Volume 15 Number 22 1987 Nucleic Acids Research

The primary structure of the predominating vicilin storage protein subunit from field bean seeds (*Vicia faba* L. var. minor cv. Fribo)

Ronald Bassüner*, Nong Van Hai[1], Rudolf Jung, Gerhard Saalbach and Klaus Müntz

Akademie der Wissenschaften der DDR, Zentralinstitut für Genetik und Kulturpflanzenforschung, DDR-4325 Gatersleben, GDR and [1]National Scientific Research Centre, Institute of Biology, Hanoi, Socialist Republic of Vietnam

Submitted September 25, 1987 Accession no.Y00462

Vicilin, the 7S seed globulin of field bean (*Vicia faba* L.) represents approximately 30% of the storage protein in mature seeds (1). Two types of subunits with M_r ~70kD and ~50kD seem to form the trimeric quarternary structure of the holoprotein (2). The 1561bp cDNA insert of clone pVi91 (3) was sequenced using the solid phase variant (4) of the chemical degradation procedure . The insert has one long open reading frame coding for the 463 amino acids (M=52,700) of the most prominent field bean previcilin subunit:

```
    1 TTAGACTTTAATTCAAATCAACATGGCAGCTACCACATTGAAAGATTCATTTCCACTTTTGACGTTGTTGGGAATCGCTTTCCTAGCTTCAGTGTGTCTTTCT
                             M  A  A  T  T  L  K  D  S  F  P  L  L  T  L  L  G  I  A  F  L  A  S  V  C  L  S↑
  104 TCTAGATCCGATCAAGATAACCCTTTTGTCTTTGAGTCTAACAGGTTTCAAACTCTTTTTGAGAATGAAAACGGTCACATTCGTCTTCTCCAAAAATTTGATCAA
       S  R  S  D  Q  D  N  P  F  V  F  E  S  N  R  F  Q  T  L  F  E  N  E  N  G  H  I  R  L  L  Q  K  F  D  Q
  209 CATTCCAAACTACTTGAGAATCTCCAAAATTACCGTCTTTTGGAATATAAATCTAAACCTCACACCATATTTCTTCCACAACAAACTGACGCCGATTTCATCCTT
       H  S  K  L  L  E  N  L  Q  N  Y  R  L  L  E  Y  K  S  K  P  H  T  I  F  L  P  Q  Q  T  D  A  D  F  I  L
  314 GTTGTCCTTAGTGGAAAAGCTATACTCACAGTGTTATTACCCAACGACAGAAACTCCTTCAGTCTTGAGCGAGGAGATACCATAAAACTTCCTGCTGGTACAATT
       V  V  L  S  G  K  A  I  L  T  V  L  L  P  N  D  R  N  S  F  S  L  E  R  G  D  T  I  K  L  P  A  G  T  I
  419 GGTTATTTAGTTAACCGAGATGACGAGGAGGATCTTAGAGTACTAGATCTTGTCATCCCCGTAAATAGACCAGGTGAACCTCAGTCTTTCTTATTGTCTGGTAAT
       G  Y  L  V  N  R  D  D  E  E  D  L  R  V  L  D  L  V  I  P  V  N  R  P  G  E  P  Q  S  F  L  L  S  G  N
  524 CAAAACCAACCGTCCATCTTATCTGGATTCAGCAAGAACATTTTAGAGGCTTCCTTCAATACCGATTATAAAGAGATAGAGAAGGTTCTCTTAGAAGAGCATGGG
       Q  N  Q  P  S  I  L  S  G  F  S  K  N  I  L  E  A  S  F  N  T  D  Y  K  E  I  E  K  V  L  L  E  E  H  G
  629 AAAGAGAAATATCACAGGAGAGGCCTTAAAGATAGGAGGCAGCGGGGCCAAGAAGAAAATGTAATAGTCAAAATATCAAGGAAACAAATTGAGGAATTGAACAAA
       K  E  K  Y  H  R  R  G  L  K  D  R  R  Q  R  G  Q  E  E  N  V  I  V  K  I  S  R  K  Q  I  E  E  L  N  K
  734 AATGCAAAGTCTAGCTCCAAAAAGAGTACATCTTCTGAATCTGAACCGTTCAATTTGAAAAGTCGTGAGCCCATCTATTCCAACAAGTTTGGTAAATTTTTTGAG
       N  A  K  S  S  S  K  K  S  T  S  S  E  S  E  P  F  N  L  K  S  R  E  P  I  Y  S  N  K  F  G  K  F  F  E
  839 ATCACCCCAAAGAGAAATCCACAACTTCAAGACTTGAATATATTTGTCAATTATGTGGAAATTAATGAGGGATCTCTTTTGTTGCCACACTACAATTCAAGGGCC
       I  T  P  K  R  N  P  Q  L  Q  D  L  N  I  F  V  N  Y  V  E  I  N  E  G  S  L  L  L  P  H  Y  N  S  R  A
  944 ATAGTGATAGTAACAGTCAATGAAGGAAAAGGGGATTTTGAACTTGTGGGTCAAAGAAATGAGAACCAACAGGGCCTGAGAGAAGAATATGACGAGGAAAAGGAA
       I  V  I  V  T  V  N  E  G  K  G  D  F  E  L  V  G  Q  R  N  E  N  Q  Q  G  L  R  E  E  Y  D  E  E  K  E
 1049 CAAGGAGAAGAAGAGATAAGGAAACAAGTGCAAAATTACAAAGCTAAGTTGTCTCCAGGTGATGTTTTAGTAATTCCAGCGGGTTATCCAGTTGCCATAAAAGCT
       Q  G  E  E  E  I  R  K  Q  V  Q  N  Y  K  A  K  L  S  P  G  D  V  L  V  I  P  A  G  Y  P  V  A  I  K  A
 1154 TCATCAAATCTCAATTTGGTGGGATTTGGCATCAATGCTGAGAACAACCAGAGATACTTCCTTGCAGGTGAGGAAGACAATGTGATAAGTCAGATACATAAACCA
       S  S  N  L  N  L  V  G  F  G  I  N  A  E  N  N  Q  R  Y  F  L  A  G  E  E  D  N  V  I  S  Q  I  H  K  P
 1259 GTGAAAGAACTTGCATTCCCTGGATCAGCTCAAGAGGTAGATACGCTACTAGAGAATCAGAAACAATCTCACTTTGCAAATGCTCAACCTCAAGAAAGAGAAAGA
       V  K  E  L  A  F  P  G  S  A  Q  E  V  D  T  L  L  E  N  Q  K  Q  S  H  F  A  N  A  Q  P  Q  E  R  E  R
 1364 GGAAGCCAAGAAATAAAGGATCATCTGTACTCAATTTTGGGTTCTTTTTAATGAGTGATCAAATGTTATGTATGTATGCTCTAAAGTACTATAGCTCATAGTGAG
       G  S  Q  E  I  K  D  H  L  Y  S  I  L  G  S  F  -
 1469 CAAATAATAAAACGACTCTCTCTTGTAACTATAATTATAACTCCACCTTTCTACTAAGAATAAATAAATTAAGAATGTTTGTGTGCAAAAAAA
```

