A number of genes that are up-regulated during cold acclimation have been identified in recent years (Lee and Chen, 1993). The gene *kin1* from *Arabidopsis thaliana* is particularly interesting because it codes for a 6.5-kD polypeptide that bears some compositional similarity to the fish阿拉冰 rich antifreeze proteins (Kurkela and Franck, 1990). This similarity as well as its increased expression during cold acclimation has led to the speculation that the product of *kin1* might be involved in enhancing plant cold tolerance. A homologous gene was more recently reported by Gilmour et al. (1992) as *cor6.6* and by Kurkela and Borg-Franck (1992) as *kin2*.

We are interested in how the expression of stress-related genes is regulated. Therefore, we set out to clone sequences containing the promoters of *kin1* and *cor6.6/kin2*. Using non-radioactive approaches, we identified genomic clones, and a 5.3-kb region was sequenced that contained both protein-coding sequences and extensive sequences 5′ to the transcribed regions (Table 1).

In the 5311-bp genomic fragment, almost 4 kb consists of two homologous sequences (direct repeats) (nucleotides 672–2628 and 2665–4597) with a short spacer (nucleotides 2629–2664). From previous reports, *kin1* (Kurkela and Franck, 1990) can be located in the region 1946 to 2758, and *cor6.6* (Gilmour et al., 1992) lies between 4006 and 4739. There are three mismatches between the sequence presented here and that reported previously (Kurkela and Franck, 1990), which has an A (instead of T) at 2065 and an A (instead of C) at 2562 and misses 2 A’s at 2175 and 2176. The same ecotype (Columbia) was used for both studies. The sequence coinciding with *cor6.6* is identical. A sequence comparison of the transcribed regions was made previously by Gilmour et al. (1992) for *cor6.6*. From the present sequence it is interesting to note that the longest transcribed *kin1* (Kurkela and Franck, 1990) ends at 2743, which is 79 nucleotides into the 5′ of the second homologous sequence. This region thus has homology to the upper 5′ region of *kin1* (nucleotides 672–751).

The 1.4-kb upstream sequences of both genes are highly homologous. The regions proximal to the transcriptional start site contain several potentially significant sequences. First, each contains a duplicated TATA box (TATAAA). The TATA box, conserved among most eukaryotic genes, is involved in the initiation of transcription. It is interesting to note that for maize *Adh1*, another stress-induced gene, duplication of the TATA box had a profound effect on organ-specific expression (Kloeckener-Gruissem et al., 1992). The effect of duplication on the transcription of these two genes remains to be determined. Second, there are three sequences homologous to the consensus G-box element CACGTG (Williams et al., 1992) within 160 bp upstream from the transcription start site, but the A in box I (proximal to the transcription start site) is replaced by a T in *kin1*. The G-box-like elements have been found in many plant promoters (see refs. in Williams et al., 1992) and have affinity for DNA-binding proteins of the

---

**Table 1. Characteristics of a 5.3-kb genomic sequence from *A. thaliana* encoding *kin1* and *cor6.6***

| Organism: | *Arabidopsis thaliana* (L.) Heyn., Columbia ecotype. |
| Location on chromosome: | Unknown. |
| Cloning: | A digoxigenin-labeled probe was produced using PCR with extracted *A. thaliana* genomic DNA as template based on the *kin1* sequence (Kurkela and Franck, 1990). This probe was used to screen *A. thaliana* EMBL3 genomic library, and the positive clones were further verified by PCR. The insert DNA from one clone was subcloned in pGEM3Zf(+). Subclones were identified using PCR with primers based on *kin1* and vector primers. |
| Characteristics of Transcribed Regions: | Purified double-stranded DNA was used directly as template for stepwise sequencing using the Applied Biosystems model 373A and multiple primers. Both strands were sequenced. DNA sequences were analyzed using Eugene software programs (Lark Sequencing, Houston, TX). |
| Characteristics of Putative Promoters: | The 1.4-kb sequences upstream to both genes are highly homologous and contain the following elements of interest: TATA box, direct repeats of TATAAA; G-box, three G-box-like elements; GATA motif, several perfect and imperfect repeats; A/T stretch, both are A/T rich and contain a continuous stretch of A/T of 22 bp. |

---

1 H.W. was supported by Saskatchewan Agriculture Development Fund grant R-89-12-0475. This is National Research Council of Canada Publication No. 37334.

2 Present address: Kuwait Institute for Scientific Research, PO Box 24885, Safat 13109, Kuwait.

* Corresponding author; fax 1–306–975-4839.
basic, Leu zipper (bZIP) type (Katagiri and Chua, 1992). In other promoters, G-boxes occur singly or in two copies (e.g. the wheat Em gene) (Williams et al., 1993). It has also been shown that two slightly different G-box-related sequences confer different expression patterns in transgenic tobacco (Salinas et al., 1993). Third, GATA sequences appear in the putative promoter region of both \( \text{kin}1 \) (1889) and \( \text{cor}6.6 \) (3817), in both cases with several imperfect repeats (G retained but one of the ATA bases altered) nearby. Closely spaced repeats of the GATA motif have been associated with leaf-specific expression and light-induced expression (Gilmartin et al., 1990). The presence of such elements are of interest, because genes associated with low-temperature acclimation would presumably be expressed in leaves and other above-ground tissue. Fourth, the entire 5′ sequences are enriched in A/T residues (68.1 and 67.7%, respectively, for \( \text{kin}1 \) [672–2069] and \( \text{cor}6.6 \) [2665–3854]), and each contains a continuous stretch of 22 bp A/T (starting at positions 1456 and 2650 for \( \text{kin}1 \) and \( \text{cor}6.6 \), respectively). Sequences enriched in A/T are associated with enhanced transcription of downstream genes and with binding to the proteinaceous chromosome scaffold (SARs) (Slatter et al., 1991). In addition, the putative promoter sequence of \( \text{kin}1 \) contains a 7-bp deletion proximal to the transcriptional start site, relative to the corresponding \( \text{cor}6.6 \) sequence.

The two genes presumably arise by a duplication of an ancestral gene. The identity of the ancestral gene cannot be established from the data presented here, because both genes contain deletions of 10 bases or more (starting at positions 1093, 1175, and 1838 in \( \text{kin}1 \) and at position 2880 in \( \text{cor}6.6 \)).

**ACKNOWLEDGMENTS**

We thank Barry Panchuk and Don Schwab for excellent assistance with synthesis of primers and DNA sequencing.

Received August 5, 1993; accepted September 3, 1993.

The GenBank accession number for the sequence reported in this article is L21929.

**LITERATURE CITED**


