Sequence of a cDNA from *Chlamydomonas reinhardii* encoding a ubiquitin 52 amino acid extension protein

J.Callis, L.Pollmann[1], J.Shanklin, M.Wettern[1] and R.D.Vierstra*

Department of Horticulture, University of Wisconsin–Madison, Madison, WI 53706, USA and [1]Botanisches Institut Technische Universitaet, Humboldtstrasse 1, D-3300 Braunscheig, FRG
Submitted September 1, 1989 EMBL accession no. X15427

Ubiquitin is one of the most highly conserved proteins known. To determine whether algal ubiquitin is identical to higher plant ubiquitin, three ubiquitin cDNAs were identified from a *Chlamydomonas reinhardii* cDNA expression library screened with human ubiquitin antibodies. The cDNAs encode a 76 amino acid (aa) ubiquitin monomer with an extension protein of 52 aa. The DNA and derived aa sequence of one is shown below and the aa sequence is compared to ubiquitin 52 aa extension proteins from *Arabidopsis* (1), yeast (2) and human (3). Arrowhead marks the ubiquitin extension junction and dashes denote aa identity. *Chlamydomonas* ubiquitin has only 1 aa substitution from the higher plant sequence (position 24) and 2 from the animal sequence (positions 19 and 57). The *Chlamydomonas* 52 aa extension protein is also highly conserved, being 86% and 81% identical to its counterparts from *Arabidopsis* and yeast, respectively. Both the positions of 4 cysteine residues (circles) and the nuclear localization signal (underlined) are conserved among extensions.

```
   Chlamyd. GGGCAGCCATGCAAATCTTCGTGAAGACCCTCACGGGTAAGACCATCACCCTCGAGGTGGAGTCTTCGGACACCATCGAG  80
   Chlamyd.         MetGlnIlePheValLysThrLeuThrGlyLysThrIleThrLeuGluValGluSerSerAspThrIleGlu
Arabidopsis         ------------------------------------------------------------------------Asp
      yeast         ------------------------------------------------------------------------Asp
      human         ------------------------------------------------------------Pro-----------
   Chlamyd. AATGTGAAGGCCAAGATTCAGGACAAGGAGGGCATCCCTCCGGACCAGCAGCGTCTGATCTTCGCCGGCAAGCAGCTGGA 160
   Chlamyd. AsnValLysAlaLysIleGlnAspLysGluGlyIleProProAspGlnGlnArgLeuIlePheAlaGlyLysGlnLeuGl
Arabidopsis ------------------------------------------------------------------------------
      yeast ---------Ser---------------------------------------------------------------------
      human ------------------------------------------------------------------------------
   Chlamyd. GGACGGCCGCACCCTGGCGGACTACAACATCCAGAAGGAGTCCACCCTCCACCTGGTGCTGCGTCTGCGCGGTGGTATCA 240
   Chlamyd. uAspGlyArgThrLeuAlaAspTyrAsnIleGlnLysGluSerThrLeuHisLeuValLeuArgLeuArgGlyGlyIleI
Arabidopsis ------------------------------------------------------------------------------
      yeast ----------------Ser-------------------------------------------------------------
      human ----------------Ser-------------------------------------------------------------
   Chlamyd. TTGAGCCCTCGCTGCAGGCCCTGGCCCGCAAGTACAACCAGGAGAAGATGATCTGCCGCAAGTGCTACGCCCGCCTGCAC 320
   Chlamyd. leGluProSerLeuGlnAlaLeuAlaArgLysTyrAsnGlnGluLysMetIleCysArgLysCysTyrAlaArgLeuHis
Arabidopsis -------------MetMet--------------------Asp-------------------------------------
      yeast -------------Lys---------Ser---------CysAsp---SerVal----------------------Pro
      human -------------ArgGln------Gln---------CysAsp-----------------------------------
   Chlamyd. CCGCGCGCCAAGAACTGCCGCAAGAAGTCGTGCGGCCACACCAACCAGCTGCGCCCCAAGAAGAAGCTCAAGTAAACCTT 400
   Chlamyd. ProArgAlaLysAsnCysArgLysLysSerCysGlyHisThrAsnGlnLeuArgProLysLysLysLeuLysEnd
Arabidopsis ---------Val-------------Lys---------Ser----------------------Ile------
      yeast ---------Thr---------------Lys---------------Asn--------------------------
      human ---Val---Val---------------Lys-----------------------------------Val------
            GGGCCCACCTCTTGGCGCCCTGGCACACGCGGGCGAGCCGCCCCGCGCCCGCGTGGCCTGCGGCAGCGGCGGCAGCAGCA 480
            GCGGCTGCGGCGGGAGGCGACGACGGCGGACGGTCCTCTCGCGGGCCGCTGCGCTTTTTTGATGCTCCGCTGTGGATGGC 560
            GCGGTGACCCCCCGGCGAGCTCTCCACTCGCCGGGGACTAGCACACGGGTCTCTGTTCTTCCGAGCGCGCGGGGTGACCC 640
            TCCGCTGTGGGACAGGAGCATCTGCGCTTTGTAACAGAAGCAGGAGTCTAAAAAAAAAAAAA                   702
```

Acknowledgements

This work was supported by USDA-CRGO grant 88-37262-3368 to R.D. Vierstra, an NSF Post-doctoral Fellowship in Plant Biology to J.C. and a German Academic Exchange Service (DAAD) grant to L.P.

*To whom correspondence should be addressed

References

1. Callis, J., Raash, J. and Vierstra, R., manuscript in preparation.
2. Ozkaynak, E. *et al.* (1987) EMBO J. 6,1429-1439.
3. Salvesen, G., Lloyd, C., and Farley, D. (1987) Nucleic Acids Res. 15, 5485.

©IRL Press