

## Update section

### Sequence

# Isolation and nucleotide sequence of a cDNA clone encoding the bread wheat (*Triticum aestivum* L.) CM17 protein

Valérie Lullien, Rémi Alary, Anne Guirao, Philippe Joudrier and Marie-Françoise Gautier  
Laboratoire de Technologie des Céréales, INRA, 2 Place Viala, 34060 Montpellier Cédex 01, France

Received 28 June 1991; accepted 1 July 1991

**Key words:** cDNA, DNA sequence, CM proteins, *Triticum aestivum* L.

A group of proteins from cereal seeds, characterized by their solubility in mixtures of chloroform and methanol (CM), have been found to be members of the trypsin/ $\alpha$ -amylase inhibitor family [3].

In *Triticum aestivum* L. five CM proteins (CM1, CM2, CM3, CM16, CM17) were isolated [3]. They are known to be subunits of a tetrameric heterologous  $\alpha$ -amylase inhibitor [6].

Recently, cDNA clones encoding *T. aestivum* CM1, CM3 and CM16 [2, 7] and cDNA clones encoding *T. durum* CM2 and CM16 [5, 4] were isolated. We present here the sequence of a cDNA clone encoding the bread wheat CM17 protein which only the N-terminal amino acid sequence was already known [1].

We used a durum wheat CM16 cDNA clone (pTd 78) as a probe to screen a  $\lambda$ gt10 *T. aestivum* (cv. Timgalen) cDNA library. From the screening a number of positive clones were detected and subcloned at the *Eco* RI site of the pbluescript II SK vector (Genofit). Sequencing of these clones allowed us to characterize a full-length cDNA clone (pTa 32.1) encoding the bread wheat CM17 protein (Fig. 1). This cDNA contains an open reading frame of 429 nucleotides flanked by a 53

nucleotide 5' untranslated sequence and a 3' non-coding sequence of 172 nucleotides. Two putative polyadenylation signals (ATAAAA) are located 26 and 137 bp upstream from the polyadenylation site. The open reading frame encodes a preprotein of 143 residues of which residues 27–53 are identical to the N-terminal amino acid sequence already determined [1] suggesting that the signal peptide cleavage site of the preprotein occurs after the glycine residue in position 26. However, alignment between the CM17 preprotein sequence and the other CM preproteins shows another possible signal peptide cleavage site after the alanine residue in position 24. Although the two putative signal peptide cleavage sites follow the (-3, -1) rule of Von Heijne [8] the last one is most likely to occur due to homologies with other CM proteins. Depending on the site of cleavage taken into consideration the *Triticum aestivum* mature CM17 protein has a molecular mass of 13275 or 13432 Da.

The amino acid sequence of *T. aestivum* CM17 preprotein is 44.6, 45.3 and 87.4% identical to those of *T. aestivum* CM1, CM3 and CM16 respectively and 47% identical to *T. durum* CM2. The present results not only confirm the predicted

-53  
CGTGGGTGCAAAAGAAAAAACACCAACGAACCTGGCCCTCCATCCAAAAAT

ATGGCGTCCAAGTCCAACACTACAATCTCCTCTTCACGGCCCTCCTAGTCTTCATCTTTGCCGCCGTTGCCGCCGTCGGCAATGAAGATTGC 30  
M A S K S N Y N L L F T A L L V F I F A A V A A V G N E D C

ACCCCATGGACGAGTACTCTGATCACTCCACTCCCAAGCTGCCGTAACATGTGGGAAGAACAGCATGTCGTATCGAAATGCCCGGGCCG 180  
T P W T S T L I T P L P S C R N Y V E E Q A C R I E M P G P

CGGTACCTCGCCAAAGCAGGAGTGTGTGAGCAGCTTGC AAAACATTCGCGACGAGTCCGATGCCAGGCCGCTGCGCTACTTCATGGGGCCG 270  
P Y L A K Q E C C E Q L A N I P Q Q C R C Q A L R Y F M G P

