Sequence of a 21 kd zein gene from maize containing an in-frame stop codon

Christine Wandelt[1] and Guenter Feix

Institute of Biology III, Schaenzlestrasse 1, D-7800 Freiburg, FRG and [1]CSIRO, Division of Plant Industry, GPO Box 1600, Black Mountain, Canberra, ACT 2601, Australia
Submitted February 16, 1989 EMBL accession nos X14334, X14335

Zein proteins of the 19 and 21 kd size classes are encoded for by a multigenic system with more than 100 members (1). So far, only a few genes and their flanking sequences have been isolated from genomic libraries and characterized by sequencing. The protein coding parts of these genes are free of intron structures. In the case of zA1, a gene of the 21 kd class, an in frame stop codon was found (2). Here, we report the sequence of pML1, a 21 kd gene of the variety A619, which codes for a 266 amino acid long protein which contains an in frame stop codon at position 397 to 399. The 5' flanking sequence of pML1 had previously been shown to be intact and functional (3). While in frame stop codons have also been found in some; other storage protein genes (4), it is open at this moment if these stop codons are functional as terminator elements. The 3' flanking sequence of pML1 shown here exhibits two putative poly(A) addition signals at positions 859 to 863 and 920 to 925 and reveals, furthermore, a 19 bp element at position 1042 to 1060 which is also present at this location in pMS2 (a gene of the 19 kd zein gene class). The sequence of pML1 and pML2 (another 21 kd gene of the variety A619) has been submitted to the EMBL data bank (Accession number: X14334 and X14335).

```
     MET A T K V L S L L A L L A L F V S A T N G F I I P Q C S L A P S A
   1 ATGGCTACCAAGGTATTATCCCTCCTTGCGCTTCTTGCCCTTTTTGTGAGCGCAACAAATGGGTTCATTATTCCACAATGCTCACTTGCTCCTAGTGCCA
     I I P K F L P P V T S MET G F E H P A V Q A Y R L Q Q A L A A S V L Q
 101 TTATTCCAAAGTTCCTCCCACCAGTTACTTCAATGGGCTTCGAACACCCAGCTGTGCAAGCCTACAGGCTACAACAAGCGCTTGCGGCGAGCGTCTTACA
       Q P I A Q L Q Q Q S L A H L T I Q T I A T Q Q Q Q Q F L P A L S H
 201 ACAACCAATTGCCCAATTACAACAACAATCCTTGGCACATCTAACAATACAAACCATCGCAACGCAACAGCAACAACAGTTCCTACCAGCACTGAGCCAC
       L A MET V N P A A Y L Q Q Q L L A S N A L A L A N V V A N Q Q Q Q *
 301 CTAGCCATGGTGAACCCTGCCGCCTACTTGCAACAACAACTGCTTGCATCCAACGCACTTGCTCTGGCAAACGTAGTTGCAAACCAGCAACAACAATAGC
       L Q Q F L P A L S Q L A MET V N P A A Y L Q Q Q Q L L S S S P L A V A
 401 TGCAACAGTTTCTGCCAGCGCTCAGTCAACTAGCCATGGTGAACCCTGCCGCCTACCTACAACAGCAACAACTGCTTTCATCTAGCCCGCTTGCTGTGGC
       N A P T Y L Q Q Q L L Q H I V P A L T Q L V V A N P A A Y L Q Q L
 501 CAATGCACCTACATACCTACAACAACAATTGTTGCAACATATTGTACCAGCTCTGACTCAGCTAGTTGTGGCAAACCCTGCTGCCTACTTGCAACAGCTG
       L P F N Q L T V S N S A A Y L Q Q R Q Q L L N P L A V A N P L V A
 601 CTTCCATTCAACCAACTGACTGTGTCGAACTCTGCTGCGTACCTACAACAGCGACAACAGTTACTTAATCCACTTGCAGTGGCTAACCCATTGGTCGCTG
       A F L Q Q Q Q L L P Y N Q F S L MET N P V L S R Q Q P I V G G A I F
 701 CCTTCCTACAGCAGCAACAATTGTTGCCATACAACCAGTTCTCTTTGATGAACCCTGTCTTGTCGAGGCAGCAACCCATCGTTGGAGGTGCCATCTTTTA
 801 GATAACATATGAGATGTACTCGATAATGGTGCCCTCATACCGGCATGTGTTTCCTAGAAATAACCAATATATTGATTGAGATTTATCTCGATATATTTCT
 901 GAACTATGTCCATCATATAAATAATTGAAAACATCAAATCGTAATTATAAACTCATGCTTGGCCAATACATAGACAATGCGATATTACTTCAACATCCCA
1001 ATGATGTCCTAGCCCAACCTCTTGAATGTTAGTGTTTTAATGTGTGAGGGTATATTTATAAGATAGACGTGACTATTTGCGCTTTTTCAGTACATACACA
1101 TATGGTATGTTGATTTGATATAGTGATGGACACATGCTTTGACCATGAATATTCAAATCAGTTGTACTTGCACGAAGCAAAACATAATATAAGTTTAGGA
1201 GTACACTTATAACTGTATCCAATCATGCTCACACAAATTCATACCACATTATAATTTTTTGGTAAATATTCCACACATGTACTTTTTACAAGAACCCAAA
1301 TTTTGCAAACAAATGTAGCATTGTAGACATGTAG
```

Acknowledgement: Financial support from the Deutsche Forschungsgemeinschaft and the Fond der Chemischen Industrie is gratefully acknowledged.

References:

1. Heidecker, G. and Messing, J. (1986) Ann. Rev. Plant Physiol. 37, 439-466
2. Spena, A., Viotti, A., Pirotta, V. (1983) J. Mol. Biol. 169, 799-811
3. Langridge, P. and Feix, G. (1983) Cell 34, 1015-1022
4. Entwistle, J. (1988) Carlsberg Res. Commun. 53, 247-258

 © IRL Press