Multiple and overlapping polyadenylation signals with plant-specific features (5) are underlined.

Previous experiments on cotranslational processing of this abundant vicilin primary translation product revealed the presence of a transient N-terminal signal peptide (3). Since serine is the predominating N-terminal amino acid of the purified holoprotein (6) the most probable signal peptide cleavage point is, in agreement with an algorithm of prediction (7) between amino acids 27 and 28, (see arrow) resulting in a mature vicilin subunit of 436 amino acids (M=49,900). The mature subunit, lacking methionyl, cysteinyl, and tryptophanyl residues, is nutritionally very imbalanced. No potential glycosylation site is present in contrast to most of the homologous vicilin-like subunits from other species (8) so far investigated.

Acknowledgement: We thank M. Wiesner and G. Jüttner for skilled technical assistance.

*To whom correspondence should be addressed

References: 1. Derbyshire E. et al. (1976) Phytochemistry *15*, 3-24. 2. Müntz K. et al. (1986) Biol. Zbl. *105*, 107-120. 3. Bassüner R. et al. (1984) FEBS Lett. *166*, 314-320. 4. Rosenthal A. et al. (1986) Gene *42*, 1-9. 5. Lycett G.W. et al. (1983) FEBS Lett.*153*, 43-46. 6. Scholz G. et al. (1983) Eur. J. Biochem. 132, 103-107. 7. von Heijne G. (1986) Nucl. Acids Res. *14*, 4683-4690. 8. Doyle J.J. et al. (1986) J. Biol. Chem. *261*, 9228-9238.