

The sequence of the *Kluyveromyces lactis* BiP gene

M.J.Lewis and H.R.B.Pelham

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Submitted September 24, 1990

EMBL accession no. X54709

The BiP gene has been shown to encode a member of the HSP 70 family which is a resident of the endoplasmic reticulum (E.R.) lumen and plays a role in protein folding and oligomerisation (2, 3). In *S. cerevisiae* BiP has been shown to be the product of the *KAR2* gene, which is required for nuclear fusion (4).

We have recently reported the cloning of the *Kluyveromyces lactis* homologue of the BiP gene (1), in order to identify the C-terminal E.R. retention signal. We present here the complete sequence and deduced amino acid sequence.

The gene was cloned from a mini library of *K. lactis* genomic DNA by low stringency hybridisation to a probe derived from the *S. cerevisiae* gene and encodes the features expected of the BiP gene, including a hydrophobic leader sequence and an E.R. retention signal, which in the case of *K. lactis* has proved to be

Asp-Asp-Glu-Leu as opposed to the His-Asp-Glu-Leu used in *S. cerevisiae*. The sequence also encodes a potential N-linked glycosylation site at amino acid 618, a feature not previously noted in BiP sequences from other species. It is not known whether this site is used *in vivo*. The deduced protein sequence is 77.3% identical to that of *S. cerevisiae* BiP and 61% to that of rat BiP.

