CCAAT specificity. J. Mol. Biol. **286**:327–337.