AAGTCTCGCCGGATCAGAGCGGCTCATGGAACCTCCCGGATGCCCTAGGGAGTGCAGATGAACTTCGTCCCAATACTCGTCACTCCG 360  
K S R P D Q S G L M E L P G C P R E V Q M N F V P I L V T P

GGTACTGCAACTTGACGACCGTTCAACAACCCCATACTGCCCTCGGTATGGAGGAGTCTCAGTGGAGCTAGAGACAATTCCTCGCTCA 450  
G Y C N L T T V H N T P Y C L G M E E S Q W S \*

.....  
TGAATAAATAAGCATGTTCGACCATACATGTGTGACATGCATATATACATATAGGACGAGCTCCGGCCGCTCATCATGTGTGTGTCTAT 540  
.....  
CTGCTATATATATGGATAATAAGAATAAAGGGAATCATTTTCGCTTCTTAAAAA

Fig. 1. Nucleotide sequence of *Triticum aestivum* L. cDNA clone pTa32.1 and deduced amino acid sequence of CM17 protein. Numbers flanking the sequence refer to nucleotides, the start of translation is defined as +1. The underlined sequence indicates sequence identical to the N-terminal amino acid sequence determined by Barber *et al.* [1]. Putative polyadenylation signals are marked with dots.

homology for the protein pair CM17–CM16 but also the fact that greater homology is found between proteins encoded by homeologous chromosomes in two different genomes rather than between CM proteins encoded by the same genome [1].

### Acknowledgements

We wish to thank Dr Greg Donovan (CSIRO, North Ryde) for the gift of the  $\lambda$ gt10 library and Jacqueline Asensi for technical assistance.

### References

- Barber D, Sanchez-Monge R, Garcia Olmedo F, Salcedo G, Mendez E: Evolutionary implications of sequential homologies among the members of the trypsin/ $\alpha$ -amylase inhibitor family (CM-proteins) in wheat and barley. *Biochim Biophys Acta* 873: 147–151 (1986).
- Garcia-Maroto F, Marana C, Mena M, Garcia-Olmedo F, Carbonero P: Cloning of cDNA and chromosomal location of genes encoding the three types of subunits of the wheat tetrameric inhibitor of insect  $\alpha$ -amylase. *Plant Mol Biol* 14: 845–853 (1990).
- Garcia-Olmedo F, Salcedo G, Sanchez-Monge R, Gomez L, Royo J, Carbonero P: Plant proteinaceous inhibitors of proteinases and  $\alpha$ -amylases. In: Mifflin B (ed) *Oxford Surveys of Plant Molecular and Cell Biology*, vol 4, pp. 275–334. Oxford University Press, Oxford (1987).
- Gautier MF, Alary R, Joudrier P: Cloning and characterization of a cDNA encoding the wheat (*Triticum durum* Desf.) CM16 protein. *Plant Mol Biol* 14: 313–322 (1990).
- Gautier MF, Alary R, Lullien V, Joudrier P: Nucleotide sequence of a cDNA clone encoding the wheat (*Triticum durum* Desf.) CM2 protein. *Plant Mol Biol* 16: 333–334 (1991).
- Gomez L, Sanchez-Monge R, Garcia-Olmedo F, Salcedo G: Wheat tetrameric inhibitors of insect  $\alpha$ -amylases: Allopolyploid heterosis at the molecular level. *Proc Natl Acad Sci USA* 86: 3242–3246 (1989).
- Lullien V, Alary R, Joudrier P, Gautier MF: Characterization of a cDNA clone encoding the *Triticum aestivum* L. CM16 protein: homology with the *Triticum durum* Desf. sequence. *Plant Mol Biol* 16: 373–374 (1991).
- Von Heijne G: A new method for predicting signal sequence cleavage sites. *Nucl Acids Res* 14: 4683–4690 (1986).