REFERENCES

- Lewis, M.J., Sweet, D.J. and Pelham, H.R.B. (1990) *Cell* **61**, 1359–1363.
- Munro, S. and Pelham, H.R.B. (1986) *Cell* **46**, 291–300.
- Pelham, H.R.B. (1989) *Ann. Rev. Cell Biol.* **5**, 1–23.
- Rose, M.D., Misra, L.M. and Vogel, J.P. (1989) *Cell* **57**, 1211–1271.

```

AGGCGGTATAATGGACACA TCACCGTTAGGAACAACAT ATCTCATTCTGTGGGACAT TGTCTGCTAAGATCTCAA AGAAAAAAAAAGAAATAA TGGTTTACGTGTCAAATGA 120
CTTTCTTCTGTCTTCTTC TTCGGTGAATTTTCATCGAT TTTCGAGGAAAATCTTGGTG GAAATATGAAGAAACACGTA TATATCACAGTTACGTATAT ATATAAAGGACAGGTATAAT 240
TTCCAAAACCTATAATAATT CTCTCGTATTATTAAGTAT AGACACATAAGATTCATITT TATTTCTTGTTCGTTCAARAT TTTGAAGGTTGATAAAACTA TCAGTGTTTTACTCAGAAAG 360
CAAAAATCAATCAATCCAAGT CTGGTTCAGTTTTTACTAGT TTCAGGTTAGTAATAGTAAG ATCAACCAAGATCATTATG TTTTCTGCCAGGAAAAGTTC AGTTGATGGCTAGTGTGCT 480
CACTAGCCGGTGTTTATGTG CTTTGTGGCCGTTATAATGCC ATTTGGCTTAAGGGGCTCTC AGTCCCTGGAGTGTGAGCT GGTGGTGTCTGAAGTCACGA AGATATGTTACCGTATTTG 600
GTATCGATTTGGTACGACA TATCTTCTGTGTGTGAT GAATAATGTTAAACTGTAAA TCTTAGCCAATGAGCAAGGT AACAGAAATTAATCTCATCTA CGTTTCATTCACTCAGCATG 720
AAGATTGATCGGTGATGCT GCTAAGAACAAGCTGATCC TAATCCAAAAGAACACCATTT TTGATATCAAGAGATTGATT GGCTTACAATACACACGATCC AACGGTTCAAAAGATATCA 840
AGCAATTCGCTATACCTG GTTAAACAAGGGCAATAAGCC ATATGTTAGATTACTGTGA AGGGTGAAGAAGAAAGTTC ACTCCAGAAGAAGTCTCTGG TATGATTTTGGTAAAGTGA 960
AGCAATTCGCTATACCTG TGGGTAAGAAGATTACTCA TCCGGTGTCACTGTACCTG CTTACTTTAAGCATGCTCAA AGACAAGCCACTAAGGACGC AGGTGCCATTGCTGGTTTGA 1080
ACATTTTGAGAAATCGTTAAC GAACCTACCGCTGACGCCAT TCGTTATGGTTTGAATAAGA CAGAAGACGAACACCAATTT ATTTGCTATGATTTGGGTGG TGGTACCTTCGATGATCTT 1200
TGTGTCTATTGAAAATGGT GTCCTTGAAGTTCAGCTAC CGCTGGTGAACACTCATTAG GTGGTGAAGATTTGCACTAC AAATTAGTCCGCTCACTTGC TCGACTGTTCCAAAAGAAGC 1320
ATGACTTGGAGCTCACCAAG AATGATAAGCCATGGCTAA GTTAAABAAGAAAGCTGTAAA AAGGCTAGCTTCTTGTGCT GGTCAAACTTCTACCGTAT TGAATATGACTCTTCTTCA 1440
AGCATATTGATCGATCGTCT TAAGTGGTGGTCTCAAGC AATTCAAAAGGTTCAACAAC TATTGGAGAAGTCTTTAAC GGTAAAGAGGCTTCAAGGG TATCAACCAGATGAAGCCG 1680
TTGCATTCGCTCAAGAAAC ACTGCAATTCACAACAAAAA ATCTCAAAATTTCTCTACTG CAAGTGTGATTCGAGTAAAT GCCTTGACTTTGGGTATTGA AACCACTGGTGTGTCAATGA 1800
AGGATAACAACCTTATGGGT AATTTGAGCTAAGTATAT TTAGAGCAGCCCAAGAGGTT TCCACAANATGAGGTCACA TTTGCTTTGATGCTAATGG TATCTTAAGCCGTTTGGCCA 2040
CCGATAAAGCACTCTGTAAG TCCGAATCTATTACATTTC CAATGACAAAGGTTAGATTGT CTCAGGATGATTCGACAGA ATGGTCAAAAGCTGTAAA GTATGAGCTGAAGTGCCA 2160
AATTCAGGCTAAGAGCAA CCTGTACACCAATTCBAAA CTTTCTCTACTCTGGAAGA TTTCTGTTAAACGGTGAATG CGCGAAATCATGGTGAAGA TGACAAGAAACCGTTTATG 2280
ACAATGCTCAATGAATCCCTA GAATGTTGGAAGAACAATC TATGTGCGCTGAAGCTGAAG ATTTGGAAGAAGATGGCT TCATTCAAGAAATCGTGA GAACAAATTTGGTCAATG 2400
CTGCATCTCAAGGCTCTACT TCAGGAGAGGGAATTCGAAGA TGAAGATGACGACGACTATT TTGATGACGAATATATCA CCTTCAATCTCCGAAACA ATGGAGACTATAAATCAA 2520
CTGAAAAGAAAAAATAAC GAATTCACAATGCTTGTGTT TCTCCGCTCTGCGAATGAA GTCACATCCGAAAGCAAAAC TTTAGTTTACTTTGCAATCT TTCAAATTCAAAATCTGAT 2640
TGCTCTGCGCATGTCAATTC TATATTTGACAAGCTCTGAT GCTGACCCGATGTCTGGAA CAACACAATCGCGAAAGTT CCATTGACTTGTACGCAACG AGAAGAATAAAGCAGCAAAA 2788
AAAAAAAAAAGAAAAGTTC TATTTTTAA
    
```

```

60
MFSARKSSVGLVSSLAFFVLLAVIMPALTGSQSRVVARAEHDEHYTVIGIDLGT
** * * * * *
MFFNRLSAGLLVPLSVLYALFVILPLQ-NSFHSNVLVRGADDVENVYTVIGIDLGT
120
TYSVAVMKGKTEILANEQGNRITPSVSVFTDDERLIGDAAKQVAASNPKNITFDIKRL
*****
TYSVAVMKGKTEILANEQGNRITPSVVAFTDDERLIGDAAKQVAANPQNTIFDIKRL
180
IGLQYNDPTVQRDIKHLFYTNNKGNKPVVEVTVKEKKEFTPEEVSQMLGKMQIAED
* * * * *
IGLYNDRSVQRDIKHLFNVNKGDKPAVEVSVKKEKFTPEEISQMLGKMQIAED
240
YLGKVTHAVVTVPAVFNDQRQATKDAGIAGLNLRLVNEPTAAAIAYGLDKTEDEHQ
*****
YLGKVTHAVVTVPAVFNDQRQATKDAGIAGLNLRLVNEPTAAAIAYGLDKSDEHQ
300
IIVYDLGGGTFDVSLLSIENGVFEVQATGDTHLGGEDFDYKLVRFHQAFLFKKHDLVTF
*****
IIVYDLGGGTFDVSLLSIENGVFEVQATSGDTHLGGEDFDYKIVRQLIKAFKKKHIDVUS
360
KNDKAMAKLKREAEKAKRSLSSQSTRIEIDSPFNGIDFSETLTKRFEELNLFKFTL
* * * * *
DNNKALAKLKREAEKAKRSLSSQSTRIEIDSPFNGIDFSETLTKRFEELNLFKFTL
420
KPVEKVLKDSGLQKEDIIDVIVGGSTRIKPVQQLLEKFFKGGKASKGINPDEAVYGA
*****
KPVEKVLQDSGLKEDVDVIVGGSTRIKPVQQLLESYDFKGGKASKGINPDEAVYGA
480
VQAGVLSGEGVEDIVLLDNALTLGIETGGVMTPLIKRNTAIPTKKSQIFSTAVDNQK
*****
VQAGVLSGEGVEDIVLLDNALTLGIETGGVMTPLIKRNTAIPTKKSQIFSTAVDNQK
540
AVRIQVYGERAMVDKNNLGNFELSDIRAAPRGVQEVTFALDANGILTVSATDKDTG
* * * * *
TVMIKVYGERAMSKDNNLNGKFEITGIPAPRGVQEVTFALDANGILTVSATDKDTG
600
KSESIITANDKGRLSQDDIDRMVEAEKFAEDAKFKAKSEARNTFENFVHYKNSVNGE
*****
KSESIITANDKGRLTQEDIDRMVEAEKFAEDASIKAKVESRNLKENYAHSLKQVSGS
660
LAEIMDEDDKETLVDNWNESLEWLEDNSVAEAEDEFKMAFPKESVPEILAKASQGS
* * * * *
LGEKLEEDKETLDAANDVLEWLDNDFETAIEDDFEFESLSKVAYITSKLYGGADG
720
TSGEG-F--EDED--DYFD-DEL K. lactis
*****
-SGAADYDDEDEDDGDFEHDEL S. cerevisiae
    
```